Advanced Information and Knowledge Processing

Series Editors
Professor Lakhmi Jain
lakhmi.jain@unisa.edu.au

Professor Xindong Wu
xwu@cs.uvm.edu

Also in this series

Gregoris Mentzas, Dimitris Apostolou, Andreas Abecker and Ron Young
Knowledge Asset Management
1-85233-583-1

Michalis Vazirgiannis, Maria Halkidi and Dimitrios Gunopulos
Uncertainty Handling and Quality Assessment in Data Mining
1-85233-655-2

Asunción Gómez-Pérez, Mariano Fernández-López and Oscar Corcho
Ontological Engineering
1-85233-551-3

Arno Scharl (Ed.)
Environmental Online Communication
1-85233-783-4

Shichao Zhang, Chengqi Zhang and Xindong Wu
Knowledge Discovery in Multiple Databases
1-85233-703-6

Jason T.L. Wang, Mohammed J. Zaki, Hannu T.T. Toivonen and Dennis Shasha (Eds)
Data Mining in Bioinformatics
1-85233-671-4

C.C. Ko, Ben M. Chen and Jianping Chen
Creating Web-based Laboratories
1-85233-837-7

Manuel Graña, Richard Duro, Alicia d'Anjou and Paul P. Wang (Eds)
Information Processing with Evolutionary Algorithms
1-85233-886-0

Colin Fyfe
Hebbian Learning and Negative Feedback Networks
1-85233-883-0

Yun-Heh Chen-Burger and Dave Robertson
Automating Business Modelling
1-85233-835-0

Dirk Husmeier, Richard Dybowski and Stephen Roberts (Eds)
Probabilistic Modeling in Bioinformatics and Medical Informatics
1-85233-778-8

Ajith Abraham, Lakhmi Jain and Robert Goldberg (Eds)
Evolutionary Multiobjective Optimization
1-85233-787-7

K.C. Tan, E.F.Khor and T.H. Lee
Multiobjective Evolutionary Algorithms and Applications
1-85233-836-9

Nikhil R. Pal and Lakhmi Jain (Eds)
Advanced Techniques in Knowledge Discovery and Data Mining
1-85233-867-9

Amit Konar and Lakhmi Jain
Cognitive Engineering
1-85233-975-6

Miroslav Kárný (Ed.)
Optimized Bayesian Dynamic Advising
1-85233-928-4

Marcus A. Maloof (Ed.)
Machine Learning and Data Mining for Computer Security
1-84628-029-X

Yannis Manolopoulos, Alexandros Nanopoulos, Apostolos N. Papadopoulos and Yannis Theodoridis
R-trees: Theory and Applications
1-85233-977-2

Sanghamitra Bandyopadhyay, Ujjwal Maulik, Lawrence B. Holder and Diane J. Cook (Eds)
Advanced Methods for Knowledge Discovery from Complex Data
1-85233-989-6

Sifeng Liu and Yi Lin

Grey Information

Theory and Practical Applications

With 60 Figures

Springer

Sifeng Liu, PhD
College of Economics and Management
Nanjang University of Aeronautics and Astronautics
Nanjing, 210016
CHINA

Yi Lin, PhD
Department of Mathematics
Slippery Rock University
Slippery Rock, PA 16057
USA

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2005924711

Advanced Information and Knowledge Processing ISSN 1610-3947
ISBN-10: 1-85233-995-0
ISBN-13: 978-185233-995-1

Printed on acid-free paper

Printed in the United States of America (MVY)

9 8 7 6 5 4 3 2 1

Springer Science+Business Media
springer.com

Preface

Rapid formation and development of new theories of systems science have become an important part of modern science and technology. For example, since the 1940s, there have appeared systems theory, information theory, fuzzy mathematics, cybernetics, dissipative structures, synergetics, catastrophe theory, chaos theory, bifurcations, ultra circulations, dynamics, and many other systems theories. Grey systems theory is also one of such systems theories that appeared initially in the 1980s.

When the research of systems science and the method and technology of systems engineering are applied in various traditional disciplines, such as management science, decision science, and various scientific disciplines, a whole new group of new results and breakthroughs are obtained. Such a historical background has provided the environment and soil for grey systems theory to form and to develop rapidly in the past 20-plus years.

More specifically, in 1982, Professor Deng Ju-Long published the first research paper in the area of grey systems in the international journal entitled *Systems and Control Letters*, published by North-Holland Co. His paper was titled "Control Problems of Grey Systems." The publication of this paper signalled the birth of grey systems theory after many years of effective research of the founding father. This new theory soon caught the attention of the international academic community and practitioners of science. Many well-known scholars, such as Chinese academicians Qian Xueshen, Song Jian, and Zhang Zhongjun. Professor Roger W. Brockett of Harvard University, a former editor-in-chief of the journal *Systems and Control Letters*, and several formal Soviet academicians, all provided very positive comments on this new theory and offered their support.

In the short time period of about two decades, the theory of grey systems has been developed and is maturing rapidly. It has been widely applied to analyses, modeling, predictions, decision making, and control, with significant consequences, of various systems, including, but not limited to, social, economic, scientific and technological, agricultural, industrial, transportation, mechanical, petrological, meteorological, ecological, hydrological, geological, financial, medical, legal, military, etc., systems. Research papers on grey systems have been cited by many scholars around the globe and been reviewed by internationally authoritative review periodicals. Currently, eighty-some universities worldwide, located in countries such as Australia, China, Japan, Taiwan, and the United States of America, have offered courses or workshops on grey systems, and hundreds of graduate students are applying the methodology of grey systems in their research and their writing of dissertations. There have been many international conferences listing grey systems as a special topic. All of these represent the fact that grey systems theory with its strong vitality has already stood in the forest of scientific theories, and the fact that its position as a transfield scientific theory has been well established.

Starting in 1982, we have gradually recognized the meaning and value of the theory of grey systems, and started to learn and to study this theory. It is no doubt that trudging in any scientific discipline is not easy, and that it is more difficult to explore and to pioneer in a new theory. To this end, we have devoted the best years of our lives.

This research has been funded in succession by the China Natural Science Foundation, Henan Province Natural Science Foundation (China), Soft Science Foundation, Science Foundation for Prominent Young Scientists, National Science Foundation for Cross-Century Academic Leaders, etc. And, our work has brought forward new progress and breakthroughs in the areas of grey sequence operators (including weakening operators and strengthening operators), generalized degrees of grey incidence (including the absolute degree of grey incidence, relative degree of grey incidence, and synthetic degree of grey incidence), finding positioned solutions of linear and nonlinear programming models with grey parameters, G–E combined models, fixed weight grey clusterings, grey incidence clusterings, measurement of grey information, etc. All these results have obtained wide acceptance in the academic community. This book is surely the crystallization of our work of many years in the past.

During the entire period of creating this book, we have always put our emphasis on the scientificability, readability, and practical applicability, tried to present the material in a logical, systematic, and simple structure, and followed the principle of eliminating all mistakes in our reasoning. This book contains a total of twelve chapters, covering the theoretical foundation of grey systems theory, fundamental methods, and the main topics in grey systems theory, including grey sequence generation, grey systems analysis, modeling, predictions, decision making, optimization, control, etc. In the final chapter, we briefly describe some main topics on numerical computations of some of the major models presented in the book.

This book can be and most parts of this book have been, in the past fifteen years, used as a textbook for upper-level undergraduate and graduate students majoring in systems science, economics, and administration, and as a self-study book for students and scholars in areas such as geoscience, engineering, agriculture, medicine, meteorology, natural sciences, bioscience, etc. Best of all, this book can be and most parts of this book have been used as a reference by state employees, politicians, administrators, planners, and policy-makers in the past years and many years to come. Here in this book, we have absorbed the research work by Professor Deng Ju-Long and many others. With its current presentation of this manuscript, the reader can expect to learn grey systems theory in a systematic fashion. And, at the finish of this book, he or she can expect to be at the cutting edge of this new and exciting theory and applications.

Over the years, many people have been involved in the research, discussion, and writing of various parts of this book, including, but not limited to, Zhu Yongda, Yang Ling, Li Xiuli, Guo Tianbang, Dong Yaoguo, Guo Hong, Hou Yunxian, Zhao Li, Jia Yong, Donald McNeil, Lin Wen, Zeng Guoqing, Roman DeNu, Narendra Patel, Liu Quanfeng, Xu Xian, Adnan Mahmood, Hector Sabelli, Sun Suan, Cao Dianli, Liu Hongbin, Shi Benguang, Kimberly Forrest, Achim Sydow, Yang Wanzai, Wang Ziliang, Tan Xuerui, Zhao Deying, Wang Lianghua, Genti Zaimi, Ye Rongjun, Li Bingjun, Li Beiyou, Xu Chaozhi, Han Jianjun, Zhang Tao, Rebecca Martin, and Wan Yagang. Our parents, wives, and children have been patient and sacrificial in supporting our research and related writing. A great deal of support and encouragement has been given to us over the years from our colleagues and the administrators at Henan Agriculture University, International Institute for General Systems Studies, and Slippery Rock University. Finally, but not least, the editors and staff members at our publisher and the referees have done a great deal for the final publication of this book. We would like to use this opportunity to express our sincere appreciation to all the people, both listed and not listed above, for their teaching, role models, guidance, support, and encouragement. Without these people, this book would have been impossible.

Sifeng Liu, Ph. D.
Nanjing University of Aeronautics and Astronautics, China

and

Yi Lin, Ph. D.
International Institute for General Systems Studies, USA

July 30, 2004

Contents

1
Introduction

1.1 Scientific Background for the Appearance of Grey Systems Theory

Based on widespread divisions in the activities of scientific research and the newest development in technology, an obvious tendency has appeared in the modern spectrum of science and technology. This tendency is strongly indicated by the rapid rise of many crossdisciplinary research activities and appearance of several important theories. The rise and appearance of these scientific phenomena surely possess some significant methodological meanings. These crossdisciplinary theories have revealed more profoundly and essentially some important internal relationships among the traditionally more or less isolated subjects. These subjects have been studied in various superficially unrelated and artificially divided fields. These crossdisciplinary studies have deeply promoted the integrative progress of modern science and technology. With the help of these newly emerging fields, many complicated problems, unsolvable before, can be resolved successfully, and a deeper understanding about nature has been brought forward. These crossdisciplinary theories include, to say a few, systems theory, information theory, and cybernetics, which were formulated during the first half of this century, the theory of dissipative structures, synergetics, and fractals, which started to be known during the end of the 1960s and the beginning of the 1970s, ultracircular theory and general and pan-systems theory, which matured after the late 1970s and 1980s. As the focus of our book here, we concentrate on one of such theories: grey systems theory. The first piece of research on grey systems, entitled "The Control Problems of Grey Systems," written by Professor Deng Julong of China, was published in 1982 in the journal *Systems and Control Letters*. This theory, unlike many of the other crossdisciplinary theories, turned out to be an important and fruitful area of research with strong and successful practical (real-life) applications.

From the dialectical materialist point of view of science and technology, it can be said that each emergence of a new research subject or theory has its own certainty and fortuity. In general, at a certain period of time or stage in the development of a new scientific law, innovative research subjects or theories arise in order to meet the needs of the development. However, at this junction, the work of abandoning relevant old theories and developing new ones is accomplished by scientists who have insightful intelligence, unusual courage, and unprecedented resourcefulness. Historically speaking, scientists with all these qualities have appeared by chance. And, they have also been products of the historical moments within which they lived or within which they are living. For instance, from the development history of natural sciences, we can draw the conclusion that some of the greatest thinkers in the history had established themselves by finding the right new theories at the right times. At the same time, however, due to the limitations of the conventional thinking of their times, they were not successful in terms of crossing over some more monumental thresholds. For example, Ernst Mach, Hendrik A. Lorentz, and Jules-Henri Poincaré had all contributed a great deal to the final establishment of one of the greatest theories of mankind, which was later named and became well-known as the theory of relativity. Even though they were thought of by historians of the following generations to be on the threshold of discovering the great theory, they did not recognize the hidden treasure and, consequently, they did not continue in that direction. Here, as we all know, Albert Einstein, a young clerk working at a patent company, was credited for the formal establishment of such a theory, which turned out to be one of the most famous theories of the twentieth century. Mach's critique on Newtonian mechanics, contained in the book entitled *Developing Mechanics*, had greatly influenced Einstein. Interestingly, when young Einstein sent his articles to Mach and called him the pioneer of the theory of relativity, Mach refused to accept this honor. Lorentz even thought that the new theory fuzzed the conventional thinking of the time and felt great pain by saying that "I would rather have been dead five years before the emergence of the theory of relativity" (for more details see Song, 1989).

Similarly, the development of modern science and societal needs have provided the environment and conditions for the development of grey systems theory. At this very moment in the history of science and technology, it was Professor Deng Julong who formally established the theory of grey systems, and it was he and his followers who both developed the theory and obtained many great successes in practical real-life applications.

Since graduating in 1955 from Huazhong University of Science and Technology, one of the finest universities in China, Professor Deng has done his teaching and research there. His main research interest lies in the area of control of multivariable systems. In the 1960s, he brought forward "Control with Abandonment of Multivariable Systems," one of his high-quality works. This theory had been accepted as a representative method by the international academic community. Starting in the late 1960s, he has worked on both forecast and control of economic systems and fuzzy systems, dealing with numerous unascertained systems with partially known and partially unknown information. The characteristics of these systems can hardly be described by fuzzy mathematics or probability or statistics. Technically, fuzzy mathematics mainly deals with problems of the phenomenon with cognitive uncertainty by experience with the help of affiliation functions. Probability and statistics need special distributions and samples of reasonable size to draw valid inferences. However, what can we really do, if we are faced with situations with which we do not have any prior experience, or we cannot establish any necessary distribution, or we are only provided with a very small sample? To perform satisfactorily any meaningful research on these unascertained systems, Professor Deng has pioneered a difficult and fruitful research. In 1982, when everything seemed to be, to Professor Deng, "at the end of all hills and streams, beyond which there seems to be no world," he was astonished with the finding of "another village with shading willows and blooming flowers." Professor Roger W. Brockett of Harvard University, the editor-in-chief of the journal *Systems and Control Letters*, commented on Professor Deng's first article about grey systems as follows: "Grey system is an initiative work" and "all the results are new."

1.2 Fundamental Concepts and Principles of Grey Systems

1.2.1 Fundational Concepts of Grey Systems

Many systems, such as those that are social, economic, agricultural, industrial, ecological, or biological in nature, are named based on the fields and ranges to which the research subjects belong. In contrast, the name *grey systems* was chosen based on the colors of the subjects under investigation. For example, in control theory, the darkness of colors has been commonly used to indicate the degree of clarity of information. One of the most well-accepted representations is the so-called "black box." It stands for an object with its internal relations or structure totally unknown to the investigator. Here, we use the word "black" to represent unknown information, "white" for completely known information, and "grey" for that information which is partially known and partially unknown. Accordingly, we name systems with completely known information as *white systems*, systems with completely unknown information as *black systems*, and systems with partially known and partially unknown information as *grey systems*, respectively.

In our daily social, economic, and scientific research activities, we often face situations involving incomplete information. For example, in some studies of agriculture, even though all the information related to the area which is planted, the quality of seeds, fertilizers, irrigation, etc., is completely known, it is still difficult to estimate the production quantity and the consequent annual income due to various unknown or vague information related to labor quality, level of technology employed, natural environment, weather conditions, etc. As for the case of insect control, we might have known very well the relationship between the special kind of insect of interest and its natural enemies, yet it might still be difficult for us to achieve the desirable effects because we do not have enough information regarding relationships between the insects of our concern and the bait, their natural enemies and the bait, one natural enemy and other natural enemies, one kind of insect and other kinds of insects, etc. For each adjustment of a price system in our economy, decision makers often face the difficulty of not having definite information on the effect of the price change on consumers, on the prices of goods, etc. All liquid pressure systems are difficult to control due to some immeasurable quantities. Electricity systems are hard to observe due to the stochastic parameters of the voltage and currents. Such difficulty is caused by not having enough knowledge of motion and parameters. In a general social or economic system, it is difficult to analyze the effect of the input on the output due to the reasons that there do not exist clear differences between the "interior" and the "exterior", the system self and its environment, and that the boundary of the system may be sometimes easy to tell or on other occasions difficult to clarify. In scholastic works, the same economic variable could be seen as endogenous by some scholars and external by some others. The wide range of these phenomena is due to the lack of modeling information, or because an appropriate systems model has not been found, or the fact that the right observation and control variables have not been employed.

Based on the discussion above, there are four possibilities for incomplete information of systems.

1. The information of elements (or parameters) is incomplete.
2. The information on structure is incomplete.
3. The information on boundary is incomplete.
4. The behavior information of movement is incomplete.

Having "incomplete information" is the fundamental meaning of being "grey". In different circumstances and from different angles, the meaning of being "grey" can still be extended. For more details, see Table 1.1.

Table 1.1. Comparison between black, grey and white systems

	Black	Grey	White
Information	Unknown	Incomplete	Known
Appearance	Dark	Grey	Bright
Process	New	Replace old with new	Old
Property	Chaos	Complexity	Order
Methodology	Negative	Transition	Positive
Attitude	Indulgence	Tolerance	Serity
Conclusion	No result	Multiple solution	Unique solution

1.2.2 Fundamental Principles of Grey Systems

During the initial establishment and the consequent development phases, many important axioms have been proposed by scholars such as Professor Deng Julong.

Axiom 1.2.1. (*Principle of Informational Differences*) *"Difference" implies the existence of information. Each piece of information must carry some kind of "difference".*

When we say that object A is different from object B, we really mean that there is some special information about the object A that is not true with regard to the object B. All the "differences" existing between natural objects and events have provided us with the elementary information in order for us to understand their nature.

If information I has changed our understanding or impression of a complicated matter, then the piece of information I is definitely different from that on which we initially understood the special matter. Great breakthroughs in science and technology have provided us with the necessary information, which we generally call knowledge and tools, to understand and change the world around us. This advanced information is surely different from that preliminary information. The more content a piece of information I contains, the more difference from the earlier version of the information the new information I brings.

Axiom 1.2.2. (*Principle of Non-Uniqueness*) *The solution to any problem with incomplete and nondeterministic information is not unique.*

Because of this principle of non-uniqueness, which is a basic law of the application of grey systems theory, one is set free to look at this problem with flexibility. With flexibility, one becomes more effective in reaching his goals.

Strategically, the principle of non-uniqueness is realized through the concept of grey targets. This concept is a unification of the concept of non-unique targets and that of non-restrainable targets. For example, if a high school graduate does not plan to enroll in any university except one specific school, then his chance of success in terms of getting into a university is greatly limited. On the other hand, if a high school graduate, who has a similar qualification as the one in the previous example, is willing to attend several additional choices of universities other than the special one, he will be more likely to succeed in terms of attending a university, because he has multiple targets with an improved chance of hitting one of the targets.

The principle of non-uniqueness can be seen as a comprehensive realization that each target can be approached, that any available information can be supplemented, that each plan made earlier can be further modified and improved, that each relationship can be harmonized, that each thinking logic can be multi-directional, that each understanding can be deepened, and that each path can be optimized. When faced with the possibility of multiple solutions, one can locate one or several satisfactory solutions through deterministic analysis and supplementation of information. Therefore, the method of finding solutions on the basis of "non-uniqueness" is one that combines both quantitative analysis and qualitative analysis.

Axiom 1.2.3. (*Principle of Minimal Information*) *One characteristic of grey systems theory is that it makes the most and best use of the available "minimal amount of information."*

The "principle of minimal information" can be seen as a dialectic unification of "a little" and "a lot." One advantage of grey systems theory is its ability to handle such uncertain problems with "small samples" and/or "poor information." Its foundation of study is the concept of "spaces of limited information." "Minimal amount of information" is the basic territory for grey systems theory to show its power. The amount of information acquirable is the dividing line between "grey" and "not grey". Making sufficient discovery and application of any available "minimal amount of information" is the basic thinking logic of problem-solving used in grey systems theory.

Axiom 1.2.4. (*Principle of Recognition Base*) *Information is the foundation on which people recognize and understand (nature).*

This principle says that all recognition must be based on information. Without information, there will be no way for people to know anything. With complete and deterministic information, people can possibly gain firm recognition. Based on incomplete and non-deterministic information, people can only possibly obtain incomplete and non-deterministic grey recognition.

Axiom 1.2.5. (*Principle of New Information Priority*) *The function of new pieces of information is greater than that of old pieces of information.*

The "principle of new information priority" is the key point of view about information applied in grey systems theory. That is, by applying additional weights on newer information, one can achieve a better effect from grey modeling, grey prediction, grey analysis, grey evaluation, and grey decision making. The model of "the new replaces the old" reflects our "principle of new information priority.” With new information becoming available, a basic motivation for the whitenization of grey elements is strengthened. The “principle of new information priority” is a materialization of the fact that information in general is time sensitive.

Axiom 1.2.6. (*Principle of Absoluteness of Greyness*) *“Incompleteness” of information is absolute.*

Incompleteness and non-determinism of information have their generality. Each completeness of information is relative and temporary. It is the moment when the original non-determinism had just disappeared, and new non-determinism is about to emerge soon. Human recognition and understanding of the objective world have been elevated time and time again through continued supplementation of information. With the endless supply of information, man’s recognition and understanding of the world also become endless. That is, the greyness of information is absolute and will never disappear.

1.3 Comparison Between Several Nondeterministic Methods

Probability and statistics, fuzzy mathematics, and grey systems theory have been the three most-often applied theories and methods employed in studies of non-deterministic systems. Even though they study objects with different uncertainties, the commonality of these theories is their ability to make meaningful sense out of incompleteness and uncertainties. It is the differences among the uncertainties studied in these theories that three areas of scientific study, each of which has its own characteristics, on uncertainties have been established.

Fuzzy mathematics has its strength in the study of problems with “recognitive uncertainties.” All objects studied using fuzzy mathematics possess the characteristic of “having a clear intension without a clear extension”. For example, the concept of “young men” is a fuzzy concept. It is because all people know the intension of being a “young man”. However, it will be extremely difficult to define such a definite range of age that within the range a man is young and out of range a man is not young. The very reason why it is so difficult to introduce such a definition for the concept of “young men” is that the extension of this concept is not clear. When faced with this kind of unascertained problem with clear intension and unclear extension, fuzzy mathematics is the theory and method to use. The main idea of fuzzy mathematics is based on the so-called membership functions established based on experiences.

Probability and statistics study those phenomena with “stochastic uncertainties” with their emphasis placed on statistical patterns existing in the historical data of the phenomena through observing the chances for each possible outcome to occur. The basis for these theories to work and to produce reliable results consists of samples of large sizes and assumptions that these samples follow certain given patterns, named distributions.

With fuzzy mathematics, probability, and statistics described earlier, what we can say here is that grey systems theory is developed to study problems of "small samples and poor information." These problems studied by grey systems theory cannot be handled successfully by using either probability or statistics. Through coverage of information and generations of series, grey systems theory looks for realistic patterns based on modeling based on a few available data. Different from fuzzy mathematics, grey systems theory focuses on such research objects that have clear extension and unclear intension. For example, the Chinese government plans to control its national population within the range between 1.5 to 1.6 billion by 2050. This range "between 1.5 to 1.6 billion" is a grey concept with its extension clearly laid out without any knowledge about the specific population size.

Based on our discussion above, let us summarize the comparison of these three theories in the following Table 1.2:

Table 1.2. Comparison between grey systems theory, probability, statistics, and fuzzy mathematics

	Grey systems theory	Probability statistics	Fuzzy mathematics
Objects of study	Poor information Uncertainty	Stochastic Uncertainty	Cognitive Uncertainty
Basic sets	Grey hazy sets	Cantor sets	Fuzzy sets
Methods	Information coverage	Probability distribution	Function of affiliation
Procedure	Grey series generation	Frequency distribution	Marginal sampling
Requirement	Any distribution	Typical distribution	Experience
Emphasis	Intention	Intention	Extension
Objective	Laws of reality	Laws of statistics	Cognitive expression
Characteristics	Small samples	Large samples	Experience

1.4 Main Contents in Grey Systems Theory

Grey numbers, grey elements, and grey relations are the main subjects of research in grey systems theory. So, the entire theory of grey systems lies on the foundation of grey numbers and their operations, grey matrices, and grey equations. Control problems from industry and grey systems analysis, modeling, forecasting, decision making, and control of intrinsic characteristics of social, economic, agricultural, ecological, etc., systems are the main research tasks of grey systems.

Each study of a grey systems problem often needs a synthesized effort of many different aspects. For example, to make a long-term development plan of an administrative district or a business, the current situation needs to be analyzed and diagnosed first. On this basis, a systems model is established so that a scientific and reliable prediction of the future can be made. With a reliable prediction in mind, plans can be made with emphases chosen so that effective decision making and control can be performed in order to reach the object of less input and more output. In the study of food chains of ecological systems, for example, one has to deal with at the same time three layers consisting of green plants, herbivorous animals, and carnivorous animals. When making development plans for livestock farming, one has to analyze quantified relations among these three layers in order to predict the development and changes of each layer under human interference. With the understanding that human interference has a cost and might bring about some possible benefits, one needs to design a procedure to reduce the cost and to obtain more benefits, and to have specific implementation schedules of the procedure. In these examples, the contents of analysis, modeling, prediction, decision making, and control have all been included.

Grey systems analysis consists mainly of grey incidence analysis, grey statistics, grey clustering, etc. Grey systems modeling is done mainly through generations of grey numbers or functions of series operators to find hidden patterns, if any. Then, the modeling is finished based on the concept of five-step-modelings. The concept of five-step-modeling consists of

1. Language model;
2. Network model;
3. Quantification of model;
4. Dynamical quantification of model; and
5. Optimization of model.

Grey prediction is a quantitative prediction based on GM(1, 1). According to effectiveness and characteristics, grey predictions can be classified in the following six classes:

1. Serial predictions;
2. Interval predictions;
3. Disaster predictions;
4. Seasonal disaster predictions;
5. Stock-market-like predictions; and
6. Systems predictions.

Grey decision making includes grey target decision making; grey incidence decision making, grey statistics, grey clustering decision making, grey situation decision making; grey stratified decision making; and grey programming. The main contents of grey control cover (1) control problems of grey systems of intrinsic characteristics, (2) control based on grey systems methodology, such as grey incidence control, and (3) control of GM(1, 1) prediction, etc.

1.5 Role of Grey Systems Theory in the Development of Science

As a new theory with its own characteristics, grey systems theory has been generally recognized by both the civil and academic communities. In the name of grey systems, Professor Deng and his followers in the past two decades have contributed a great deal to the development of science and technology. For example, successful applications have been found in a great many areas of human endeavor, including agriculture, industry, energy resources, transportation, geology, meteorology, hydrology, ecology, environment, medicine, military science, economy, societal issues, and so on. Similarly, in manufacturing areas, profits have also been brought forward with successful applications of the new grey systems theory. Promoted by the theory of grey systems, some other new transfield scientific subjects have also been brought into the scene of academic activities, such as grey hydrology, grey geology, grey breeding, grey medicine, grey control theory, grey chaos, and grey analysis of regional economic systems. For more details, see the special volume *New Methods of Grey Systems*, as listed in Liu's paper (1993).

At the present time, many (both rookie and established) political and community leaders in several countries and differential geographical regions, including, but not limited to, Austria, Australia, Canada, England, Germany, Hongkong, Japan, Russia, Singapore, Taiwan, and the United States of America, have been engaging in the research and applications of grey systems theory. Courses named grey systems have been offered in over one hundred universities around the globe. Hundreds of Ph. D. candidates and Master's Degree candidates have been applying the thinking logic of grey systems to their scientific research for their academic degrees. At the time of this writing, in China alone, more than 160 scientific research projects reflecting the use of grey systems theory have been completed, 142 of which have won awards at Chinese national, provincial, or ministry levels. New findings on grey systems can be found in 201 different national and international periodicals. Surely, the idea and new results concerning grey systems have been reported at various international conferences and gathering places.

According to our incomplete statistics, in recent years, more than five hundred papers by leading Chinese scholars have been indexed by journals such as *SCI*, *EI*, *ISTP*, and *MR*. A commercial consulting firm has been established in Germany to collect and to translate papers on grey systems and to supply advisory services. In summary, during the short period of time of several years, the theory of grey systems was born, has grown, and now is standing on its own feet, based on its magnificent and versatile successes in the spectrum of modern science and technology as a new transfield subject matter.

1.6 Positions of Grey Systems Theory in the Spectrum of Interdisciplinary Sciences

No matter what an object is, different people will have different views on the same object. So, there exist different ways to divide each subject system under consideration. Based on human abilities, such as memory, imagination, and judgment, Francis Bacon concluded that science should be divided into three parts: history, poetry and art, and philosophy. Later, St. Simon and George F.W. Hegel put forward their way of dividing science based on metaphysics and idealism. During the final period of the nineteenth century, Friedrich Engels pointed out that subjects of knowledge should be divided according to their different forms of motion of matter and their innate sequences. Therefore, according to Engels, the structure of the science system can be formed and can lay a foundation for the division of subject matter. For more details, see Engels' work (1971).

In modern China, scholars divide science into two parts: liberal arts and science, or into three parts: natural science, mathematics, and social science. And, the foundation of natural science is thought to have six parts: mathematics, science, chemistry, geography, biology, and meteorology. Professor Xuesen Qian (1988) reasoned that the whole system of science should be divided into natural science, social science, systems science, science of thinking, science of the human body, and science of mathematics. Each scientific area should also be divided into three parts named basic science, technical science, and engineering science. In the book entitled *Outline of Science*, published in 1981, Professor Xipu Guan claimed that science and technology should be divided into three areas: natural science, social science, and science of thinking, and that each of the branches was made up of transfield subjects formed with mutual connections of both infiltrative subjects and synthetic subjects.

Here, from the point of view of grey systems, we first classify scientific problems according to complexity and uncertainty, and then point out the transfield subjects with methodological meanings in terms of the quality of scientific problems. In doing so, we can correctly position grey systems theory in the spectrum of all interdisciplinary sciences.

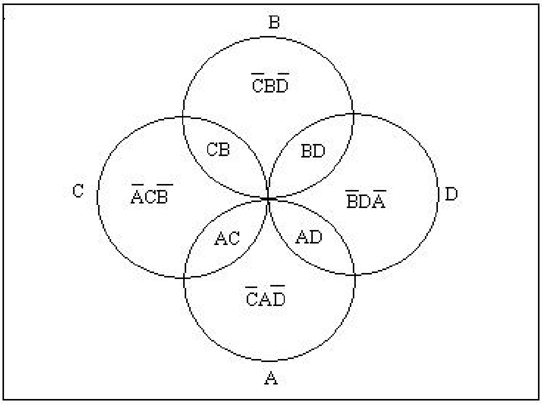

FIGURE 1.1. A four-leaf graph

We depict our division of science with a Venn diagram, (Figure 1.1), where the universal set consists of all possible subjects in the world of learning, and the four circles, labeled A, B, C, and D, stand for aggregations of simple subjects, complex subjects, certain subjects, and uncertain subjects, respectively. We can then obtain the four-leaf-graph as shown in Figure 1.1 providing a different way of division of all scientific endeavors. Each enclosed part of the four-leaf graph represents a scientific endeavor dealing with problems of the following features:

1. $\overline{C}B\overline{D}$: semi-certain complex problems,
2. $\overline{A}C\overline{B}$: certain semi-complex problems,
3. $\overline{C}A\overline{D}$: semi-certain simple problems,
4. $\overline{B}D\overline{A}$: uncertain semi-complex problems,
5. CB: certain complex problem,
6. BD: uncertain complex problems,
7. AD: uncertain simple problems, and
8. AC: certain simple problems.

Accordingly, with consideration of different problem-solving methods, we can obtain a division of all interdisciplinary sciences as follows:

1. $\overline{C}B\overline{D}$: theory of self-organization,
2. $\overline{A}C\overline{B}$: operations research,
3. $\overline{C}A\overline{D}$: theory of logic and intuition,

4. $\overline{B}D\overline{A}$: theory of grey systems,
5. CB: theory of general systems,
6. BD: nonlinear science,
7. AD: probability, statistics and fuzzy mathematics, and
8. AC: mathematics.

Therefore, we can draw the following conclusion: Under comparison with probability, statistics, and fuzzy mathematics, which study simple problems with uncertainty, the theory of grey systems, which studies uncertain and semi-complex problems, represents a new height in human learning. And, problems with complexity and uncertainty will be resolved only with breakthroughs in nonlinear science.

1.7 Grey Systems in the Content of Uncertain Information

It has been shown in the scientific literature that when predictions of zero probability events are concerned, no theoretical methods so far have been successful. According to recent publications, such a lack of success is mainly due to how information and consequent uncertainties are handled. In this section we show how grey systems compare to other concepts of uncertain information.

Because the problem of unsuccessful applications of traditional scientific theories involving uncertain information occurs mainly in the area of predictions of zero probability events, let us first look at the concept of a true historical process. According to George Soros, a legendary investor and financial guru, a true historical process is a process that leads to either unexpected outcomes or non-traditional beliefs. Consequently, a brand new page in human history is created. In each true historical process, the past creates expectations for all human participants so that predictions about what will happen in the near future are made. To different participants the expected events are viewed differently. For example, some wish the expected to happen whereas the others might like to avoid it as much as possible. So consequently, various human participants will behave differently from the past in order to materialize or to avoid the expectations and predictions. Consequent to these adjustments in human behavior the expected and predicted future is in general altered. That is, the expected events in general do not actually occur in the time frame or in the magnitude expected. That is, in order to be successful in predicting the outcome of a true historical process we must be able to handle uncertainties created by either accurate or inaccurate information. It is because all information, either accurate or inaccurate, brings forward uncertainties.

Historically, the concept of information has been defined in many different ways. To see clearly how grey systems compare with other kinds of uncertain information, we establish various information on a common ground.

A piece of tidings is meant the totality of a special form of objective motion. It is an objective entity, which reduces a human's level of ignorance. For example, the statement that "It will snow today," is a piece of tidings. This statement improves our outlook about the weather condition of the day.

Let A be a piece of tidings; then $A \cup \overline{A}$ means no tidings, because it contains the universal description of the tidings A and the opposite $\overline{A}$. For example, let A = the stock market will go up. Then $\overline{A}$ = the market will go down or sideways. Now, $A \cup \overline{A}$ = the stock market will go in some direction. At the same time the combined tidings $A \cap \overline{A}$ = the market will not go in any direction.

We use lowercase letters x, y, z, ... to stand for unknowns, which could be variables or statements, and A a piece of tidings. The notation $A\Delta x$ represents that the tidings A can make people know the value of the unknown x. Otherwise we write $A\nabla x$ to mean A is a piece of unrelated tidings about x; for example, given two pieces of tidings,

A = everyone has gone to watch a movie in the theatre;

B = the Dow Jones Industrial Average has gone down 400 points.

Now we are concerned with the unknown x = "where is Joe?" So, $A\Delta x$, because A provides an answer to x, even though we do not know whether the answer is true or false. That is, based on A, we know that Joe went to the theatre. At the same time, $B\nabla x$, because B does not provide any value and answer to x.

Assume x is an unknown, $\mathcal{U}$ a piece of tidings, and S a set of Cantor type. If $\mathcal{U}$ makes people realize $x \in S$, $\mathcal{U}$ is called a piece of x-position tidings. Each piece of tidings $A \subseteq \mathcal{U}$ is called a piece of information regarding the position tidings $\mathcal{U}$, or just information for short. The totality of all pieces of information of $\mathcal{U}$ is called an informational hierarchy.

Each so-called informational uncertainty stands for uncertainties related to information, or the quality of information. As indicated by Soros's Reflexivity Theory, in a true historic process, all information involved in the formation of predictions about the future could be very certain and definite. However, it is the certainty and the definiteness of the information, and the accuracy and preciseness of the predictions that make the future different and more uncertain. So, informational uncertainties are different from practical uncertainties.

Based on published studies, we have the following types of uncertainties:

1. Grey uncertainty,

2. Stochastic uncertainty,
3. Unascertainty,
4. Ascertainty,
5. Fuzzy uncertainty,
6. Rough uncertainty,
7. Soros reflexive uncertainty, and
8. Blind uncertainty.

1.7.1 Grey Uncertainties

Suppose A is a piece of grey information defined as follows. Let x be an unknown, $S \neq \varnothing$ a set, $S\prime$ a subset of S, $\mathcal{U}$ = "x belongs to S" and A = "x belongs to $S\prime$ ". Then the so-called grey uncertainty stands for the uncertainty of which specific value of the unknown x should take. For example, suppose we are given that $\mathcal{U}$ = "x belongs to S", S = "$\mathbf{R}$ is the set of all real numbers" , $S\prime$ = the interval [2,3], and A = "x belongs to $S\prime$". Then the piece of grey information A brings about the following uncertainty: we know that x is a number between 2 and 3 inclusive. However, we do not know which value x really assumes.

Example 1.7.1. In the negotiation process of buying a car, the buyer knows she would pay no more than $30,000. If x stands for the final negotiated price of the car she likes, then x is a number between $0 and $30,000. Here, the grey uncertainty is the uncertainty about the final purchase price.

1.7.2 Stochastic Uncertainty

If x is unknown, S a nonempty set, $\mathcal{U}$ = "x belongs to S" and A = "x belongs to S and the possibility for $x = e \in S$ is α_e, where $0 \leq \alpha_e \leq 1$ and $\sum_{e \in S} \alpha_e = 1$." In this case A is called a piece of stochastic information. When a piece of stochastic information is given, the consequent uncertainty is called stochastic uncertainty. Such uncertainty is created because the piece of stochastic information A can only spell out how likely the unknown x equals a special element $e \in S$. This implies that the probability α_e can be very close to 1 or equal to 1, however, the large probability does not guarantee that $x = e$ will definitely be true.

Example 1.7.2. In the business of commodity trading, we can compute based on the historical price data that the market for the S&P500 has a 90% chance to go up on a certain Thursday. So some traders would buy in to the S&P500 futures contracts on Thursday and sell out on the calculated day, which may be the next Monday or Tuesday. However, the 90% possibility of a rising S&P500 futures market does not guarantee this time when we buy on Thursday the market will go up as expected.

Example 1.7.3. The current commercial weather forecasting business tends to provide services as follows. The chance of snow for tomorrow is 70%. If it does snow the next day, the weather forecasting service is correct because they said it would. On the other hand, if it does not snow the next day, the weather forecasting service is still correct because it only provided 70% likelihood of a forthcoming snow. Now, if the figure 70% is replaced by 100%, the same thing can be still be said about the weather forecasting service since the service only stated the chance of snowing was 100%, which was not a guarantee.

1.7.3 Unascertainty

If in the definition of a piece of stochastic information A, we replace the condition that

$$\sum_{e \in S} \alpha_e = 1$$

by

$$\sum_{e \in S} \alpha_e \leq 1$$

then A is called a piece of unascertained information.

The main difference between stochastic and unascertained information is that the former concept is developed on the assumption that all possible outcomes of an experiment are known, whereas for unascertained information, we assume that only some possible outcomes of the experiment are known to the researcher.

Example 1.7.4. A group of researchers have a scheduled meeting at 11:30 AM Thursday. However, at around 11:45 AM, Genti, as a key member of the group, did not show up. So the rest of the group need to decide where to find him for their urgent business decision making. Now, we face two possible situations.

Situation 1: The group knows Genti very well. So the members come up with a definite list of possible places Genti could be at the moment. Because they know Genti so well, they could also attach a probability to each place on the list. So, to locate Genti successfully, they only need to check these places in the order from the largest probability to the smallest probability. This is an example of stochastic uncertainties.

Situation 2: No member of the group knows Genti well enough to come up with a list of all possible places and relevant probabilities where Genti could be at this very moment. This is an example of unascertainties.

The second situation explains that the whereabouts of Genti at the special moment was certain because as a living being, he must be at *some place*. However, the decision makers did not know the true state of Genti or relevant information used by Genti in his decision about where to go and to be at the special moment. That is, the concept of unascertainty deals with the situation that no matter whether an objective event is definite or not, whether it has already occurred or not, it will be "unascertained" as long as the decision maker does not completely understand the essential information. □

1.7.4 Fuzzy Uncertainty

A piece A of tidings is called a piece of fuzzy information, if A satisfies: x is an unknown, S a nonempty set, the position tiding $\mathcal{U}$ = "x belongs to S", and A = "x belongs to S and the degree of the membership for $x = e \in S$ is α_e, $0 \leq \alpha_e \leq 1$."

Example 1.7.5. Jacklin Ruscitto is an official member of many committees. Due to the nature of these committees, Jacklin does not have enough time to be involved 100% in all of the committee works and relevant decision making. Let us look at one committee, say committee A. If Jacklin is listed as a member but did not ever do anything for the committee, then her degree of membership in the committee would be very close to zero. If Jacklin was not listed as a member on the committee and did not do anything either for the committee then her degree of membership in the committee is zero. Even though she is not a listed member of the committee, if she has been involved in activities of the committee, then her degree of membership in the committee should be greater than zero.

Now, the so-called fuzzy uncertainty will be that for a given piece of fuzzy information A = "x belongs to S and the degree of membership for $x = e \in S$ is α_e, $0 \leq \alpha_e \leq 1$," one has no clue on that for a given variable y, should y be considered with the set S or not, even though he knows the degree of membership of y in S is α_e.

For example, Jacklin is listed as an official member of Committee A and has been involved in all committee activities. So her degree of membership in committee A is 1. Now, the fuzzy uncertainty implies that her degree 1 of membership in Committee A does not guarantee her 100% involvement or membership in Committee A in the future. On the other hand, even though John Opalanko is not a listed member of Committee A, it is very likely that because Committee A is involved in a special project that looks extremely important in John's eyes, John may very well get involved in the project. In this case John's degree of membership in Committee A should be more than zero even though his previous degree of membership in Committee A was zero.

1.7.5 Rough Uncertainty

Let $\mathcal{U}$ be a set of elements. And, a subset $r \subseteq p(\mathcal{U})$, the power set of $\mathcal{U}$, is called a partition of $\mathcal{U}$, if the following conditions hold true:

1. $\cup r = \cup\{x : x \in r\} = U$,

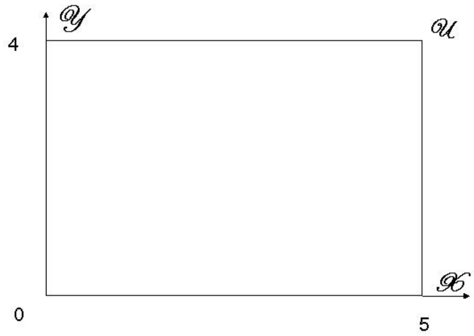

FIGURE 1.2. The given rectangular area $\mathcal{U}$

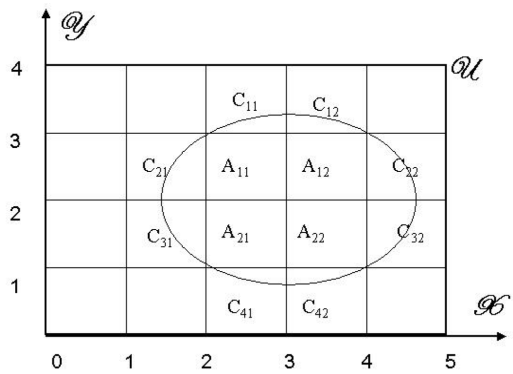

FIGURE 1.3. The partition r of the area $\mathcal{U}$

2. $\forall A, B \in r$, if $A \neq B$, then $A \cap B = \varnothing$.

Let $K = (\mathcal{U}, \mathbf{R})$ be a knowledge base over $\mathcal{U}$, where $\mathcal{U}$ is the universal set of all objects involved in a study, and $\mathbf{R}$ is a given set of partitions of the set $\mathcal{U}$, called a knowledge base over $\mathcal{U}$. A subset $X \subseteq \mathcal{U}$ is called exact in K, if there exists a $\mathbf{P} \subseteq \mathbf{R}$ such that X is the union of some elements in $\cap\mathbf{P}$. Otherwise, X is said to be rough in K.

Example 1.7.6. Let $\mathcal{U}$ be the rectangular area $\mathcal{U} = \{$(x,y) : $0 \leq x \leq 5$, $0 \leq y \leq 4\}$ (Figure 1.2).

A given partition r of $\mathcal{U}$ is defined as follows: each element x in r is a smaller rectangular area as shown in Figure 1.3

such that (1) if x is not on the bottom row, x includes all interior points and points on the upper and left borders; (2) if x is on the bottom row, then x includes all points as described in (1) and those on the bottom border; (3) if x is located on the far right column but not on the bottom x contains all points as in (1) and the points on the right border; and (4) if

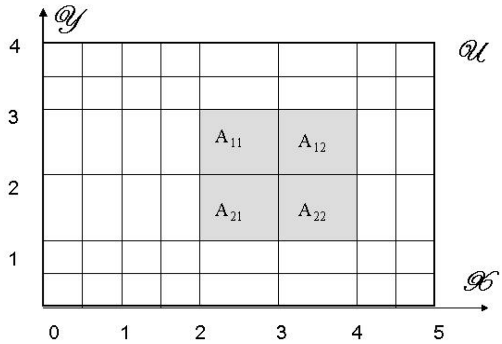

FIGURE 1.4. The partition s of the area $\mathcal{U}$

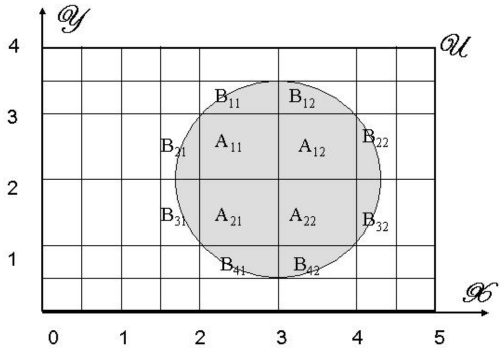

FIGURE 1.5. Rough approximations of the area A

x is located at the lower right corner then x contains all the points within and on the border of the rectangular area.

Assume that s is another partition of $\mathcal{U}$ as shown in Figure 1.4 with border points classified in a similar way as in the partition of r above.

Then the shaded area of $\mathcal{U}$ is an exact subset of $\mathcal{U}$ because it is the union of elements A_{11}, A_{12}, A_{21}, and A_{22} of the partition r. Now the shaded area A in Figure 1.5 is considered a rough subset of $\mathcal{U}$, because it does not equal the union of any combination of elements in r and/or s. It implies that in order to know more about this rough set A, one has to approximate it in two different ways if the partition s is applied.

1. $\underline{s}A = A_{11} \cup A_{12} \cup A_{21} \cup A_{22} = \cup\{A_{ij} : i, j = 1, 2\}$.
2. $\overline{s}A = \cup\{A_{ij} : i, j = 1, 2\} \cup \{B_{ij} : i = 1, ..., 4, j = 1, 2\}$.

If the partition r is applied, the following approximations could be employed.

3. $\underline{r}A = A_{11} \cup A_{12} \cup A_{21} \cup A_{22} = \cup\{A_{ij} : i, j = 1, 2\}$.
4. $\overline{r}A = \cup\{A_{ij} : i, j = 1, 2\} \cup \{C_{ij} : i = 1, ..., 4, j = 1, 2\}$.

Because elements in s are finer than those in r, approximations 1 and 2 of the rough set A are expected to be more accurate than those of 3 and 4. From this example, one can see that no matter how fine a partition is, its approximations to a given rough set always contain uncertainty. This uncertainty is in the area of $\overline{s}A - \underline{s}A$ or $\overline{r}A - \underline{r}A$, where

$$\overline{s}A - \underline{s}A = \cup\{B_{ij} : i = 1, ..., 4, j = 1, 2\}$$

and

$$\overline{r}A - \underline{r}A = \cup\{C_{ij} : i = 1, ..., 4, j = 1, 2\}$$

1.7.6 Soros Reflexive Uncertainty

Let x be the unknown path a true historic process will eventually take and $S =$ all possible outcomes of this historical process. Then, a Soros reflexive uncertain information is defined as follows: $\mathcal{U}$ = "$x \in S$" and $A \subseteq S$ is a piece of information regarding the position of x in S defined by $A =$ "if it is expected $x = e \in S$ with a degree of credence α_e, $0 \leq \alpha_e \leq 1$, then $x = e \in S$ has a degree of credence $1 - \alpha_e$." Now the uncertainty associated with a piece of Soros reflexive uncertain information is that the more accurate a prediction about a true historical process is, the more uncertain the expected future will become.

Example 1.7.7. George Soros started in the early 1960s to apply the concept of reflexivity to understand finance, politics, and economics. Based on his success in practice and over 40 years of theoretical study, he drew the following conclusion: "Statements whose truth value is indeterminate are more significant than statements whose truth value is known or definite. The latter constitute knowledge. They help us understand the world as it is. But the former, expressions of our inherently imperfect understanding, help to shape the world in which we live."

Based on his theory of reflexivity, Mr. Soros always looks out for situations where he sees an investment opportunity differently from the prevailing wisdom. He assumes all companies and industries to be flawed and tries to find what the flaws are. As soon as there are signs to show that the flaws are becoming a problem he will take financial positions so that he puts himself ahead of the investment game. More specifically, when a company has a superior market position, competent management, and exceptional profit margins, the stock may be overvalued. Now, the management may become complacent and the competitive or regulatory environment may change. When Soros looks for flaws, he establishes a hypothesis on which he would invest. Each of his hypotheses satisfies the following conditions.

(1) The hypothesis has to differ from the widely accepted wisdom. The greater the difference is, the greater the profit potential. (2) Each of his hypotheses does not have to be true in real life to be profitable, as long as it could be generally accepted. (3) The length of time the beneficial effects of the hypothesis can provide depends on whether the underlying flaws are recognized and corrected.

In 1972, Mr. Soros sensed that a change was about to occur in the banking industry because banks had the worst reputations at the time. Their employees were considered stodgy and dull, and few believed that the banks would rouse themselves from their deep slumber. That was why investors showed no interest in their stocks. After doing his homework, Soros realized that a new generation of bankers were quietly ready and in place to act aggressively on behalf of their employers. Because the new generation of bank managers was using new financial instruments, the banks earnings performances were looking up. However, the bank stocks were sold at virtually no premium. Many banks had reached their leveraging limits. So, in order to grow, they needed more equity. When the First National City Bank hosted a dinner for security analysts in 1972, even though Mr. Soros was not invited, he sensed the forthcoming aggressive change in the way banks would conduct their business. So, he wrote a brokerage report arguing that because bank shares had been going nowhere they were about to take off, contrary to what many others thought. Timing the publication of his report to coincide with the bank's dinner he recommended getting behind some of the better-managed banks. As expected, bank stocks began to rise and he made a handsome 50% profit in a short period of time. The legendary success of Mr. Soros' lifelong investing has evidenced a widespread existence of our so-called Soros reflexive uncertainty in real-life situations with human participants.

Our presentation in this section shows that the study of grey systems is only the start of an ambitious scientific project: Establish a unified information theory. With unascertained information well understood, we expect that many problems facing scientific practitioners, which have not been addressed successfully in the past, can be resolved with desirable satisfaction.

2

Grey Numbers and Their Operations

2.1 Grey Numbers

Each grey system is described with grey numbers, grey equations, grey matrices, etc. Here, grey numbers are the elementary "atoms" or "cells".

A *grey number* is such a number whose exact value is unknown but a range within that the value lies is known. In applications, a grey number in general is an interval or a general set of numbers.

Following are several classes of grey numbers.

1. *Grey numbers with only lower limits.*

The grey numbers with lower limits but not upper limits are denoted as

$$\otimes \in [\underline{a}, \infty) \text{ or } \otimes(\underline{a}),$$

where $\underline{a}$ represents the lower limit of the grey number $\otimes$, that is a fixed value. We call $[\underline{a}, \infty)$ the value field of the grey number $\otimes$ or briefly a grey field.

For example, the weight of a living tree is a grey number with a lower limit, because the weight of the tree must be greater than zero. However, the exact value for the weight cannot be obtained through normal means. If we use the symbol $\otimes$ to represent the weight of the tree, we then have that $\otimes \in [0, \infty)$.

2. *Grey numbers with only the upper limits.*

The grey numbers with only upper limits are written as

$$\otimes \in (-\infty, \overline{a}] \text{ or } \otimes(\overline{a}),$$

where $\overline{a}$ stands for the upper limit of the grey number $\otimes$ and is a fixed number.

For example, for an investment opportunity, there always exists a upper limit representing the maximum amount of money that can be mobilized. For an electrical equipment, there must be a maximum critical value for the equipment to function normally. The critical value could be for a maximum voltage or for a maximum amount of current allowed to be applied to the equipment. So, the amount of dollars, that can be used for a specific investment opportunity, the voltage and the current requirements for electrical equipment are all examples of grey numbers with only upper limits.

3. *Interval grey numbers.*

A grey number with both a lower limit $\underline{a}$ and a upper limit $\overline{a}$ is called an interval grey number, denoted as $\otimes \in [\underline{a}, \overline{a}]$.

For example, the weight of a seal is between 20 and 25 kg. A specific person's height is between 1.8 and 1.9 meters. These two grey numbers can be respectively written as

$$\otimes_1 \in [20, 25] \ \text{and} \otimes_2 \in [1.8, 1.9] .$$

4. *Continuous grey numbers* and *discrete grey numbers.*

The grey numbers taking on a finite number of values or a countable number of values in an interval are called discrete grey numbers. And, those continuously taking values that covers an interval are continuous grey numbers.

For example, if a person's age is between 30 and 35, his or her age could be one of the values 30, 31, 32, 33, 34, 35. So, age is a discrete grey number. As for a person's height, weight, etc., they are continuous grey numbers.

5. *Black* and *white numbers.*

When $\otimes \in$ (-∞, ∞) or $\otimes \in$ ($\otimes_1$, $\otimes_2$), that is, when $\otimes$ has neither an upper limit nor lower limit, or the upper and the lower limits are all grey numbers, $\otimes$ is called a black number.

When $\otimes \in [\underline{a}, \overline{a}]$ and $\underline{a} = \overline{a}$, $\otimes$ is called a white number.

For the sake of convenience in our discussion, we treat black and white numbers as special grey numbers.

6. *Essential grey numbers* and *non-essential grey numbers.*

An essential grey number is a grey number that is impossible or temporarily not possible to find a white number to represent. For example, a general forecast value, the total amount of energy in the universe, an "age" with accuracy to seconds or milliseconds, etc., are all examples of essential grey numbers.

A non-essential grey number is a grey number $\otimes$ that can be described with a white number as its "representative," where the white number is determined by using either previously known information or through some other means. This white number is called the whitenization (value) of the relevant grey number, denoted as $\tilde{\otimes}$. And, $\otimes(a)$ is used to stand for the grey number with a as its whitenization value. For example, we ask somebody to help buy a snow coat for about \$100. This number 100 can be treated as the whitenization value of the future coat price $\otimes(100)$, denoted as $\tilde{\otimes}(100) = 100$.

Fundamentally, the set of all grey numbers can be divided into three classes of different types.

1. *Grey numbers of information type* are those whose values cannot be certain due to temporary shortage of information. For example, an estimate indicates that the summer crop production of a certain region for the current year is over 100,000 tons. That is, $\otimes \in [100{,}000, \infty)$. It is estimated that a local branch bank would have a total of deposits in dollars somewhere between \$700,000 and \$900,000. That is, $\otimes \in [700,000, 900,000]$. It is predicted that the maximum temperature in the greater Pittsburgh area of Pennsylvania in May would not go beyond 102^{o}F. That is, $\otimes \in [0, 120]$. In all these examples, we have dealt with some grey numbers of information type. Due to a temporary shortage of information, we cannot be certain about the exact values of these grey numbers. However, after a certain period of time, through addition of new information, the relevant grey numbers can be whitenized. Specifically, in the previous three examples, as soon as the time periods involved in the prediction are over, the grey numbers will all become complete determined numbers.

2. Among all *conceptual grey numbers*, some are also called *grey numbers of wish type*, that means that these grey numbers are formed based on people's wishes and thoughts. For example, a scientist wishes to obtain a research grant of at least \$10,000 and the more the better. That is, $\otimes \in [10,000, \infty)$. A manufacturing business has a history of producing 1% defective products. In the hope of increasing its profit margin, the administration of the business wishes to reduce the rate and the smaller the rate the better. So, $\otimes \in [0, 0.01]$. All these grey numbers are conceptual type.

3. *Grey numbers of layer type* are those formed by changing layers. Some numbers are white, if seen from the height of the system's level. That is the macrocosmic level, the level of the whole, or the level of cognition. However, they might be grey, if seen at some lower level, say, at a microcosmic level of the system, the level of parts, or from the depth of cognition. For example, human height is white if measured in centimeters, and is grey if measured with the accuracy of 10,000th of a millimeter. Also, for some numbers, they are white within a small domain, and grey when considered in a larger range. For example, the number of persons named Tom could be 1 if one only considers the pool of people in a certain class of a certain university. It could be between 10 and 25 if the entire university is considered. So, $\otimes \in [10, 25]$ has already become a grey number. However,

if we consider the entire world, no one really knows what the answer would be. That is, $\otimes \in [1, \infty)$.

2.2 Whitenization of Grey Numbers and Degree of Greyness

There is a class of grey numbers that vibrate around a base value. This class of grey numbers can be whitenized relatively easily, because we can use the base value as the main whitenization value. A grey number with a base value a can be denoted as

$$\otimes(a) = a + \delta_a \text{ or } \otimes(a) \in (-, a, +),$$

where δ_a stands for the vibration variable. The whitenization value of this grey number is $\otimes(a) = a$. For example, the grey number that this year's research expense of a small group of scientists from a local university will be about \$10,000 can be expressed as

$$\otimes(10,000) = 10,000 + \delta$$

or

$$\otimes(10,000) \in (-, 10,000, +),$$

whose whitenization value is 10,000.

For a general interval grey number $\otimes \in [a, b]$, we take its *whitenization value* $\tilde{\otimes}$ as

$$\tilde{\otimes} = \alpha a + (1-\alpha) b, \alpha \in [0, 1].$$

Definition 2.2.1. *The whitenization of the form* $\tilde{\otimes} = \alpha a + (1-\alpha) b$, $\alpha \in [0, 1]$, *is called equal weight whitenization.*

Definition 2.2.2. *In an equal weight whitenization, the whitenization value, obtained when taking* $\alpha = \frac{1}{2}$, *is called an equal weight mean whitenization.*

When the distribution information of an interval grey number is hardly known, we often use the equal weight mean whitenization.

Definition 2.2.3. *Assume that the interval grey numbers* $\otimes_1 \in [a, b]$ *and* $\otimes_2 \in [c, d]$ *have the whitenizations*

$$\tilde{\otimes}_1 = \alpha a + (1-\alpha) b, \alpha \in [0, 1]$$

and

$$\tilde{\otimes}_2 = \beta c + (1-\beta) d, \beta \in [0, 1].$$

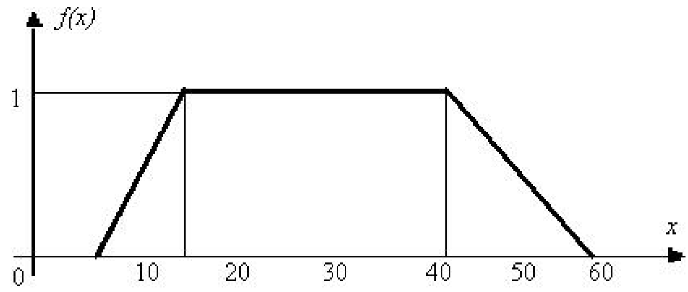

FIGURE 2.1. The weight function of whitenization for the content of quick-acting nitrogen in soil

When $\alpha = \beta$, we say that the grey numbers $\otimes_1$ and $\otimes_2$ are synchronous. When $\alpha \neq \beta$, we say that the grey numbers $\otimes_1$ and $\otimes_2$ are not synchronous.

When the distribution information of a grey number is known, we often use non-equal weight whitenization. For example, the statement that a person's age is possibly between 40 and 60 is a grey number $\otimes \in [40, 60]$. Based on some research, it is found that he spent 12 years in pre-college education, and entered into a university in the mid-1960s. So, it is more likely that he is about 50 years old. Or, in other words, it is more likely for him to be between 45 and 55 years old. For this grey number, if one uses an equal weight whitenization, it is obviously not very reasonable. Therefore, we employ a *whitenization weight function* to describe the degree of preference of a grey number to take values in its range.

For example, the contents of chemical elements nitrogen, phosphorus, and potassium are all grey numbers. To obtain a normal growing soil condition, the contents of quick-acting nitrogen should be between 15 to 40 ppm. So, we can use the *weight function of whitenization* as shown in Figure 2.1 to describe the content of quick-acting nitrogen in the soil of interest. Here, the flat top with weight 1, represents the optimal content of quick-acting nitrogen. The left slope stands for the content of quick-acting nitrogen from 5 to 15 ppm so that the higher the better the effects would be. The right slope indicates the content from 40 to 60 ppm such that the higher the content is the worse the effect would be on the production of a certain crop. The curve starts at 5 ppm and ends at 60 ppm, that implies that content of less than 5 ppm or more than 60 ppm is not allowed for the production of the crop in the area of consideration.

For grey numbers of conceptual type, representing wishes, their weight functions of whitenization in general are designed as monotonic increasing functions. The weight functions of whitenization, contained in Figure 2.2, represent the grey number of an amount of a loan and related degree of "preferences", where the straight line represents the "normal wish" (or the

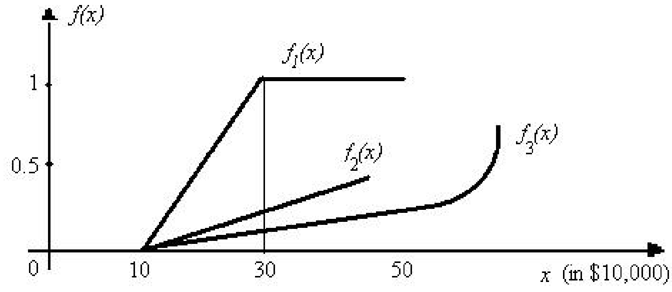

FIGURE 2.2. Weight functions of whitenization for loans and preferences

"preference" is proportional to the amount of the loan of concern with the slope describing the degree of desire to obtain the loan). $f_1(x)$ implies that the desire of obtaining a loan is not very strong with the understanding that a loan of less than $100,000 is not enough, a loan of around $200,000 is relatively satisfactory, and a loan of $300,000 is enough. Function $f_2(x)$ represents such a strong desire that a loan of $350,000 only carries a 20% degree of satisfaction. Function $f_3(x)$ shows that even for a loan of $400,000, the satisfaction level is only 10%. However, a loan of $500,000 is enough. More specifically, this function states that a loan of about $500,000 must be obtained otherwise there would be no other choices.

In general, the weight function of whitenization of a grey number is designed based on known information and experience of the researcher. It has no fixed procedure to follow except that the starting and ending points of the function should have some significance. For example, during an international business negotiation, there exists a process of change from grey to white. One country claims that its exporting must be at least, say, $5 billion, and the other side insists that the importing must be limited to a scale under $3 billion. So, the final deal successfully agreed upon will be somewhere between $3 billion to $5 billion. The weight function of whitenization can have a start at $3 billion and an end at $5 billion.

Definition 2.2.4. *The continuous functions with fixed starting and ending points and increasing on the left and decreasing on the right are called typical weight functions of whitenization.*

Typical weight functions of whitenization in general looks as shown in Figure 2.3a, where

$$y = f_1(x) = \begin{cases} L(x), & x \in [a_1, b_1) \\ 1, & x \in [b_1, b_2] \\ R(x), & x \in (b_2, a_2] \end{cases} .$$

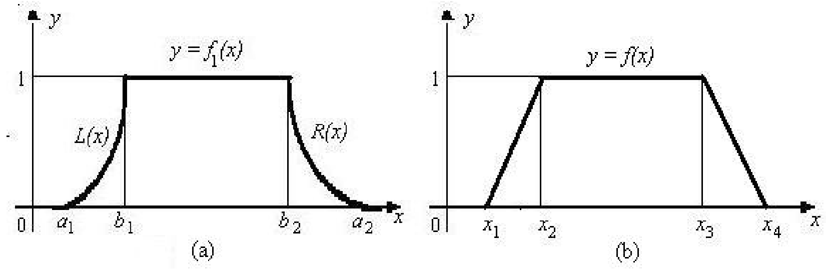

FIGURE 2.3. Typical weight function of whitenization

We call the function $L(x)$ the left increasing function, $R(x)$ the right decreasing function, $[b_1, b_2]$ the peak area, a_1 the starting point, a_2 the ending point, and b_1 and b_2 the turning points.

In applications, for the sake of convenience for computer programming and calculation, the functions $L(x)$ and $R(x)$ are often simplified as straight lines (see Figure 2.3b). That is, we have

$$f(x) = \begin{cases} L(x) = \dfrac{x - x_1}{x_2 - x_1}, & x \in [x_1, x_2) \\ 1, & x \in [x_2, x_3] \\ R(x) = \dfrac{x_4 - x}{x_4 - x_3}, & x \in (x_3, x_4] \end{cases} .$$

Theorem 2.2.1. *Assume that X is a space of all real numbers, $Y \subseteq [0, 1]$ and $f : X \to Y$ satisfies the conditions of typical weight functions of whitenization, then f satisfies the following conditions.*

1. *$f(\emptyset) = \emptyset$;*
2. *$f(X) = Y$;*
3. *$A, B \subseteq X$, if $A \subseteq B$, then $f(A) \subseteq f(B)$;*
4. *if $A \neq \emptyset$, then $f(A) \neq \emptyset$;*
5. *$f(A \cup B) = f(A) \cup f(B)$;*
6. *$f(A \cap B) \subseteq f(A) \cap f(B)$.*

Definition 2.2.5. *Suppose f is a typical weight function of whitenization,*

$$Y \subseteq [0, 1],\ f^{-1}(y) = \{x |\ f(x) = y\},\ y \in Y.$$

Then f^{-1} is called the inverse function of f.

Definition 2.2.6. *For the typical weight function of whitenization as shown in Figure 2.3a, the following*

$$g^o = \frac{2\,|b_1 - b_2|}{b_1 + b_2} + \max\left\{\frac{|a_1 - b_1|}{b_1}, \frac{|a_2 - b_2|}{b_2}\right\}$$

is called the degree of greyness of the grey number $\otimes$.

The expression of g^o is a sum of two parts. The first part represents the effect of the size of the peak area on the degree of greyness, and the second part shows the effect of the size of the areas under the functions $L(x)$ and $R(x)$ on the degree of greyness. In general, the greater the peak area is and the greater the areas under the functions $L(x)$ and $R(x)$, the greater g^o.

When

$$\max\{\frac{|a_1 - b_1|}{b_1}, \frac{|a_2 - b_2|}{b_2}\} = 0,$$

then

$$g^o = \frac{2|b_1 - b_2|}{b_1 + b_2}.$$

So, the weight function of whitenization becomes a horizontal line.

When

$$\frac{2|b_1 - b_2|}{b_1 + b_2} = 0,$$

the grey number $\otimes$ is such that it has a basic value, that equals $b = b_1 = b_2$.

When $g^o = 0$, $\otimes$ is a white number.

2.3 Operations of Interval Grey Numbers

In this section we discuss various operations of interval grey numbers.

Assume that we have grey numbers

$$\otimes_1 \in [a, b]\,,\ a < b$$

and

$$\otimes_2 \in [c, d]\,,\ c < d.$$

If we use the symbol $*$ to represent an operation between $\otimes_1$ and $\otimes_2$, and if

$$\otimes_3 = \otimes_1 * \otimes_2,$$

then $\otimes_3$ should also be an interval grey number. So, we should have

$$\otimes_3 \in [e, f]\,,\ e < f$$

and for any $\tilde{\otimes}_1$ and $\tilde{\otimes}_2$, $\tilde{\otimes}_1 * \tilde{\otimes}_2 \in [e, f]$.

Rule 2.3.1. *Assume that* $\otimes_1 \in [a, b]$, $a < b$, *and* $\otimes_2 \in [c, d]$, $c < d$. *The sum of* $\otimes_1$ *and* $\otimes_2$, *written* $\otimes_1 + \otimes_2$, *is defined as follows.*

$$\otimes_1 + \otimes_2 \in [a + c,\ b + d] .$$

Example 2.3.1. Given $\otimes_1 \in [3, 4]$ and $\otimes_2 \in [5, 8]$, then

$$\otimes_1 + \otimes_2 \in [8,\ 12] .$$

Rule 2.3.2. *Assume that* $\otimes \in [a, b]$, $a < b$. *The* negative *inverse of* $\otimes$, *written* $-\otimes$, *is defined as follows.*

$$-\otimes = [-b, -a] .$$

Rule 2.3.3. *Assume* $\otimes_1 \in [a, b]$, $a < b$, *and* $\otimes_2 \in [c, d]$, $c < d$. *The* $difference$ *of* $\otimes_1$ *with* $\otimes_2$ *is defined as follows.*

$$\otimes_1 - \otimes_2 = \otimes_1 + (-\otimes_2) \in [a - d,\ b - c] .$$

Example 2.3.2. Given $\otimes_1 \in [3,\ 4]$, and $\otimes_2 \in [1,\ 2]$, then we have

$$\begin{aligned} \otimes_1 - \otimes_2 &\in [3 - 2,\ 4 - 1] = [1,\ 3] \\ \otimes_2 - \otimes_1 &\in [1 - 4,\ 2 - 3] = [-3,\ -1]. \end{aligned}$$

Rule 2.3.4. *Assume* $\otimes \in [a, b]$, $a < b$, *and* $ab > 0$. *The reciprocal of* $\otimes$, *written* $\otimes^{-1}$, *is defined as follows.*

$$\otimes^{-1} \in [\frac{1}{b}, \frac{1}{a}].$$

Example 2.3.3. Given $\otimes \in [2,\ 4]$, we have $\otimes^{-1} \in [0.25,\ 0.5]$.

Rule 2.3.5. *Assume* $\otimes_1 \in [a, b]$, $a < b$, *and* $\otimes_2 \in [c, d]$, $c < d$. *The product of* $\otimes_1$ *and* $\otimes_2$ *is defined as follows.*

$$\otimes_1 \cdot \otimes_2 \in [\min\{ac, ad, bc, bd\},\ \max\{ac, ad, bc, bd\}].$$

Example 2.3.4. For $\otimes_1 \in [3,\ 4]$, and $\otimes_2 \in [5,\ 10]$, we have

$$\begin{aligned} \otimes_1 \cdot \otimes_2 &\in [\min\{15, 30, 20, 40\}, \max\{15, 30, 20, 40\}] \\ &= [15,\ 40]. \end{aligned}$$

Rule 2.3.6. *Assume* $\otimes_1 \in [a, b]$, $a < b$, *and* $\otimes_2 \in [c, d]$, *satisfying* $c < d$ *and* $cd > 0$. *The quotient of* $\otimes_1$ *divided by* $\otimes_2$ *is as defined as follows.*

$$\otimes_1 / \otimes_2 = \otimes_1 \cdot \otimes_2^{-1}.$$

That is,

$$\frac{\otimes_1}{\otimes_2} \in \left[\min\{\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d}\}, \max\{\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d}\}\right].$$

Example 2.3.5. For $\otimes_1 \in [3,\ 4]$, and $\otimes_2 \in [5,\ 10]$, we have

$$\frac{\otimes_1}{\otimes_2} \in \left[\min\{\frac{3}{5}, \frac{3}{10}, \frac{4}{5}, \frac{4}{10}\}, \max\{\frac{3}{5}, \frac{3}{10}, \frac{4}{5}, \frac{4}{10}\}\right] = [0.3, 0.8].$$

Rule 2.3.7. *Assume that* $\otimes \in [a, b]$, $a < b$, *and* k *is a positive real number. The scalar multiplication of* k *and* $\otimes$ *is defined as follows.*

$$k \cdot \otimes \in [ka, kb].$$

Theorem 2.3.1. *Interval grey numbers cannot in general be cancelled additively or multiplicatively. More specifically, the difference of any two grey numbers is generally not zero, except in the case that they are identical. And the division of any two grey numbers is generally not 1 except in the case when they are identical.*

Example 2.3.6. For $\otimes \in [2, 5]$, we have

$$\otimes - \otimes \begin{cases} = 0, & \text{if synchronous} \\ \in [-3, 3], & \text{if not synchronous} \end{cases}$$

and

$$\frac{\otimes}{\otimes} \begin{cases} = 1, & \text{if synchronous} \\ \in \left[\frac{2}{5}, \frac{5}{2}\right] & \text{if not synchronous} \end{cases}.$$

Definition 2.3.1. *Assume that* $R(\otimes)$ *is a set of grey numbers. If for any* $\otimes_i, \otimes_j \in R(\otimes)$,

$$\otimes_i + \otimes_j,\ \otimes_i - \otimes_j,\ \otimes_i \cdot \otimes_j,\ \otimes_i / \otimes_j$$

all belong to $R(\otimes)$ *(when division is considered, the conditions in Rule 2.3.6 need to be satisfied), then* $R(\otimes)$ *is called a field of grey numbers.*

Theorem 2.3.2. *The totality of all interval grey numbers constitutes a field.*

Definition 2.3.2. *Assume that* $E(\otimes)$ *is a set of grey numbers. If for any* $\otimes_i$, $\otimes_j$ *and* $\otimes_k \in E(\otimes)$, *the following holds true,*

1. $\otimes_i + \otimes_j = \otimes_j + \otimes_i$;
2. $(\otimes_i + \otimes_j) + \otimes_k = \otimes_i + (\otimes_j + \otimes_k)$;
3. *There exists a zero element* $0 \in E(\otimes)$ *such that* $\otimes_i + 0 = \otimes_i$;
4. *For any* $\otimes \in E(\otimes)$, *there exists* $-\otimes \in E(\otimes)$ *such that*

$$\otimes + (-\otimes) = 0;$$

5. $\otimes_i \cdot (\otimes_j \cdot \otimes_k) = (\otimes_i \cdot \otimes_j) \cdot \otimes_k$;
6. *There exists a unit element* $1 \in E(\otimes)$ *such that* $1 \cdot \otimes_i = \otimes_i \cdot 1 = \otimes_i$;
7. $(\otimes_i + \otimes_j) \cdot \otimes_k = \otimes_i \cdot \otimes_k + \otimes_j \cdot \otimes_k$; *and*
8. $\otimes_i \cdot (\otimes_j + \otimes_k) = \otimes_i \cdot \otimes_j + \otimes_i \cdot \otimes_k$,

then $E(\otimes)$ *is called a grey linear space.*

Theorem 2.3.3. *The totality of all synchronous interval grey numbers constitutes a grey linear space.*

2.4 Measures of Grey Numbers

In this section, we discuss the concept of measures or greyness of grey numbers. The greyness of a grey number to a certain degree reflects how much the researcher does not know about the behavioral characteristics of the grey system of concern. In an earlier section, we have learned the definition of greyness of interval grey numbers such that their weight functions of whitenization are known. However, in many practical applications, we will often face grey numbers with unknown weight functions of whitenization. For example, it is extremely difficult for the researcher to construct the weight functions of whitenization for all such grey numbers that are made up of predictions about behavioral characteristics of a general grey system.

After carefully checking through various cases, it can be seen that the greyness of a grey number is mainly related to the length of the information field on that the grey number is defined, and its basic value. For example, let us consider a grey number near 4000. If we construct two grey number estimations $\otimes_1 \in [3998, 4002]$ and $\otimes_2 \in [3900, 4100]$, then it is obvious that $\otimes_1$ is a more valuable estimate than $\otimes_2$. That is, the greyness of $\otimes_1$ is smaller than that of $\otimes_2$. Now, let us consider another grey number with its basic value at 4. If we construct a grey number estimate $\otimes_3 \in [2, 6]$ for this second unknown number, we can easily see that the lengths of the information fields for $\otimes_1$ and $\otimes_3$ are both 4. However, the greyness of $\otimes_1$ is obviously smaller than that of $\otimes_3$.

Definition 2.4.1. *Assume that a grey number* $\otimes$ *is defined on the information field* $[a, b]$. *That is,* $\otimes \in [a, b]$, $a < b$. *Then,* $\ell(\otimes) = |b - a|$ *is called the length of the information field of* $\otimes$.

It can be seen readily that when two grey numbers have the same basic value, the grey number with the longer information field will have a greater greyness than the other grey number with a shorter information field.

Definition 2.4.2. *Assume that a given grey number $\otimes$ is defined on the information field $[a, b]$; that is, $\otimes \in [a, b]$, $a < b$.*

1. When the weight function of whitenization of $\otimes$ is known, $\widehat{\otimes} = E(\otimes)$ is called the mean-value whitenization (number) of the grey number $\otimes$, where $E(\otimes)$ stands for the expected value of the grey number $\otimes$, if the grey number is a random variable.

2. When the weight function of whitenization of the grey number $\otimes$ is unknown, (i) if $\otimes$ is a continuous grey number, then $\widehat{\otimes} = \frac{1}{2}(a + b)$ is called the mean-value whitenization (number) of the grey number $\otimes$. (ii) If $\otimes$ is a discrete grey number such that $a_i \in [a, b]$, for $i = 1, 2, ...$, are all the possible values of $\otimes$, then

$$\widehat{\otimes} = \begin{cases} \frac{1}{n}\sum_{i=1}^{n} a_i, & \otimes \text{has a finite number of possible values} \\ \lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} a_i, & \otimes \text{takes a countable number of possible values} \end{cases}$$

is called the mean-value whitenization (number) of the grey number $\otimes$.

Note. If $a_i(\otimes)$ is also a grey number such that $a_i(\otimes) \in [a_i, b_i]$ with $a_i < b_i$, then one can take $a_i = \widehat{a}_i(\otimes)$.

We continue to use $g^o(\otimes)$ to represent the greyness of the grey number $\otimes$. In the following, we establish an axiomatic system for the concept of greyness of grey numbers.

Axiom 2.5.1. *For any grey number $\otimes \in [a, b]$, $a < b$, $g^o(\otimes) \geq 0$.*

Axiom 2.5.2. *When $a = b$, that is, when $\ell(\otimes) = 0$, $g^o(\otimes) = 0$; that is, the greyness of the grey number $\otimes$ is zero.*

Axiom 2.5.3. *When either $a \to -\infty$ or $b \to +\infty$, $g^o(\otimes) \to \infty$; that is, the greyness approaches ∞.*

Axiom 2.5.4. $g^o(k\otimes) = g^o(\otimes)$.

This axiom says that when a grey number $\otimes$ is multiplied by a real number k, the greyness of the original grey number $\otimes$ is not changed.

Axiom 2.5.5. *$g^o(\otimes)$ is directly proportional to $\ell(\otimes)$, and inversely proportional to $\widehat{\otimes}$.*

Definition 2.4.3. *Assume that a grey number $\otimes \in [a, b]$, $a < b$, is given. Then,*

$$g^o(\otimes) = \frac{\ell(\otimes)}{|\widehat{\otimes}|} \tag{2.1}$$

is called the greyness of the grey number $\otimes$, where $\ell(\otimes)$ stands for the length of the information field of $\otimes$ and $\widehat{\otimes}$ the mean-value whitenization value.

Theorem 2.4.1. *The greyness, as defined in Definition 2.4.3, of grey numbers satisfies the five axioms established for that of grey numbers in Axioms 2.5.1 to 2.5.5.*

Definition 2.4.4. *When $\widehat{\otimes} = 0$, $\otimes$ is called a grey number with center zero.*

Proposition 2.4.1. *The greyness of any grey number with center zero equals ∞.*

In the following, we look at some results regarding the relationship between combinations of grey numbers and their greyness.

Theorem 2.4.2. *For given grey numbers $\otimes_1 \in [a, b]$ and $\otimes_2 \in [c, d]$ satisfying $a < b$, $c < d$, and either $a \geq 0$, $c \geq 0$ or $b \leq 0$ and $d \leq 0$, the following holds true:*

$$g^o(\otimes_1 + \otimes_2) \leq g^o(\otimes_1) + g^o(\otimes_2). \tag{2.2}$$

Proof.

$$\begin{aligned} g^o(\otimes_1 + \otimes_2) &= \frac{\ell(\otimes_1 + \otimes_2)}{|(\otimes_1 + \otimes_2)|} \\ &= \frac{2|b+d-a-c|}{|b+d+a+c|} \\ &\leq \frac{2|b-a| + |d-c|}{|b+d+a+c|} \\ &\leq \frac{2|b-a|}{|b+a|} + \frac{2|d-c|}{|d+c|} \\ &= \frac{\ell(\otimes_1)}{|\widehat{\otimes}_1|} + \frac{\ell(\otimes_2)}{|\widehat{\otimes}_2|} \\ &= g^o(\otimes_1) + g^o(\otimes_2). \ \square \end{aligned}$$

Theorem 2.4.3. *If the grey numbers $\otimes_1 \in [a, b]$ and $\otimes_2 \in [c, d]$ with $a < b$, $c < d$ satisfy one of the following conditions.*

1. $a \geq 0$ and $c \geq 0$,

2. $a \geq 0$, $d \leq 0$,

3. $b \leq 0$, $c \geq 0$, or

4. $b \leq 0$, $d \leq 0$,

then the following hold true,

$$g^o(\otimes_1 \cdot \otimes_2) \geq g^o(\otimes_1), \tag{2.3}$$

$$g^o(\otimes_1 \cdot \otimes_2) \geq g^o(\otimes_2). \tag{2.4}$$

Proof. If condition 1 holds true, we have $\otimes_1 \cdot \otimes_2 \in [ac, bd]$. If condition 2 holds true, we have $\otimes_1 \cdot \otimes_2 \in [bc, ad]$. If condition 3 holds true, we have $\otimes_1 \cdot \otimes_2 \in [ad, bc]$. If condition 4 holds true, we have $\otimes_1 \cdot \otimes_2 \in [bd, ac]$. In the following, we provide a detailed proof for condition 1. The proofs for the other cases are similar and omitted. Now, if condition 1 holds true, we have

$$\begin{aligned} g^o(\otimes_1 \cdot \otimes_2) &= \frac{\ell(\otimes_1 \cdot \otimes_2)}{|(\otimes_1 \cdot \otimes_2)|} \\ &= \frac{2|bd - ac|}{|bd + ac|} \\ &\geq \frac{2|bd - ad|}{|bd + ad|} \\ &= \frac{2|b - a|}{|b + a|} \\ &= \frac{\ell(\otimes_1)}{\left|(\widehat{\otimes}_1)\right|} \\ &= g^o(\otimes_1). \end{aligned}$$

Similarly, we have

$$\begin{aligned} g^o(\otimes_1 \cdot \otimes_2) &= \frac{\ell(\otimes_1 \cdot \otimes_2)}{|(\otimes_1 \cdot \otimes_2)|} \\ &= \frac{2|bd - ac|}{|bd + ac|} \\ &\geq \frac{2|bd - bc|}{|bd + bc|} \\ &= \frac{2|d - c|}{|d + c|} \\ &= \frac{\ell(\otimes_2)}{\left|(\widehat{\otimes}_2)\right|} \\ &= g^o(\otimes_2). \ \square \end{aligned}$$

Theorem 2.4.4. *If the grey numbers $\otimes_1 \in [a, b]$ and $\otimes_2 \in [c, d]$ with $a < b$, $c < d$ satisfy one of the following conditions,*

1. *$a > 0$ and $cd < 0$;*
2. *$b \leq 0$, $cd < 0$;*
3. *$ab < 0$, $cd < 0$, and $\frac{|a|}{b} \geq \max\left\{\frac{|c|}{d}, \frac{d}{|c|}\right\}$; or*
4. *$ab < 0$, $cd < 0$, and $\frac{b}{|a|} \geq \max\left\{\frac{|c|}{d}, \frac{d}{|c|}\right\}$,*

then the following holds true,

$$g^o(\otimes_1 \cdot \otimes_2) = g^o(\otimes_2). \tag{2.5}$$

Proof. For the cases 1 and 4, we have $\otimes_1 \cdot \otimes_2 \in b[c, d]$. For the cases 2 and 3, we have $\otimes_1 \cdot \otimes_2 \in a[c, d]$. So, for all these four cases, we have $\otimes_1 \cdot \otimes_2 = k \otimes_2$. Therefore,

$$g^o(\otimes_1 \cdot \otimes_2) = g^o(k\otimes_2) = g^o(\otimes_2). \ \square$$

Similarly, one can prove the following Theorem 2.4.5.

Theorem 2.4.5. *If the grey numbers $\otimes_1 \in [a,b]$ and $\otimes_2 \in [c,d]$ with $a < b$ and $c < d$ satisfy one of the following conditions,*

1. $c > 0$ and $ab < 0$;

2. $d \leq 0$, $ab < 0$;

3. $ab < 0$, $cd < 0$, and $\frac{|c|}{d} \geq \max\left\{\frac{|a|}{b}, \frac{b}{|a|}\right\}$; or

4. $ab < 0$, $cd < 0$, and $\frac{d}{|c|} \geq \max\left\{\frac{|a|}{b}, \frac{b}{|a|}\right\}$,

then the following holds true,

$$g^o(\otimes_1 \cdot \otimes_2) = g^o(\otimes_1). \tag{2.6}$$

Theorem 2.4.6. *If the grey numbers $\otimes_1 \in [a,b]$ and $\otimes_2 \in [c,d]$ with $a < b$ and $c < d$ satisfy one of the following conditions,*

1. $a \geq 0$ and $c > 0$;

2. $a \geq 0$, $d < 0$;

3. $b \leq 0$, $c > 0$; or

4. $b \leq 0$, $d < 0$,

then the following hold true,

$$g^o(\otimes_1 \div \otimes_2) \geq g^o(\otimes_1), \tag{2.7}$$

$$g^o(\otimes_1 \div \otimes_2) \geq g^o(\otimes_2). \tag{2.8}$$

Proof. For case 1, we have $\otimes_1 \div \otimes_2 \in \left[\frac{a}{d}, \frac{b}{c}\right]$. For case 2, we have $\otimes_1 \div \otimes_2 \in \left[\frac{b}{d}, \frac{a}{c}\right]$. For case 3, we have $\otimes_1 \div \otimes_2 \in \left[\frac{a}{c}, \frac{b}{d}\right]$. And, for case 4, we have $\otimes_1 \div \otimes_2 \in \left[\frac{b}{c}, \frac{a}{d}\right]$. In the following, we focus on the proof of case 1^o. The proofs for the other cases are similar and omitted.

When condition 1 holds true, we have

$$\begin{aligned} g^o(\otimes_1 \div \otimes_2) &= \frac{\ell(\otimes_1 \div \otimes_2)}{|\otimes_1 \div \otimes_2|} \\ &= \frac{2\left|\frac{b}{c} - \frac{a}{d}\right|}{\left|\frac{b}{c} - \frac{a}{d}\right|} \\ &= \frac{2(bd - ac)}{|bd + ac|} \\ &= \frac{\ell(\otimes_1 \cdot \otimes_2)}{|(\otimes_1 \cdot \otimes_2)|} \\ &= g^o(\otimes_1 \cdot \otimes_2). \end{aligned}$$

Now, from Theorem 2.4.3, it follows that $g^o(\otimes_1 \div \otimes_2) \geq g^o(\otimes_1)$ and $g^o(\otimes_1 \div \otimes_2) \geq g^o(\otimes_2)$.□

Theorem 2.4.7. *If the grey numbers $\otimes_1 \in [a,b]$ and $\otimes_2 \in [c,d]$ with $a < b$ and $c < d$ satisfy one of the following conditions,* *1. $ab < 0$ and $c > 0$, or*

2. $ab < 0$, $d < 0$,

then the following hold true,

$$g^o(\otimes_1 \div \otimes_2) = g^o(\otimes_1). \tag{2.9}$$

Proof. For case 1, we have $\otimes_1 \div \otimes_2 \in \left[\frac{a}{c}, \frac{b}{c}\right]$. For case 2, we have $\otimes_1 \div \otimes_2 \in \left[\frac{b}{d}, \frac{a}{d}\right]$. So, in both of these cases, we have

$$\otimes_1 \div \otimes_2 = k \otimes_1 .$$

Therefore,

$$g^o(\otimes_1 \div \otimes_2) = g^o(k\otimes_1) = g^o(\otimes_1). \square$$

Theorem 2.4.8. *If the grey numbers $\otimes_1 \in [a, b]$ and $\otimes_2 \in [c, d]$ with $a < b$, $c < d$ and $cd < 0$, then*

$$g^o(\otimes_1 \div \otimes_2) = \infty. \tag{2.10}$$

Proof. 1. When $a \geq 0$, we have $\otimes_1 \div \otimes_2 \in \left(-\infty, \frac{a}{c}\right] \cup \left[\frac{a}{d}, +\infty\right)$. 2. When $b \leq 0$, we have $\otimes_1 \div \otimes_2 \in \left(-\infty, \frac{b}{d}\right] \cup \left[\frac{b}{c}, +\infty\right)$. And, 3. when $ab < 0$, we have $\otimes_1 \div \otimes_2 \in (-\infty, +\infty)$. That is, $\otimes_1 \div \otimes_2$ is either a black number or a combination of two black numbers. Therefore, $g^o(\otimes_1 \div \otimes_2) = \infty$. $\square$

2.5 Information Content of Grey Numbers

Materials, energies, and information have been seen as three elementary bases for the evolution of the natural world and the development of human society. Our ability to understand nature increases rapidly with our capability to better comprehend and control information. During the Civil War of the United States of America, people used telegrams to transmit information. The speed of information transmission was 30 words per minute. At the time, in order to control an area of the size of 10 square kilometers, 38,830 soldiers were needed. During World War I of the last century, the functions of telegrams had been improved so that the speed and accuracy of information transmission had reached a new level. During that period of time, to control an area of the size of 10 square kilometers, 4040 soldiers were needed. During World War II, the technology of the telex was so well understood that 66 words could be transmitted per minute. And, to control an area of 10 square kilometers, 360 soldiers were needed. During the Gulf War in 1991, the coalition forces used computers to transmit information reaching a speed as high as 192,000 words per minute. And, only 23.4 soldiers were needed to control an area of 10 square kilometers. It is estimated that by 2010, with the improvement of computers, the speed of information transmission can reach 1,500 billion words a minute. And, 2.4 soldiers will be able to control an area of 10 square kilometers.

In 1948, C.E. Shannon introduced the following formula for the computation of information measures on the space made up of random discrete systems,

$$I = -\sum_{i=1}^{n} P_i \log P_i \text{ where } \sum_{i=1}^{n} P_i = 1,$$

that is generally known as *Shannon* (information) *entropy*. In 1996, Qishan Zhang introduced the concept of (information) entropy for the difference information sequence $X = (x_1, x_2, ..., x_s)$ using structural image sequence $Y = (y_1, y_2, ..., y_s)$:

$$I(X) = -\sum_{j=1}^{s} y_j \cdot \ln y_j. \tag{2.11}$$

The concept of grey numbers is established to express characteristic behaviors of grey systems. The information content contained in grey numbers reflects the degree of comprehension of the researcher of a specific grey system. The measure of the information content cannot be unrelated to the background on that relevant grey numbers were initially created. For example, if there is no explanation on the background and field of definition of a grey number and no explanation on the grey system about that the grey number was initially introduced, it will be very difficult for us to discuss the amount of information content contained in the grey number. For example, for the given grey number $\otimes \in [160, 200]$, without knowing the background where it was initially introduced, it is very difficult for anyone to tell the amount of useful information the grey number actually carries. When we know that the grey number stands for an estimate of a male adult's height in centimeters, its information content is still nearly zero, because the interval $[160, 200]$ coincides almost entirely with the range of male adult heights. If a criminal is sought by a law enforcement agent, and the only known information is that the criminal is a male as tall as somewhere between 160 to 200 centimeters, then the law officer will surely need additional information to break the case. On the other hand, if $\otimes_1 \in [160, 200]$ stands for the blood pressure of a patient, it surely provides quite a bit of useful information for the patient and his doctor.

Definition 2.5.1. *Assume that the background of introduction of a grey number $\otimes$ is Ω with $\otimes \subset \Omega$. Then $\overline{\otimes} = \Omega - \otimes$ is called the remanent set of $\otimes$.*

Let $\mu(\otimes)$ be the measure of the field on that the grey number $\otimes$ is defined and $I(\otimes)$ the information content of the grey number. Then, $I(\otimes)$ satisfies the following axioms.

Axiom 2.5.1. $0 \leq I(\otimes) \leq 1$.

Axiom 2.5.2. $I(\Omega) = 0$.

Axiom 2.5.3. *$I(\otimes)$ is directly proportional to $\mu(\overline{\otimes})$ and inversely proportional to $\mu(\Omega)$.*

At this junction, Axiom 2.5.1 limits the information content of a grey number to the range $[0,1]$. The closer to zero $I(\otimes)$ is, the less information content is contained in the grey number $\otimes$. On the other hand, the closer to 1 $I(\otimes)$ is, the more information content the grey number $\otimes$ contains. Axiom 2.5.2 stipulates that the information content of the background on that a grey number was initially introduced is zero. That is because in general the background is commonly known to people and covers the entire field on that the grey number is defined. That is why knowledge of the background Ω does not provide much, if any at all, useful information to the researcher. For example, the proposition that "a train is able to pull more than zero pounds" does not provide much useful information, because $\Omega = (0, +\infty)$ represents the background of all possible weights. Axiom 2.5.3 states that when the background Ω is fixed, the larger the measure $\mu(\overline{\otimes})$ of the remanent set $\overline{\otimes}$, the larger is the information content contained in the grey number $\otimes$. That is, the smaller the measure of the grey number $\otimes$ itselfis, the larger its information content. For example, if a grey number $\otimes$ stands for an estimate for a specific real number value, then when the reliability is fixed, the smaller the measure of $\otimes$, the more meaningful an estimate the grey number $\otimes$ represents.

Definition 2.5.2. *Assume that a grey number $\otimes$ is initially introduced on the background Ω. Then*

$$I(\otimes) = \frac{\mu(\overline{\otimes})}{\mu(\Omega)} \tag{2.12}$$

is called the information content contained in the grey number $\otimes$.

Theorem 2.5.1. *The concept of information content of grey numbers, as defined in Definition 2.5.2, satisfies Axioms 2.5.1 to 2.5.3, introduced for the measurement of information content of grey numbers.*

Proof. 1. From the fact that $\overline{\otimes} \subset \Omega$ and properties of measures, it follows that $0 \leq \mu(\otimes) \leq \mu(\Omega)$. So, we have $0 \leq I(\otimes) \leq 1$.

2. When $\otimes = \Omega$, $\overline{\otimes} = \varnothing$. So, $\mu(\overline{\otimes}) = \mu(\varnothing) = 0$. That is, $I(\Omega) = 0$.

3. is obvious. □

Theorem 2.5.2. *If $\otimes_1 \subset \otimes_2$, then $I(\otimes_1) \geq I(\otimes_2)$.*

Proof. From the assumption that $\otimes_1 \subset \otimes_2$ and properties of measures, it follows that $\mu(\otimes_1) \leq \mu(\otimes_2)$. So,

$$\begin{aligned} \mu(\otimes_1) &= \mu(\Omega) - \mu(\otimes_1) \\ &\geq \mu(\Omega) - \mu(\otimes_2) = \mu(\overline{\otimes}_2). \end{aligned}$$

Therefore, we have

$$I(\otimes_1) \geq I(\otimes_2). \ \square$$

Because grey numbers can be combined, it is necessary for us to study the information content of combined grey numbers.

Definition 2.5.3. *For grey numbers $\otimes_1 \in [a,b]$ and $\otimes_2 \in [c,d]$ satisfying $a < b$ and $c < d$,*

$$\otimes_1 \cup \otimes_2 = \{\xi |\ \xi \in [a,b] \, or \, \xi \in [c,d]\} \tag{2.13}$$

is called the union of the grey numbers $\otimes_1$ and $\otimes_2$.

The concept of union of grey numbers is similar to putting a group of grey numbers together without any specific order. The consequence of piling grey numbers together in this fashion is that the more grey numbers one piles together, the weaker the total information becomes.

Theorem 2.5.3. $I\left(\otimes_1 \cup \otimes_2\right) \leq I\left(\otimes_k\right),\ k = 1, 2.$

Proof. From the fact that $\otimes_1 \cup \otimes_2 \supset \otimes_k$, $k = 1, 2$, and Theorem 2.5.2, Theorem 2.5.3 follows readily. □

Definition 2.5.4. *For grey numbers $\otimes_1 \in [a,b]$ and $\otimes_2 \in [c,d]$ satisfying $a < b$ and $c < d$,*

$$\otimes_1 \cap \otimes_2 = \{\xi |\ \xi \in [a,b] \, and \, \xi \in [c,d]\} \tag{2.14}$$

is called the intersection of the grey numbers $\otimes_1$ and $\otimes_2$.

The intersection of grey numbers is like a combination of several grey numbers through synthetic organization so that useful information can be extracted and understanding about the background of the grey numbers can be deepened.

Theorem 2.5.4. $I\left(\otimes_1 \cap \otimes_2\right) \geq I\left(\otimes_k\right),\ k = 1, 2.$

Proof. From Theorem 2.5.2 and the fact that $\otimes_1 \cap \otimes_2 \subset \otimes_k$, $k = 1, 2$, it follows that Theorem 2.5.4 holds true. □

Theorem 2.5.5. *If $\otimes_1 \subset \otimes_2$, then $I\left(\otimes_1 \cup \otimes_2\right) = I\left(\otimes_2\right)$ and $I\left(\otimes_1 \cap \otimes_2\right) = I\left(\otimes_1\right)$.*

Proof. From the assumption that $\otimes_1 \subset \otimes_2$, it follows that

$$\otimes_1 \cup \otimes_2 = \otimes_2$$

and

$$\otimes_1 \cap \otimes_2 = \otimes_1.$$

Therefore,

$$I\left(\otimes_1 \cup \otimes_2\right) = I\left(\otimes_2\right)$$

and

$$I\left(\otimes_1 \cap \otimes_2\right) = I\left(\otimes_1\right). \ \square$$

When two grey numbers $\otimes_1$ and $\otimes_2$ are independent of the measure μ, we have the following more satisfactory result.

Theorem 2.5.6. *If $\mu(\Omega) = 1$ and the grey numbers $\otimes_1$ and $\otimes_2$ are independent of the measure μ, then the following hold true,*

1. $I(\otimes_1 \cup \otimes_2) = I(\otimes_1) I(\otimes_2)$; *and*

2. $I(\otimes_1 \cap \otimes_2) = I(\otimes_1) + I(\otimes_2) - I(\otimes_1) I(\otimes_2)$.

Proof. 1.From the assumption that $\mu(\Omega) = 1$, it follows that

$$\begin{aligned} I(\otimes_1 \cup \otimes_2) &= \mu\left(\overline{\otimes_1 \cup \otimes_2}\right) \\ &= \mu\left(\overline{\otimes}_1 \cap \overline{\otimes}_2\right) \\ &= \mu\left(\overline{\otimes}_1\right) \mu\left(\overline{\otimes}_2\right) \\ &= I(\otimes_1) I\otimes_2). \end{aligned}$$

2.

$$\begin{aligned} I(\otimes_1 \cap \otimes_2) &= \mu\left(\overline{\otimes_1 \cap \otimes_2}\right) \\ &= \mu\left(\overline{\otimes}_1 \cup \overline{\otimes}_2\right) \\ &= \mu\left(\overline{\otimes}_1\right) + \mu\left(\overline{\otimes}_2\right) - \mu\left(\overline{\otimes}_1 \cap \overline{\otimes}_2\right) \\ &= \mu\left(\overline{\otimes}_1\right) + \mu\left(\overline{\otimes}_2\right) - \mu\left(\overline{\otimes}_1\right) \mu\left(\overline{\otimes}_2\right) \\ &= I(\otimes_1) + I(\otimes_2) - I(\otimes_1) I(\otimes_2). \ \square \end{aligned}$$

For example, let us consider rolling a die and recording the number of dots showing upward. In this case,

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Assume that we have two grey numbers $\otimes_1 \in \{1, 2\}$ and $\otimes_2 \in \{2, 3, 4\}$ and μ is the probability measure. Then,

$$\mu(\otimes_1) = \frac{1}{3},\ \mu(\otimes_2) = \frac{1}{2}, \text{and } \mu(\otimes_1 \cap \otimes_2) = \frac{1}{6}$$

satisfy the condition of independence. Also, for this case, we have the following

$$I(\otimes_1 \cup \otimes_2) = \frac{1}{3} = I(\otimes_1) I(\otimes_2)$$

and

$$I(\otimes_1 \cap \otimes_2) = \frac{5}{6} = I(\otimes_1) + I(\otimes_2) - I(\otimes_1) I(\otimes_2),$$

that coincide with the results in Theorem 2.5.6.

Theorems 2.5.3 to 2.5.6 can all be generalized to the cases of union and intersection of n grey numbers, where n can be any arbitrary natural number.

Different ways on that grey numbers are combined affect the information contents and the reliability of the information content of the consequent grey numbers. In general, when grey numbers are unioned, the consequent information content decreases and the reliability of the information content increases. On the other hand, when grey numbers are intersected, the information content increases and its reliability decreases. When one faces a practical problem with the necessity to process a large number of grey numbers, she can consider combining these grey numbers at different levels so that useful information can be extracted through each level. In the process of combining the available grey numbers, she can also apply the concepts of union and intersection across different levels so that the final extracted information can satisfy her desires in terms of reliability and content.

3

Grey Equations and Grey Matrices

3.1 Grey Algebraic Equations and Grey Differential Equations

Because applications of grey systems theory in general deal with relations between grey variables and numbers, in this section, we focus on such basic concepts as various grey equations.

Definition 3.1.1. *Algebraic equations with grey coefficient(s) are called grey algebraic equations.*

Definition 3.1.2. *An n-dimensional vector with grey component(s) is called a grey n-dimensional vector.*

An n-dimensional grey vector is denoted as

$$X(\otimes) = (\otimes_1, \otimes_2, ..., \otimes_n)$$

In general, white equations containing no grey elements can be solved relatively more easily. For example, the solution of a white linear equation in one variable is a fixed point on the real number line. And, the solution of a system of white equations of multivariables also has definite expressions. However, it is relatively complicated to discuss the solution of a grey equation. Strictly speaking, a grey equation is not simply one equation, but a symbol for many equations. The number of equations, represented by a grey equation, is determined by how the grey elements in the equation take values. If all the grey elements take a finite number of values in a bounded grey field, then the grey equation represents a finite number of white equations. If the grey elements in the grey equation take an infinite number of values, then the grey equation represents an infinite number of white equations. For example, the solution of the grey equation

$$\otimes_1 x + \otimes_2 = 0$$

is $x = -\otimes_2 \cdot \otimes_1^{-1}$. When $\otimes_1 \in \{1, 2, 4\}$ and $\otimes_2 = 1$, the solution set is

$$X \in \{-0.25, -0.5, -1\}.$$

When $\otimes_1 \in [1, 4]$ and $\otimes_2 = 1$, the solution set is

$$X \in [-1, -0.25]\,.$$

Different values of the grey coefficients correspond to different values in the set of solutions. In general, the solution set of a linear grey equation in one variable consists of a few points or a grey interval on the real number line. And, each solution of a system of linear grey equations in n variables is an n-dimensional grey vector. The forms of solutions of other types of grey algebraic equations are also similar to those of the corresponding white algebraic equations.

Definition 3.1.3. *Grey equations with grey derivative(s) or grey differential(s) are called grey differential equations.*

Definition 3.1.4. *The differential equations, containing grey coefficient(s) and white derivative(s), are called white differential equations with grey coefficients.*

Grey derivatives and grey differential equations constitute the foundation for grey systems modeling. We study more detailed concepts and related problems in Chapter 7. In general, white differential equations with grey coefficients can be solved by going through integration or using characteristic equations. All details are omitted here.

3.2 Grey Matrices and Their Operations

When a huge amount of values and variables are involved in a study, one will have to make use of the concepts of matrices and related results.

Definition 3.2.1. *Matrices, containing grey entries, are called grey matrices, denoted as $A(\otimes)$. And, the grey entry at the location of the ith row and the jth column is denoted as $\otimes_{ij}$ or $\otimes(i, j)$.*

For example,

$$A(\otimes) = \begin{bmatrix} \otimes_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

represents a 2×2 grey matrix with $\otimes_{11} \in [\underline{a}_{11}, \overline{a}_{11}]$, $\underline{a}_{11} < \overline{a}_{11}$; the other entries a_{12}, a_{21}, and a_{22} are all white numbers.

Definition 3.2.2. *Assume that $A(\otimes)$ is an $n\times\ m$ grey matrix; G stands for the number of grey entries contained in the matrix $A(\otimes)$. We then call*

$$\theta = \frac{G}{mn-G}$$

the absolute grey degree of entries of the grey matrix $A(\otimes)$, and

$$\omega = \frac{G}{mn}$$

the relative grey degree of entries of the grey matrix $A(\otimes)$.

Each relative grey degree takes values in the interval $[0,1]$. When $\omega = 1$, the matrix $A(\otimes)$ contains no white entry; when $\omega = 0, A(\otimes)$ is a white matrix. Both grey degrees can be used to indicate the numbers of grey entries appearing in the matrix $A(\otimes)$.

The symbol $G^{m\times n}$ is used to represent the set of all $m\times n$ grey matrices. An $m \times n$ grey matrix $A(\otimes)$ with the $(i,\, j)$ entry $\otimes_{ij}$ is written as

$$A(\otimes) = [\otimes_{ij}]_{m\times n}\,.$$

Two grey matrices, satisfying certain conditions, can be added, subtracted, and multiplied, etc. In the rest of this section, we assume that all grey matrices contain interval grey numbers only as entries.

Definition 3.2.3. *Assume that*

$$A(\,\otimes\,) = \, [\otimes_{ij}]_{m\times n}$$

and

$$B(\,\otimes\,) = \, \left[\otimes'_{ij}\right]_{m\times n}$$

are two given grey matrices. If all the corresponding entries of $A(\otimes)$ and $B(\otimes)$ are identical, that is,

$$\otimes_{ij} = \otimes'_{ij},$$

$i = 1,2,...,m, j = 1,2,...,n$, then the matrices $A(\otimes)$ and $B(\otimes)$ are said to be equal, denoted $A(\otimes)\ = B(\otimes)$.

Definition 3.2.4. *If $A(\otimes)$ and $B(\otimes)$ are the same as defined in Definition 3.2.3, then*

$$A(\otimes) + B(\otimes) = \left[\otimes_{ij} + \otimes'_{ij}\right]_{m\times n}$$

is called the sum of $A(\otimes)$ and $B(\otimes)$, and

$$A(\otimes) - B(\otimes) = \left[\otimes_{ij} - \otimes'_{ij}\right]_{m\times n}$$

the difference of $A(\otimes)$ *and* $B(\otimes)$*, and*

$$-A(\otimes) = [-\otimes_{ij}]_{m\times n}$$

the additive inverse of $A(\otimes)$.

Proposition 3.2.1. *The addition and subtraction of grey matrices satisfy the following,*

1. $A(\otimes) + B(\otimes) = B(\otimes) + A(\otimes)$;
2. $(A(\otimes) + B(\otimes)) + C(\otimes) = A(\otimes) + (B(\otimes) + C(\otimes))$;
3. $A(\otimes) - B(\otimes) = A(\otimes) + (-B(\otimes))$.

Definition 3.2.5. *Assume that* $\otimes$ *is a grey number and*

$$A(\otimes) = [\otimes_{ij}]_{m\times n}$$

is given. Then the following

$$\otimes\cdot A(\otimes) = [\otimes\cdot\otimes_{ij}]_{m\times n}$$

is called the scalar multiplication of the grey number $\otimes$ *and the grey matrix* $A(\otimes)$.

Proposition 3.2.2. *Scalar multiplication of grey numbers and matrices satisfies the following,*

1. $(\otimes_1\cdot\otimes_2)\cdot A(\otimes) = \otimes_1\cdot(\otimes_2\cdot A(\otimes)) = \otimes_2\cdot(\otimes_1\cdot A(\otimes))$;
2. $(\otimes_1+\otimes_2)\cdot A(\otimes) = \otimes_1\cdot A(\otimes) + \otimes_2\cdot A(\otimes)$;
3. $\otimes\cdot(A(\otimes)\pm B(\otimes)) = \otimes\cdot A(\otimes)\pm\otimes B(\otimes)$.

When $\otimes = -1$, $\otimes\cdot A(\otimes) = -A(\otimes)$ *is the additive inverse of* $A(\otimes)$.

Definition 3.2.6. *Assume that two grey matrices*

$$A(\otimes) = [\otimes_{ij}]_{m\times s}$$

and

$$B(\otimes) = \left[\otimes'_{ij}\right]_{s\times n}$$

are given. Then the following matrix

$$A(\otimes)\cdot B(\otimes) = \left[\otimes''_{ij}\right]_{m\times n},$$

where

$$\begin{aligned}\otimes''_{ij} &= \otimes_{i1}\otimes'_{1j} + \otimes_{i2}\otimes'_{2j} + \otimes_{is}\otimes'_{sj}\\ &= \sum_{k=1}^{s}\otimes_{ik}\otimes'_{kj},\end{aligned}$$

$i = 1, 2, ..., m; j = 1, 2, ..., n$, *is called the product of the grey matrices* $A(\otimes)$ *and* $B(\otimes)$. *It can be shown that the product of an* $m\times s$ *grey matrix and an* $s\times n$ *grey matrix is an* $m\times n$ *grey matrix.*

It needs to be seen that only when the number of columns of the former grey matrix equals the number of rows of the latter grey matrix, these two matrices can be multiplied. The product of a $1 \times s$ grey matrix and an $s \times 1$ grey matrix is a grey number. Even when both $A(\otimes) \cdot B(\otimes)$ and $B(\otimes) \cdot A(\otimes)$ are meaningful, in general, they are not equal. That is to say, the multiplication of grey matrices is not commutative.

Proposition 3.2.3. *When all operations involved are well defined, the multiplication of grey matrices satisfies the following,*

1. $(A(\otimes) \cdot B(\otimes)) \cdot C(\otimes) = A(\otimes) \cdot (B(\otimes) \cdot C(\otimes))$;

2. $A(\otimes) \cdot (B(\otimes) + C(\otimes)) = A(\otimes) \cdot B(\otimes) + A(\otimes) \cdot C(\otimes)$,

and $(A(\otimes) + B(\otimes)) \cdot C(\otimes) = A(\otimes) \cdot C(\otimes) + B(\otimes) \cdot C(\otimes)$;

3. $\otimes \cdot (A(\otimes) \cdot B(\otimes)) = (\otimes \cdot A(\otimes)) \cdot B(\otimes) = A(\otimes) \cdot (\otimes \cdot B(\otimes))$.

Definition 3.2.7. *Assume that $A(\otimes)$ is a square grey matrix of size n,*

$$A^k(\otimes) = \underbrace{A(\otimes) \cdot A(\otimes) \cdot \quad \cdots \quad \cdot A(\otimes)}_{k\ times}$$

is called the kth power of the grey square matrix $A(\otimes)$.

Proposition 3.2.4. *The exponents (or powers) of a square grey matrix $A(\otimes)$ satisfy the following,*

1. $A^m(\otimes) \cdot A^n(\otimes) = A^{m+n}(\otimes)$;

2. $(A^m(\otimes))^n = A^{mn}(\otimes)$.

where m and n are positive integers.

Because the multiplication of grey matrices does not satisfy the law of commutativity, for $A(\otimes)$ and $B(\otimes) \in G^{n\times n}$, in general

$$(A(\otimes) \cdot B(\otimes))^m \neq A^m(\otimes) \cdot B^m(\otimes).$$

Definition 3.2.8. *Assume the grey matrix*

$$A(\otimes) = \begin{bmatrix} \otimes_{11} & \otimes_{12} & \cdots & \otimes_{1n} \\ \otimes_{21} & \otimes_{22} & \cdots & \otimes_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \otimes_{m1} & \otimes_{m2} & \cdots & \otimes_{mn} \end{bmatrix}$$

is given. The following grey matrix

$$A(\otimes)^T = \begin{bmatrix} \otimes_{11} & \otimes_{21} & \cdots & \otimes_{m1} \\ \otimes_{12} & \otimes_{22} & \cdots & \otimes_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ \otimes_{1n} & \otimes_{2n} & \cdots & \otimes_{mn} \end{bmatrix}$$

is called the transposition of the grey matrix $A(\otimes)$.

The (i, j) entry of the transposition of a grey matrix equals the (j, i) entry of the original grey matrix.

Proposition 3.2.5. *When all operations involved are well defined, the transposition of grey matrices satisfies the following,*

1. $(A(\otimes)^T)^T = A(\otimes)$;

2. $(A(\otimes) + B(\otimes))^T = A(\otimes)^T + B(\otimes)^T$;

3. $(\otimes \cdot A(\otimes))^T = \otimes \cdot A(\otimes)^T$;

4. $(A(\otimes) \cdot B(\otimes))^T = B(\otimes)^T \cdot A(\otimes)^T$.

3.3 Several Special Grey Matrices

Similar to the case of applying ordinary matrices, one also needs to know the properties of special matrices well.

Definition 3.3.1. *All grey matrices of the form*

$$A(\otimes) = \begin{bmatrix} \otimes_{11} & & & \\ & \otimes_{22} & & \\ & & \ddots & \\ & & & \otimes_{nn} \end{bmatrix}$$

are called diagonal grey matrices, where all unshown entries are assumed to be zero. Diagonal grey matrices are also written as $diag[\otimes_{11} \otimes_{22} \cdots \otimes_{nn}]$.

Proposition 3.3.1. *All diagonal grey matrices satisfy the following properties,*

1. The sum and difference of grey diagonal matrices of the same size are still diagonal.

2. The scalar product of a grey number and a grey diagonal matrix is still diagonal.

3. The multiplication of two diagonal grey matrices of the same size is still diagonal and is commutative.

4. Each grey diagonal matrix and its transpose are equal.

Definition 3.3.2. *All diagonal grey matrices with* $diag[1\ 1\ \cdots\ 1]$ *as their matrix of whitenization, are called unit grey matrices, denoted* $E(\otimes)$.

The normal unit matrix is written as E.

$$E = diag[1\ 1 \cdots\ 1].$$

Definition 3.3.3. $diag[\otimes \otimes \cdots \otimes]$ *is called a grey scalar matrix.*

Definition 3.3.4. *Grey matrices of the forms*

$$\begin{bmatrix} \otimes_{11} & \otimes_{12} & \cdot & \cdot & \cdot & \otimes_{1n} \\ & \otimes_{22} & \cdot & \cdot & \cdot & \otimes_{2n} \\ & & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot \\ & & & & & \otimes_{nn} \end{bmatrix}$$

and

$$\begin{bmatrix} \otimes_{11} & & & & & \\ \otimes_{21} & \otimes_{22} & & & & \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \otimes_{n1} & \otimes_{n2} & \cdot & \cdot & \cdot & \otimes_{nn} \end{bmatrix}$$

are called upper triangular and lower triangular grey matrices, respectively. And, they are all called triangular grey matrices.

Proposition 3.3.2. *The sum, difference, and product of two triangular grey matrices of the same type, are still a triangular matrix of the original type.*

Definition 3.3.5. *If $A(\otimes) \in G^{n\times n}$ satisfies $A(\otimes)^T = A(\otimes)$, then the grey matrix $A(\otimes)$ is called a symmetric grey matrix.*

It can be readily shown that the entries of a symmetric matrix satisfy the condition that

$$\otimes_{ij} = \otimes_{ji},$$

for $i, j = 1, 2, ..., n$.

Definition 3.3.6. *If $A(\otimes) \in G^{n\times n}$ satisfies the condition that*

$$-A(\otimes)^T = A(\otimes),$$

then $A(\otimes)$ is called a skew-symmetric grey matrix.

It can be shown easily that the entries of a skew-symmetric matrix satisfy the condition that

$$\otimes_{ij} = -\otimes_{ji},$$

for $i, j = 1, 2, ..., n$. Obviously, all the entries on the main diagonal must be zero.

Proposition 3.3.3. *The sum, difference, and scalar product of symmetric (resp., skew-symmetric) grey matrices are still symmetric (resp., skew-symmetric).* □

The product of two symmetric (resp., skew-symmetric) grey matrices might not be a symmetric (resp., skew-symmetric) grey matrix.

Definition 3.3.7. *If a grey matrix $A(\otimes) \in G^{n\times n}$ satisfies*

$$A(\otimes) \cdot A(\otimes)^T = A(\otimes)^T \cdot A(\otimes) = E(\otimes),$$

then $A(\otimes)$ is called an orthogonal grey matrix.

Proposition 3.3.4. *The product of two orthogonal grey matrices is still orthogonal.*

3.4 Singularities of Grey Matrices

When solving systems of equations, one will have to deal with the concept of singularity. In this section, we focus on this and relevant concepts.

Definition 3.4.1. *For a grey matrix $A(\otimes)\in G^{n\times n}$, if there exist grey matrices $B(\otimes)$ and $C(\otimes)\in G^{n\times n}$ such that*

$$A(\otimes)\cdot B(\otimes)=E(\otimes)$$

and

$$C(\otimes)\cdot A(\otimes)=E(\otimes),$$

then $B(\otimes)$ is called a grey right inverse matrix of $A(\otimes)$ and $C(\otimes)$ a grey left inverse matrix of $A(\otimes)$. When $B(\otimes)=C(\otimes)$, the matrix $A(\otimes)$ is said to be invertible and $B(\otimes)$ is called the inverse matrix of $A(\otimes)$, denoted as $A(\otimes)^{-1}$.

For a general grey matrix $A(\otimes)$, it is very difficult to find its inverse. In this section, we discuss singularities of grey matrices with the help of whitenization matrices of the grey matrix of interest.

Definition 3.4.2. *Assume $A(\otimes)\in G^{n\times n}$. The matrix, obtained from $A(\otimes)$ by whitenizing all grey entries of the matrix $A(\otimes)$, is called a whitenization (matrix) of $A(\otimes)$, denoted $\widetilde{A}(\otimes)=\left[\widetilde{\otimes}_{ij}\right]_{n\times n}$.*

The upper limit matrix $\overline{A}$ and the lower limit matrix $\underline{A}$ are special whitenization matrices. The set of all whitenization matrices of $A(\otimes)$ is written as $\left\{\widetilde{A}\right\}$.

Definition 3.4.3. *For $A(\otimes)\in G^{n\times n}$,*

1. if for any $\widetilde{A}\in\left\{\widetilde{A}\right\}$, it is always true that $\det\widetilde{A}=0$, then $A(\otimes)$ is called a singular grey matrix;

2. if for any $\widetilde{A}\in\left\{\widetilde{A}\right\}$, it is always true that $\det\widetilde{A}\neq 0$, then $A(\otimes)$ is called a non $-$ singular grey matrix.

Both singular and non-singular grey matrices are all called *grey matrices of determinable singularity*. Others are called *grey matrices of non-determinable singularity*.

Proposition 3.4.1. *For any $A(\otimes)\in G^{n\times n}$, $A(\otimes)$ is a grey matrix of non-determinable singularity, if, and only if, there exist $\widetilde{A}_1$ and $\widetilde{A}_2\in\left\{\widetilde{A}\right\}$ such that*

$$\det\widetilde{A}_1=0,\ \ \det\widetilde{A}_2\neq 0.$$

Example 3.4.1. Assume that

$$A(\otimes) = \begin{bmatrix} \otimes_{11} & 1 \\ 1 & 1 \end{bmatrix},$$

where $\otimes_{11} \in [-2, 1]$. Take $\widetilde{\otimes}_{11} = 1$ or $\widetilde{\otimes}_{11} = -1$; one can obtain, respectively,

$$\widetilde{A}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

and

$$\widetilde{A}_2 = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

So, it is easy to show that $\det\widetilde{A}_1 = 0$ and $\det\widetilde{A}_2 \neq 0$. Therefore, $A(\otimes)$ is a grey matrix of nondeterminable singularity.

Proposition 3.4.2. *Assume that $A(\otimes)$ is a triangular grey matrix. If all entries of the diagonal of $A(\otimes)$, $\otimes_{ii} \in [\underline{a}_{ii}, \overline{a}_{ii}]$ $(i = 1, 2, ..., n)$ satisfy $\underline{a}_{ii} \cdot \overline{a}_{ii} > 0$, then $A(\otimes)$ is a grey matrix of determinable singularity.*

Proposition 3.4.3. *Let $A(\otimes) \in G^{n\times n}$. If all cofactors of grey entries in $A(\otimes)$ equal zero, then $A(\otimes)$ is a grey matrix of determinable singularity.*

Example 3.4.2. Determine the singularity of the following grey matrices, assuming that $a_{ij} \neq 0$.

$$A_1(\otimes) = \begin{bmatrix} \otimes_{11} & 1 \\ 1 & \otimes_{22} \end{bmatrix}, \quad A_2(\otimes) = \begin{bmatrix} \otimes_{11} & 1 \\ 1 & 0 \end{bmatrix},$$

$$A_3(\otimes) = \begin{bmatrix} a_{11} & \otimes_{12} & 0 \\ \otimes_{21} & \otimes_{22} & a_{23} \\ a_{31} & \otimes_{32} & 0 \end{bmatrix}, \quad A_4(\otimes) = \begin{bmatrix} a_{11} & \otimes_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & \otimes_{32} & a_{33} \end{bmatrix}.$$

Solution. 1. For $A_1(\otimes)$, $\otimes_{11}$ and $\otimes_{22}$ are cofactors of each other and not equal to zero. So, the conditions in Proposition 3.4.3 are not satisfied. Consider

$$\det A_1(\otimes) = \begin{vmatrix} \otimes_{11} & 1 \\ 1 & \otimes_{22} \end{vmatrix} = \otimes_{11} \cdot \otimes_{22} - 1 = \otimes - 1.$$

When $\otimes_{ii}, i = 1, 2$, take different values, $\otimes$ will take different values. Hence $\det \widetilde{A}_1$ may be zero and may not be zero. So, $A_1(\otimes)$ is nondeterminably singular.

2. $A_2(\otimes)$ contains only one grey entry $\otimes_{11}$, whose cofactor is zero. From Proposition 3.4.3, $A_2(\otimes)$ is determinably singular. In fact,

$$\det A_2(\otimes) = \begin{vmatrix} \otimes_{11} & 1 \\ 1 & 0 \end{vmatrix} = -1,$$

for any $\widetilde{A}_2 \in \left\{\widetilde{A}_2\right\}$, $\det \widetilde{A}_2$ never equals zero. So $A_2(\otimes)$ is non-singular.

3. For $A_3(\otimes)$, because the cofactor of the grey entry $\otimes_{12}$

$$\begin{vmatrix} \otimes_{12} & a_{23} \\ a_{31} & 0 \end{vmatrix} = -a_{23}a_{31} \neq 0$$

does not satisfy the conditions in Proposition 3.4.3, consider

$$\begin{aligned} \det A_3(\otimes) &= -a_{23}\begin{vmatrix} a_{11} & \otimes_{12} \\ a_{31} & \otimes_{32} \end{vmatrix} \\ &= -a_{32}a_{11}\otimes_{32} + a_{23}a_{31}\otimes_{12}. \end{aligned}$$

So, for different $\widetilde{A}_3 \in \left\{\widetilde{A}_3\right\}$, $\det \widetilde{A}_3$ may be zero and/or may not be zero. Therefore, $A_3(\otimes)$ is non-determinably singular.

4. $A_4(\otimes)$ contains two grey entries, $\otimes_{12}$ and $\otimes_{32}$. The cofactor of $\otimes_{12}$

$$\begin{vmatrix} 0 & 0 \\ 0 & a_{33} \end{vmatrix} = 0,$$

and the cofactor of $\otimes_{32}$

$$\begin{vmatrix} a_{11} & 0 \\ 0 & 0 \end{vmatrix} = 0.$$

Therefore, $A_4(\otimes)$ is a grey matrix of determinable singularity. In fact, for any $\widetilde{A}_4 \in \left\{\widetilde{A}_4\right\}$, we always have

$$\det \widetilde{A}_4 = a_{11}a_{22}a_{33} \neq 0.$$

So, $A_4(\otimes)$ is nonsingular.

3.5 Grey Characteristic Values and Vectors

To prepare for practical applications of grey systems theory, in this section we look at grey characteristic values and vectors and their properties.

Definition 3.5.1. *Assume that $A(\otimes) \in G^{n\times n}$, and $\lambda(\otimes)$ is a grey number. If there is a nonzero grey vector*

$$X(\otimes) = (\otimes_1, \otimes_2, ..., \otimes_n)^T,$$

such that

$$A(\otimes)\cdot X(\otimes) = \lambda(\otimes)\cdot X(\otimes),$$

then $\lambda(\otimes)$ is called a grey characteristic value (or grey eigenvalue) of $A(\otimes)$ and $X(\otimes)$ a characteristic vector (or eigenvector) of $A(\otimes)$ associated with $\lambda(\otimes)$.

Definition 3.5.2. *Assume that* $A(\otimes)\in\ G^{n\times n}$ *and* $\lambda(\otimes)$ *is an unknown grey number. The grey matrix*

$$\lambda(\otimes)\cdot E - A(\otimes)$$

is called the grey characteristic matrix of $A(\otimes)$*. Its determinant*

$$|\lambda(\otimes)\cdot E - A(\otimes)|$$

is called the grey characteristic polynomial of $A(\otimes)$*, and the equation*

$$|\lambda(\otimes)\cdot E - A(\otimes)| = 0$$

the grey characteristic equation of $A(\otimes)$*.*

Proposition 3.5.1. *Assume* $A(\otimes)\in\ G^{n\times n}$*. Then* $\lambda(\otimes)$ *is a grey characteristic value of* $A(\otimes)$ *and* $X(\otimes)$ *a grey characteristic vector of* $A(\otimes)$ *associated with* $\lambda(\otimes)$*, if, and only if,* $\lambda(\otimes)$ *is a root of the grey characteristic equation*

$$|\lambda(\otimes)\cdot E - A(\otimes)| = 0$$

and $X(\otimes)$ *is a non-zero solution of the homogeneous system of grey linear equations*

$$[\lambda(\otimes)\cdot E - A(\otimes)]\,X(\otimes) = 0.$$

Proposition 3.5.2. *Assume* $A(\otimes)\in G^{n\times n}$*. Then transposition grey matrices* $A(\otimes)^T$ *and* $A(\otimes)$ *have the same characteristic values.*

Definition 3.5.3. *Assume* $A(\otimes)\in\ G^{n\times n}$ *such that*

$$A(\otimes)\ =\ [\otimes_{ij}]_{n\times n}\,,$$

where $\otimes_{ij}\in\left[\underline{a}_{ij},\overline{a}_{ij}\right]$, *and* $\underline{a}\ _{ij}<\overline{a}_{ij}$*. If*

$$|\underline{a}_{ii}| < \sum_{j=1,j\neq i}^{n} \left|\underline{a}_{ij}\right|,$$

$i=1,2,...,n$*, then* $A(\otimes)$ *is called a diagonally favorable grey matrix of lower limits. If*

$$|\overline{a}_{ii}| < \sum_{j=1,j\neq i}^{n} |\overline{a}_{ij}|\,,$$

$i=1,2,...,n$*, then* $A(\otimes)$ *is called a diagonally favorable grey matrix of upper limits.*

Proposition 3.5.3. *Assume* $A(\otimes)\in\ G^{n\times n}$*. Then, the following hold true,* *1. If $A(\otimes)$ is a diagonally favorable grey matrix of lower limits, then the matrix $\underline{A}$ of lower limits of $A(\otimes)$ is non-singular.*

2. If $A(\otimes)$ is a diagonally favorable grey matrix of upper limits, then the matrix $\overline{A}$ of upper limits of $A(\otimes)$ is non-singular.

3. If $A(\otimes)$ is both diagonally favorable in terms of lower limits and upper limits, and $\underline{a}_{ii}\overline{a}_{ii} > 0, i = 1, 2, ..., n$, then $A(\otimes)$ is non-singular.

4

Generation of Grey Sequences

4.1 Introduction

One of the main tasks facing the theory of grey systems is to seek a mathematical relationship among factors, based on behavioral data of social, economic, ecological, etc., systems. In grey systems theory, each stochastic process is seen as a grey quantity taking values on a certain range or changing on a certain range of time. And, stochastic processes are treated as grey processes.

As a matter of fact, when investigating the behavioral characteristics of a system, the data obtained are often a sequence of definite white numbers. It will be essentially the same whether we see the sequence either as an orbit or reality of a certain stochastic process, or as the whitenization values of a grey process. However, if the characteristic data of systems' behaviors are used to study the laws of development of the systems, different methodological thoughts will lead to different theoretical outcomes.

The theory of stochastic processes is based on prior laws to uncover the statistical laws, if any, implied by the data. This method is established on large sets of available data. However, sometimes even with a large quantity of data, there might not be any statistical laws to be found, because there are only a few typical distributions employed in theories of probability and stochastic processes. As for non-typical distributions, it is often difficult to make much progress.

In the theory of grey systems, it is through organization of raw data that the researcher must sort out development or governing laws, if any. This

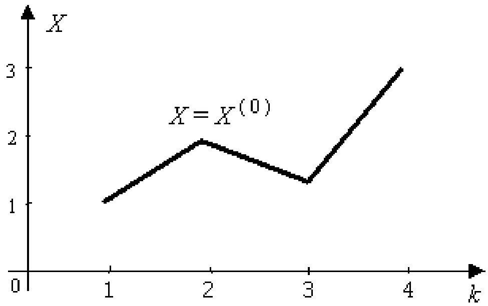

FIGURE 4.1. A graphical representation of the data set $X^{(0)}$

is a path of finding out realistic governing laws from the available data. This path is called a generation of grey sequences. It is believed in the theory of grey systems that even though objective systems phenomena can be complicated and related data chaotic, they always represent a whole, hence, implicitly contain the underlying governing laws. The key for us to uncover and to make use of these laws is how to choose appropriate method(s). The randomness of all grey sequences can be weakened to show its regularities through some processing, also called generation of the given sequences.

For example, the following sequence

$$X^{(0)} = (1, 2, 1.5, 3) = \left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}\right)$$

is given. It does not clearly show any regularity or pattern. Now, we depict the data set with the graph in Figure 4.1. From this graph, it can be seen that the curve of $X^{(0)}$ undulates with relatively large amplitude.

If we apply accumulating generation once to the original data set $X^{(0)}$, and denote the resultant sequence as $X^{(1)}$, then we have

$$X^{(1)} = (1, 3, 4.5, 7.5) = \left(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(1)}\right),$$

where for $i = 1, 2, 3, 4$, $x_i^{(1)}$ is given by

$$x_i^{(1)} = \sum_{j=1}^{i} x_j^{(1)}.$$

Now, this processed sequence $X^{(1)}$ clearly shows a growing tendency (see Figure 4.2 for more details).

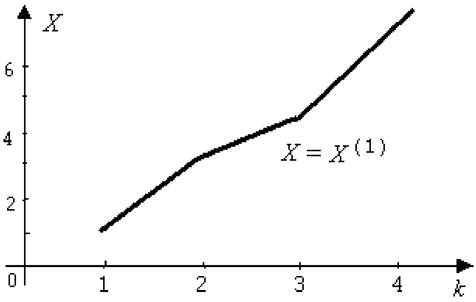

FIGURE 4.2. The resultant sequence $X^{(1)}$ with a clear tendency of growth

4.2 Generation Based on Average

When collecting data, often due to some unconquerable difficulties, there appear some blanks in the data sequence collected. There also exist such data sequences that even though the data are complete, some abnormal values are included due to dramatic behavioral changes of the system under investigation. These abnormal data values bring more difficulties to the researcher. However, if these abnormal values are deleted, some blanks in the data sequence will be created. Therefore, how to effectively fill these blanks naturally becomes the first problem the researcher has to face when dealing with sequential data. Here, the method of generation based on average is commonly used to construct new data values to fill blanks existing in a given sequence of data, and to generate new sequences.

Definition 4.2.1. *Assume that*

$$X = (x(1), x(2), ..., x(k), x(k+1), ..., x(n))$$

is a given sequence. Then, $x(k)$ and $x(k+1)$ are called consecutive neighbors of X, $x(k)$ the predecessor, and $x(k+1)$ the successor. If $x(n)$ is a piece of new data (or information), then for any $k \leq n-1, x(k)$ is called a piece of old data (or information).

Definition 4.2.2. *If a sequence X contains a blank at the location k, then the blank entry is denoted as $\varnothing(k)$. That is, the sequence X looks like*

$$X = (x(1), x(2), ..., x(k-1), \varnothing(k), x(k+1), ..., x(n)).$$

In this case, $x(k-1)$ and $x(k+1)$ are called limit values of $\varnothing(k)$ with $x(k-1)$ being the predecessor limit and $x(k+1)$ the successive limit. When $\varnothing(k) = x(k)$ is generated from $x(k-1)$ and $x(k+1)$, the value $x(k)$ is called an internal point of

$$[x(k-1), x(k+1)] \text{ or } [x(k+1), x(k-1)].$$

Definition 4.2.3. *Assume that $x(k)$ and $x(k-1)$ are two consecutive values in a sequence X. If*

1. *$x(k-1)$ is a piece of old information and $x(k)$ new information, and*
2. *$x^*(k) = \alpha x(k) + (1-\alpha)\, x(k-1), \alpha \in [0,1]$,*

then $x^*(k)$ is called a generated value of the new and old information with generation coefficient (or weight) α. When $\alpha > 0.5$, the generation of $x^*(k)$ is said to have "emphasis more on new and less on old information." When $\alpha < 0.5$, the generation of $x^*(k)$ is said to have "emphasis more on old and less on new information." And when $\alpha = 0.5$, the generation of $x^*(k)$ is said to be "no preference."

The exponential smoothing method widely used when predicting time series is a generation with emphasis more on old and less on new information. It is because the smoothing value

$$s_k^{(1)} = \alpha x_k + (1-\alpha) s_{k-1}^{(1)}$$

is the weighted sum of new information and the smoothing value of old information with the α value limited between 0.01 and 0.3.

Definition 4.2.4. *Assume that*

$$X = (x(1), x(2), ..., x(k-1), \emptyset(k), x(k+1), ..., x(n))$$

is a sequence with a blank $\emptyset(k)$ at location k. Then the value

$$x^*(k) = 0.5x(k-1) + 0.5x(k+1)$$

is called a mean generated value of non-consecutive neighborhood values. The sequence, obtained from using mean generated values of non-consecutive neighbors to fill blanks, is called a generated sequence of non-consecutive neighbors.

When $x(k+1)$ is a piece of new information, the mean generation of non-consecutive values is an equal weight generation based on both new and old information. When it is difficult to measure the reliability of new and old information due to a shortage of needed background information, the method of generation of equal weight, which is the mean generation, is often employed.

Definition 4.2.5. *Assume that a data sequence*

$$X = (x(1),\ x(2), ...,\ x(n))$$

is given. If

$$x^*(k) = 0.5x(k-1) + 0.5x(k),$$

then $x^*(k)$ is called a generated mean value of consecutive neighbors.

In grey systems modeling (GM), we often use the mean generation of consecutive neighbors. It is a method based on the raw sequence of data to construct new sequences in order to reveal the underlying pattern, if any.

Assume that

$$X = (x(1), x(2), ..., x(n))$$

is an n-tuple, and Z the mean sequence generated using consecutive neighbors of X. Then Z is an $(n-1)$-tuple

$$Z = (z(2), z(3), ..., z(n)) .$$

In fact, we have no way to generate $z(1)$ from X, because from the definition of generation of consecutive neighbors, we should have

$$z(1) = 0.5x(0) + 0.5x(1).$$

However, $x(0) = \emptyset(0)$ is a blank in X. If no information expansion is done, we only have the following three choices,

1. Treat $x(0)$ as a grey number without a definite value;
2. Let $x(0)$ be zero or an arbitrary value; or
3. Let $x(0)$ be a value related to $x(1)$,

where case 2 does not have any scientific background, and "letting $x(0) = 0$" in case 2, and cases 1 and 3 do not belong to the category of equal weight mean generations.

4.3 Operators of Sequences

Due to interference of some uncontrollable shock waves, the data set collected sometimes may show too fast or too slow development tendencies, which do not reflect the true development tendency of the system under consideration. If this kind of data is used to build models and to make predictions without first eliminating the effect of the uncontrollable interference, the conclusions obtained are often not usable. The purpose of introducing sequence operators is to eliminate the interference of shock waves in order to show the true face of the data collected, and, based on conclusions of qualitative analysis, to strengthen or to weaken the development tendency of the raw sequences so that resultant prediction accuracy can be improved.

Definition 4.3.1. *Assume that the sequence of data, representing a system's behavior, is given as*

$$X = (x(1), x(2), ..., x(n))$$

1. X is called a monotonic increasing sequence if $\forall\ k = 2, 3, ..., n$, $x(k) - x(k-1) > 0$;

2. X is called a monotonic decreasing sequence if $\forall\ k = 2, 3, ..., n$, $x(k) - x(k-1) < 0$;

3. X is called a vibrational sequence if $\exists\ i, j \in \{2, 3, \cdots, n\}$ such that $x(i) - x(i-1) > 0$ and $x(j) - x(j-1) < 0$. If

$$M = \max\{x(k)|k = 1, 2, ..., n\}$$

and

$$m = \min\{x(k)|k = 1, 2, ..., n\},$$

then $M - m$ is called the amplitude of X.

Definition 4.3.2. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data, D is an operator worked on X, and the sequence, obtained by having D worked on X, is denoted as

$$XD = (x(1)d, x(2)d, ..., x(n)d).$$

Then D is called a sequence operator, and XD the first-order sequence worked on by the operator D.

A sequence operator can be applied as many times as needed. If D_1, D_2, and D_3 are all sequence operators, we call D_1D_2 a *second-order sequence operator* and

$$XD_1D_2 = (x(1)d_1d_2, x(2)d_1d_2, ..., x(n)d_1d_2)$$

a sequence worked on by a second-order operator.

$D_1D_2D_3$ is called a *third-order sequence operator*, and

$$XD_1D_2D_3 = (x(1)d_1d_2d_3, x(2)d_1d_2d_3, ..., x(n)d_1d_2d_3)$$

a sequence worked on by a third order operator. In the same manner, *nth-order sequence operators* and sequences worked on by an nth-order operator can be defined.

Sequence operators and their orders, in applications, can be appropriately chosen and defined based on how much the sequence of the raw data is interfered with by uncontrollable shock waves.

Axiom 4.3.1. *(Axiom of Fixed Points) Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data and D an arbitrary sequence operator. Then D must satisfy

$$x(n)d = x(n).$$

This axiom of fixed points says that under the effect of a sequence operator of our choice, the last datum in the sequence of the raw data must be kept unchanged. This axiom is established on the fact or understanding that $x(n)$ is a starting point or foundation for any future development and is an objective reality, because it is the last data value collected.

According to conclusions of a qualitative analysis, we can also make sure that several other entries near $x(n)$ are kept unchanged under the effect of a desirable sequence operator, so that more emphasis can be addressed to weaken any interference of shock waves on early data values in the sequence. For example, let

$$x(1)d \neq x(1)$$

$$x(2)d \neq x(2)$$

$$\cdots\cdots\cdots\cdots$$

$$x(k-1)d \neq x(k-1)$$

$$x(k)d = x(k)$$

$$x(k+1)d = x(k+1)$$

$$\cdots\cdots\cdots\cdots$$

$$x(n)d = x(n).$$

Axiom 4.3.2. *(Axiom on Sufficient Usage of Information) When a sequence operator is applied, all the information contained in each datum* $x(k)$, $k = 1, 2,..., n$, *of the sequence* X *of the raw data should be sufficiently applied, and any effect of each entry* $x(k)$, $k = 1, 2, ..., n$, *should also be directly reflected in the sequence resulted from the usage of the operator.*

The axiom on sufficient usage of information implies that when we define a sequence operator in an application, we must build or define the operator on the foundation of the data available, and not other way around.

Axiom 4.3.3. *(Axiom of Analytic Representations) For any* $k = 1, 2, ..., n$, $x(k)d$ *can be described with a uniform and elementary analytic representation in* $x(1), x(2), ..., x(n)$.

The axiom of analytic representations requires that the procedure to obtain a new sequence from the original raw data by applying an operator is clear, standardized, unified, and simplified as much as possible so that the actual computation of the new sequence can be relatively easily implemented on a computer.

Axioms 4.3.1, 4.3.2, and 4.3.3 are jointly called three axioms of buffer operators. All sequence operators, satisfying these three axioms, are called *buffer operators*; and the sequences, obtained by applying first, second, third, ..., orders of buffer operators, are referred to as a first-, second-, third-, ..., order buffer sequences, respectively.

Definition 4.3.3. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data, D a buffer operator, and

$$XD = (x(1)d, x(2)d, ..., x(n)d)$$

a D's buffer sequence. When X is respectively a monotonic increasing, decreasing, or vibrational sequence,

1. If the buffer sequence XD increases or decreases more slowly or vibrates with a smaller amplitude than the original sequence X, the buffer operator D is termed a weakening operator;

2. If the buffer sequence XD increases or decreases more rapidly or vibrates with a greater amplitude than the original sequence X, the buffer operator D is termed a strengthening operator.

Theorem 4.3.1. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data, and

$$XD = (x(1)d, x(2)d, ..., x(n)d)$$

one of its buffer sequences. When X is a monotonic increasing sequence, the following hold true,

1. If D is a weakening operator, then $x(k)d \geq x(k), k = 1, 2, ..., n$;

2. If D is a strengthening operator, then $x(k)d \leq x(k), k = 1, 2, ..., n$.

That is, the data in a monotonic increasing sequence expand when a weakening operator is applied, and shrink when a strengthening operator is applied.

Proof: Assume that

$$r(k) = \frac{x(n) - x(k)}{n - k + 1},$$

$k = 1, 2, ..., n$, represents the average rate of increase from $x(k)$ to $x(n)$ in the sequence X of raw data, and

$$r(k)d = \frac{x(n)d - x(k)d}{n - k + 1},$$

$k = 1, 2, ..., n$, the average rate of increase from $x(k)d$ to $x(n)d$ in the buffer sequence XD. From the condition that

$$x(n)d = x(n),$$

it follows that

$$\begin{aligned} r(k) - r(k)d &= \frac{[x(n) - x(k)] - [x(n)d - x(k)d]}{n-k+1} \\ &= \frac{x(n) - x(k) - x(n) + x(k)d}{n-k+1} \\ &= \frac{x(k)d - x(k)}{n-k+1}. \end{aligned}$$

1. If D is a weakening operator, then $r(k) \geq r(k)d$; that is, $r(k) - r(k)d \geq 0$. Therefore, $x(k)d - x(k) \geq 0$; that is, $x(k)d \geq x(k)$.

2. If D is a strengthening operator, then $r(k) \leq r(k)d$; that is, $r(k) - r(k)d \leq 0$. Therefore, $x(k)d - x(k) \leq 0$; that is, $x(k)d \leq x(k)$. □

Example 4.3.1. The condition that $x(k)d \geq x(k)$, $k = 1, 2, ..., n$, in Theorem 4.3.1, 1 is not sufficient to guarantee that D is a weakening operator. To this end, let us look at the following example. For a given increasing positive sequence $X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)}\right)$, meaning that each entry is positive, define a sequence operator D as follows. $X^{(0)}D = \left(x_1^{(1)}, x_2^{(1)}, ..., x_n^{(1)}\right)$ such that

$$x_i^{(1)} = \sqrt{x_i^{(0)} x_n^{(0)}}, \ i = 1, 2, ..., n.$$

It can be checked that $x_n^{(1)} = x_n^{(0)}$ and D is reversible with the inverse given by

$$x_i^{(0)} = \frac{\left(x_i^{(1)}\right)^2}{x_n^{(1)}}, \ i = 1, 2, ..., n.$$

Because $x_i^{(1)} = \sqrt{x_i^{(0)} x_n^{(0)}} \geq x_i^{(0)}$, $i = 1, 2, ..., n$, we have

$$\begin{aligned} x_2^{(1)} - x_1^{(1)} &= \sqrt{x_2^{(0)} x_n^{(0)}} - \sqrt{x_1^{(0)} x_n^{(0)}} \\ &= \sqrt{x_n^{(0)}} \left(\sqrt{x_2^{(0)}} - \sqrt{x_1^{(0)}} \right) \\ &> x_2^{(0)} - x_1^{(0)}, \end{aligned}$$

if $\sqrt{x_n^{(0)}} > \sqrt{x_2^{(0)}} + \sqrt{x_1^{(0)}}$. That is, the buffer operator D may not be a weakening operator.

Example 4.3.2. The condition that $x(k)d \leq x(k), k = 1, 2, ..., n$, in Theorem 4.3.1, 2 is not sufficient to guarantee that D is a strengthening operator. That is, there is a monotonically increasing sequence $X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)}\right)$ and buffer operator D such that $X^{(1)} = X^{(0)}D =$

$\left(x_1^{(1)}, x_2^{(1)}, ..., x_n^{(1)}\right)$ with $x_i^{(1)} \leq x_i^{(0)}$, $i = 1, 2, ..., n$, and D is not a strengthening operator. For instance, if $X^{(0)}$ is a positive sequence and D the inverse operator of that as studied in Example 4.3.1; that is,

$$x_i^{(1)} = \frac{\left[x_i^{(0)}\right]^2}{x_n^{(0)}};$$

then, $x_n^{(1)} = x_n^{(0)}$ and D is reversible with the inverse given by

$$x_i^{(0)} = \sqrt{x_i^{(1)} x_n^{(1)}}$$

and

$$x_i^{(1)} = \frac{\left[x_i^{(0)}\right]^2}{x_n^{(0)}} \leq \frac{\left[x_i^{(0)}\right]^2}{x_i^{(0)}} = x_i^{(0)}, \ i = 1, 2, ..., n.$$

Because it is possible that $x_n^{(0)} > x_2^{(0)} + x_1^{(0)}$, this inequality implies that it is possible that

$$\begin{aligned} x_2^{(1)} - x_1^{(1)} &= \frac{\left[x_2^{(0)}\right]^2 - \left[x_1^{(0)}\right]^2}{x_n^{(0)}} \\ &< \frac{\left(x_2^{(0)} + x_1^{(0)}\right)\left(x_2^{(0)} - x_1^{(0)}\right)}{x_2^{(0)} + x_1^{(0)}} \\ &= x_2^{(0)} - x_1^{(0)}. \end{aligned}$$

That is, D is not a strengthening operator.

Theorem 4.3.2. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data, and

$$XD = (x(1)d, x(2)d, ..., x(n)d)$$

one of its buffer sequences. When X is a monotonic decreasing sequence, the following hold true.

1. If D is a weakening operator, then $x(k)d \leq x(k), k = 1, 2, ..., n$;

2. If D is a strengthening operator, then $x(k)d \geq x(k), k = 1, 2, ..., n$.

That is, the data in a monotonic decreasing sequence shrink when a weakening operator is applied, and expand when a strengthening operator is applied.

The proof of this theorem is similar to that of Theorem 4.3.1 and is omitted here.

As suggested by Examples 4.3.1 and 4.3.2, it can be seen that the conclusions 1 and 2 in Theorem 4.3.2 cannot be rewritten as necessary and sufficient conditions for D to be a weakening and a strengthening operator, respectively.

Theorem 4.3.3. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data and

$$XD = (x(1)d, x(2)d, ..., x(n)d)$$

one of its buffer sequences. When X is a vibrational sequence, the following hold true.

1. If D is a weakening operator, then

$$\max_{1 \le k \le n} \{x(k)d\} \le \max_{1 \le k \le n} \{x(k)\}$$

and

$$\min_{1 \le k \le n} \{x(k)d\} \ge \min_{1 \le k \le n} \{x(k)\}.$$

2. If D is a strengthening operator, then

$$\max_{1 \le k \le n} \{x(k)d\} \ge \max_{1 \le k \le n} \{x(k)\}$$

and

$$\min_{1 \le k \le n} \{x(k)d\} \le \min_{1 \le k \le n} \{x(k)\}.$$

The proof is omitted here.

Proposition 4.3.1. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data and

$$XD = (x(1)d, x(2)d, ..., x(n)d)$$

one of its buffer sequences, where for any $k = 1, 2, ..., n$,

$$x(k)d = \frac{1}{n-k+1}\left[x(k) + x(k+1) + ... + x(n)\right].$$

Then when X is a monotonic increasing, a monotonic decreasing, or a vibrational sequence, D is always a weakening operator.

This proposition is a straightforward consequence of the definition of $x(k)d$. The detailed proof is omitted.

The weakening operator D in Proposition 4.3.1 possesses some very good properties and has been applied widely in modeling and prediction of the systems with interference of uncontrollable shock waves.

Proposition 4.3.2. *For the weakening operator D as defined in Proposition 4.3.1, let*

$$XD^2 = XDD = \left(x(1)d^2, x(2)d^2, ..., x(n)d^2\right)$$

and

$$x(k)d^2 = \frac{1}{n-k+1}\left[x(k)d + x(k+1)d + \cdots + x(n)d\right],$$

for $k = 1, 2, ..., n$. Then D^2 is always a second-order weakening operator for monotonic increasing, monotonic decreasing, and vibrational sequences.

Proposition 4.3.3. *For a sequence of raw data and one of its buffer sequences*

$$X = (x(1), x(2), ..., x(n))$$

and

$$XD = (x(1)d, x(2)d, ..., x(n)d),$$

where for any $k = 1, 2, ..., n-1$,

$$x(k)d = \frac{1}{2k-1}\left[x(1) + x(2) + \cdots + x(k-1) + kx(k)\right]$$

and

$$x(n)d = x(n),$$

when X is either monotonic increasing or monotonic decreasing, D is always a strengthening operator.

Proposition 4.3.4. *For the strengthening operator D as defined in Proposition 4.3.3, let*

$$XD^2 = XDD = \left(x(1)d^2, x(2)d^2, ..., x(n)d^2\right),$$

where for any $k = 1, 2, ..., n-1$,

$$x(k)d^2 = \frac{1}{2k-1}\left[x(1)d + x(2)d + \cdots + x(k-1)d + kx(k)d\right]$$

and

$$x(n)d^2 = x(n)d = x(n).$$

Then D^2 is always a second-order strengthening operator for monotonic increasing and monotonic decreasing sequences.

Example 4.3.3. When a slight modification to the mean generation of consecutive neighbors as given in Definition 4.2.5 is made, we can obtain a kind of buffer operator. More specifically, we can define a buffer operator D as follows,

$$XD = (x(1)d, x(2)d, ..., x(n)d),$$

where for any $k = 1, 2, ..., n-1$,

$$x(k)d = 0.5x(k+1) + 0.5x(k)$$

and

$$x(n)d = x(n).$$

It can be seen easily that D is always a weakening operator for monotonic increasing and monotonic decreasing sequences.

Of course, we can also consider constructing other applicable buffer operators. As a matter of fact from our experience, we know that it is not an easy matter to construct weakening and strengthening operators with good and useful properties. Buffer operators can be used not only in grey systems modeling, but also in a wide range of modeling of various kinds. Before an actual modeling, based on conclusions of some qualitative analysis, some buffer operators are applied to the original sequence of raw data to weaken the effect of any shock vibration in order to achieve desirable results potentially possible from the consequent modeling.

Example 4.3.4. The overall business revenue of a county, located in Henan Province of The People's Republic of China, for the years from 1983 to 1986 was recorded as

$$X = (10155, 12588, 23480, 35388)$$

which showed a tendency of rapid growth. The average rate of growth for these years was 51.6%, and the average rate of growth for the years of 1984 $\sim$ 1986 was 67.7%. All those people involved in the economic planning of the county, including some politicians, scholars, related experts, and residents, commonly believed that the overall revenue of this county could not keep up with this record speed of growth in the coming years. If these data were directly used to build models and make predictions, no body could accept the resultant conclusions. After numerous rigorous analyses and discussions, all parties involved recognized that the reason for the high growth rate to have appeared was mainly due to the low baseline, and the low baseline was a consequence of the fact that in the past, the policies relevant to private enterprises had been either not in existence, or encouraged or applied thoroughly. To weaken the growth rate of the sequence of the raw data, it is necessary to artificially add all favorable environmental factors, created from the introduction of the related policies, for the development of private enterprise to the past years. With this goal in mind, we introduced the second-order weakening operator, as defined in Proposition 4.3.3, and obtained the following second order buffer sequence

$$XD^2 = (27260, 29547, 32411, 35388).$$

Now, the consequent modeling, based on XD^2, produced predictions for the years 1987 $\sim$ 2000 for the county's growth of business revenue. These predictions indicated an average 9.4% annual growth. When we look back today, this predicted rate of growth agreed very well with the recorded values over the time span of our predictions.

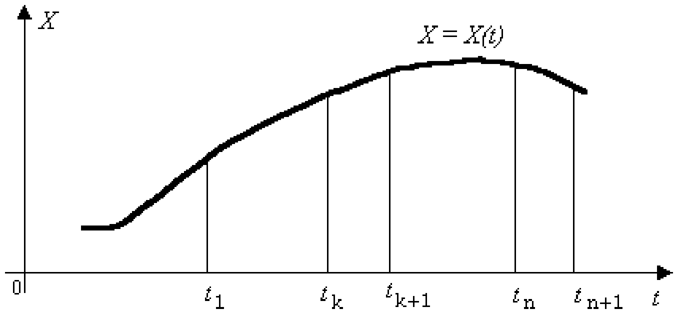

FIGURE 4.3. The curve of a nomotonic increasing and continuous function

4.4 Smooth Sequences

One major characteristic of smooth continuous functions is being differentiable everywhere. Because sequences of data consist of isolated numbers, there is no way for us to talk about differentiability in the traditional sense as described in calculus. Due to this reason, we cannot use derivatives and related methods to study the smoothness of data sequences. However, if we look at the characteristics of smooth continuous functions from a different angle, we may see the following idea: if a sequence of data possesses similar characteristics to those of smooth continuous functions, the sequence may be treated as being smooth.

Without loss of generality, let $X(t)$ be a monotonic increasing and continuous curve, as shown in Figure 4.3.

Partition the interval of consideration into n subintervals with the following points,

$$t_1 < t_2 < \cdots < t_k < t_{k+1} < \cdots < t_{n+1}.$$

Let $\Delta t_k = t_{k+1} - t_k$, $k = 1, 2, ..., n$. Accordingly, the curve $X(t)$ is partitioned into n pieces. Pick an arbitrary point $x(k)$ on the kth piece

$$[x(t_k), x(t_{k+1})],$$

for $k = 1, 2, ..., n$ so that a sequence of some internal points is obtained

$$X = (x(1), x(2), ..., x(n)).$$

Now, by picking the lower limit points of these small pieces, we obtain a sequence of lower limit points:

$$X_0 = (x(t_1), x(t_2), ..., x(t_n)).$$

If $X(t)$ is a smooth continuous function, then when the partition is relatively fine, we have the following.

1. Any two sequences of internal points are sufficiently close to each other;

2. Any sequence of internal points and the relevant lower limit points are sufficiently close.

Definition 4.4.1. *Let* $[a, b]$ *be an interval on the real number line, which is generally seen as a window of time, called a time interval. Assume that this interval is divided into* n *sub-time-intervals* Δt_k, $k = 1, 2, ..., n$. *If the division satisfies*

1. $\Delta t_k = [t_k, t_{k+1}]$;

2. $\bigcup_{k=1}^{n} \Delta t_k = [a, b]$;

3. $\Delta t_i \cap \Delta t_j = \varnothing$, *if* $i \neq j$;

then Δt_k $(k = 1, 2, ..., n)$ *is called a partition of the interval* $[a, b]$.

Based on the previous discussion, a definition in terms of sequences for smooth continuous functions can be given as follows.

Definition 4.4.2. *Assume that* $X(t)$ *is a continuous function defined on interval* $[a, b]$. *Inserting points into the interval* $[a, b]$

$$a = t_1 < t_2 < \cdots < t_k < t_{k+1} < \cdots < t_{n+1} = b$$

gives a partition of the interval $[a, b]$ *as follows,* $\Delta t_k = [t_k, t_{k+1}]$, $k = 1, 2, ..., n$. *At the same time, let us use the same symbol* Δt_k *to indicate the length of the interval* $[t_k, t_{k+1}]$,

$$\Delta t_k = t_{k+1} - t_k,$$

$k = 1, 2, ..., n$. *Picking a point* $x(k)$ *in the subinterval* $[t_k, t_{k+1}]$ *gives a sequence*

$$X = (x(1), x(2), ..., x(n))$$

and write

$$X_0 = (x(t_1), x(t_2), ..., x(t_n))$$

as the sequence of lower limit points. Let $\Delta t = \max_{1 \leq k \leq n} \{\Delta t_k\}$. *Assume that* d *is the distance function in the* n*-dimensional Euclidean space* $\mathbf{R}^n$, *where*

$$\mathbf{R}^n = \{(x_1, x_2, ..., x_n) \mid x_i \text{ is a real number}, 1 \leq i \leq n\},$$

and X^* *is a representative sequence of a chosen differentiable function. If when* $\Delta t \to 0$, *no matter how the time interval* $[a, b]$ *is partitioned, and how an internal point is picked from each subinterval, the following always hold true,*

1. for any two sequences X_i *and* X_j *of internal points,*

$$d(X^*,\ X_i) = d(X^*,\ X_j),$$

2. $d(X^*,\ X) = d(X^*,\ X_0)$,

then $X(t)$ *is termed as a smooth continuous function.*

In the following, we introduce the concept of smooth sequences of data. When defining smooth continuous functions, we mainly focused on whether the sequences of internal points and lower limit points agree, and whether the sequences of internal and lower limit points agree with a sequence of representative points of a differentiable function. As for sequences, there only exist endpoints between which there exist blank spaces. To resolve this problem, we make use of the method of mean generation to create internal points from endpoints.

Definition 4.4.3. *Assume that*

$$X = (x(1), x(2), ..., x(n), x(n+1))$$

is a sequence, and Z *a sequence obtained by mean generation*

$$Z = (z(1), z(2), ..., z(n)),$$

where $z(k) = 0.5x(k) + 0.5x(k+1), k = 1, 2, \cdots, n$. *Assume again that* X^* *is a representative sequence of a differentiable function, and d the distance function in the n-dimensional Euclidean space* $\mathbf{R}^n$. *If the sequence, obtained from deleting* $x(n+1)$ *from* X, *is still denoted as* X, *and* X *satisfies*

1. When k *is sufficiently large,* $x(k) < \sum_{i=1}^{k-1} x(i)$;

2. $\max_{1 \le k \le n} |x^*(k) - x(k)| \ge \max_{1 \le k \le n} |x^*(k) - z(k)|$;

then X *is called a smooth sequence.*

Definition 4.4.4. *Assume that* X *is a smooth sequence,* Z *the sequence of mean generation based on* X, X^* *a representative sequence of a fixed differentiable function, and* d *the distance function in the* n*-dimensional Euclidean space* $\mathbf{R}^n$. *If there exists* $\varepsilon \in [0, 1]$ *such that*

$$|d(X^*, X) - d(X^*, Z)| \le \varepsilon,$$

then the degree of smoothness of the sequence X *is said to be greater than* $\frac{1}{\varepsilon}$, *and*

$$|d(X^*, X) - d(X^*, Z)|^{-1}$$

is referred to as the degree of smoothnesss of the sequence X, *denoted* $S(d)$. *When*

$$|d(X^*, X) - d(X^*, Z)| = 0,$$

that is, when $S(d) = \infty$, *we call* X *an infinitely smooth sequence.*

Proposition 4.4.1. *If the data values in a sequence* X *are distributed on a straight line, then* X *is an infinitely smooth sequence.*

4.5 Stepwise and Smooth Ratios

For a given sequence $X = (x(1), x(2), ..., x(n))$, when the starting or the ending entries $x(1)$ or $x(n)$ are blank, that is, $x(1) = \varnothing(1)$ or $x(n) = \varnothing(n)$, there is no way to use the method of mean generation to fill these blank(s). In this case, other methods will have to be considered to fill the starting and/or ending blank(s). In this situation, the methods of stepwise ratio generation and smooth ratio generation are often used.

Definition 4.5.1. *Let*

$$X = (x(1), x(2), ..., x(n))$$

be a sequence. Then

$$\sigma(k) = \frac{x(k)}{x(k-1)},$$

$k = 2, 3, ..., n$, *are called stepwise ratios of the sequence* X, *and,*

$$\rho(k) = \frac{x(k)}{\sum_{i=1}^{k-1} x(i)},$$

$k = 2, 3, ..., n$, *smooth ratios of the sequence* X.

Definition 4.5.2. *Assume that* X *is a sequence with blanks at the two ends. That is,*

$$X = (\varnothing(1), x(2), ..., x(n-1), \varnothing(n)).$$

If the stepwise ratio (or smooth ratio) of the right-side neighbor of $\varnothing(1)$ *is used to generate* $x(1)$, *and the stepwise ratio (or smooth ratio) of the left-side neighbor of* $\varnothing(n)$ *is used to generate* $x(n)$, *then* $x(1)$ *and* $x(n)$ *are said to be stepwise ratio generated (or smooth ratio generated). The sequence, with blanks filled by stepwise ratio generation (or smooth ratio generation), is called a sequence generated with stepwise ratios (or smooth ratios).*

Proposition 4.5.1. *Assume that* X *is a sequence with blank ends.*

1. If stepwise ratio generation is applied, then

$$x(1) = \frac{x(2)}{\sigma(3)}, x(n) = x(n-1)\sigma(n-1).$$

2. If smooth ratio generation is used, then

$$x(1) = \frac{x^2(2)}{x(3) - x(2)}, x(n) = x(n-1)\left(1 + \rho(n-1)\right).$$

Proposition 4.5.2. *The following equation establishes the relationship between stepwise ratios and smooth ratios.*

$$\sigma(k+1) = \frac{\rho(k+1)}{\rho(k)}\left(1 + \rho(k)\right),$$

for $k = 2, 3, ..., n$.

Proposition 4.5.3. *If*

$$X = (x(1),\ x(2), ...,\ x(n))$$

is an increasing sequence, satisfying that

1. For $k = 2, 3, ..., n, \sigma(k) < 2$*;*

2. For k = 2, 3, ..., n,

$$\frac{\rho(k+1)}{\rho(k)} < 1,$$

that is, the smooth ratio is decreasing, then for any fixed real number $\varepsilon \in [0, 1]$ *and* $k = 2, 3, ..., n$*, when* $\rho(k) \in [0, \varepsilon]$*, it must be that* $\sigma(k+1) \in [0, 1+\varepsilon]$.

Example 4.5.1. Given a sequence of data

$$X = (2.874, 3.278, 3.337, 3.390, 3.679),$$

we have

$$\sigma(2) = \frac{x(2)}{x(1)} = \frac{3.278}{2.874} = 1.14$$

and

$$\sigma(3) = 1.1017, \sigma(4) = 1.1015, \sigma(5) = 1.1085$$

satisfying that for $k = 2, 3, 4, 5, \sigma(k) < 2$.

$$\rho(2) = \frac{x(2)}{x(1)} = 1.14,$$

$$\rho(3) = \frac{x(3)}{x(1) + x(2)} = 0.5425,$$

$$\rho(4) = \frac{x(4)}{x(1) + x(2) + x(3)} = 0.3573,$$

and

$$\rho(5) = \frac{x(5)}{x(1) + x(2) + x(3) + x(4)} = 0.2856,$$

satisfying that for $k = 2, 3, 4, 5,$

$$\frac{\rho(k+1)}{\rho(k)} < 1.$$

If $\rho(2)$ is not seen as a smooth ratio, then when $k = 3, 4, 5,$

$$\rho(k) \in [0, 0.5425] = [0, \varepsilon],$$

and $\sigma(k+1) \in [0, 1.085] \subset [0, 1+\varepsilon]$, $k = 2, 3, 4$.

Definition 4.5.3. *If a sequence* $X = (x(1), x(2), ..., x(n))$ *satisfies that*

1. For $k = 2, 3, ..., n-1$,

$$\frac{\rho(k+1)}{\rho(k)} < 1;$$

2. For $k = 3, 4, ..., n$, $\rho(k) \in [0, \varepsilon]$; *and*

3. $\varepsilon < 0.5$,

then X is said to be a quasi-smooth sequence.

Definition 4.5.4. *Assume that X is a sequence with blank entries. If a new sequence generated based on X satisfies the conditions of being quasi-smooth, then the related generation is said to be quasi-smooth generation.*

4.6 Accumulating and Inverse Accumulating Generation Operators

The so-called accumulating generation is a method to whitenize a grey process. It occupies a very important position in the theory of grey systems. Through accumulation, the development situation and tendency of a grey quantity can be clearly seen so that special characteristics or laws, hidden in the chaotic raw data, can be sufficiently revealed. For example, let us consider the total expense of a family. If it is computed daily, there may not be any obvious pattern. If computed monthly, some patterns of expenses of the family may well be shown, which may, for example, very well relate to the family's monthly income. If we consider the weight of a single piece of wheat, in general, we might not find any useful information. That is why people often use the weight of a thousand pieces of wheat to be an evaluation index. For a manufacturing business of heavy equipment, because the production of each piece of equipment requires a relatively long period of time, it will be useless to do daily analysis. On the other hand, it will be very important to do an annual evaluation on the overall production and relevant revenue.

The inverse accumulating generation is often used when additional information is needed. At the same time, it also plays the role of returning the data, after an accumulating generation process was applied, to the original condition. So, accumulating and inverse accumulating generations are a pair of inverse sequence operators.

Definition 4.6.1. *Assume that* $X^{(0)}$ *is a sequence of raw data and D a sequence operator satisfying that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

and

$$X^{(0)}D = \left(x^{(0)}(1)d, x^{(0)}(2)d, ..., x^{(0)}(n)d\right),$$

where

$$x^{(0)}(k)d = \sum_{i=1}^{k} x^{(0)}(i),$$

for $k = 1, 2, ..., n$. *Then the sequence operator* D *is called a* $(first$*-order*$)$ *accumulating generator of* $X^{(0)}$, *denoted 1-ADO. The* r*th-order operator* D^r *of* $X^{(0)}$ *is obtained by applying the 1-ADO* D r *times, denoted* r*-AGO.*

Conventionally, we write

$$X^{(0)}D = X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\right)$$

and

$$X^{(0)}D^r = X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), ..., x^{(r)}(n)\right),$$

where $x^{(r)}(k) = \sum_{i=1}^{k} x^{(r-1)}(i), k = 1, 2, ..., n.$

Definition 4.6.2. *Assume that* $X^{(0)}$ *is a sequence of raw data and* D *is a sequence operator such that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

and

$$X^{(0)}D = \left(x^{(0)}(1)d, x^{(0)}(2)d, ..., x^{(0)}(n)d\right),$$

where

$$x^{(0)}(1)\,d = x^{(0)}(1)\,,\ \text{and}$$

$$x^{(0)}(k)d = x^{(0)}(k) - x^{(0)}(k-1),$$

for $k = 1, 2, ..., n$. *Then* D *is called a* $(first - order)$ *inverse accumulating generator of* $X^{(0)}$. *The* r*th order operator* D^r *is called an (*r*th-order) inverse accumulating generator of* $X^{(0)}$.

Again, conventionally, we use the following notation:

$$X^{(0)}D = \alpha^{(1)}X^{(0)} = \left(\alpha^{(1)}x^{(0)}(1), \alpha^{(1)}x^{(0)}(2), ..., \alpha^{(1)}x^{(0)}(n)\right),$$

and

$$X^{(0)}D^r = \alpha^{(r)}X^{(0)} = \left(\alpha^{(r)}x^{(0)}(1), \alpha^{(r)}x^{(0)}(2), ..., \alpha^{(r)}x^{(0)}(n)\right),$$

where

$$\alpha^{(r)}x^{(0)}(k) = \alpha^{(r-1)}x^{(0)}(k) - \alpha^{(r-1)}x^{(0)}(k-1),$$

for $k = 2, 3, ..., n$, and $\alpha^{(r)}x^{(0)}(1) = \alpha^{(r-1)}x^{(0)}(1)$.

Based on the previous definitions, it is obvious that

Theorem 4.6.1. *The inverse accumulating generator is an inverse operator of the accumulating generator. That is,*

$$\alpha^{(r)} X^{(r)} = X^{(0)}.$$

Because of this result, we denote an inverse accumulating generator as *IAGO*.

Example 4.6.1. Assume that the annual data records of a grey number are given in the following sequence

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right).$$

And, based on the monthly and yearly records, the data values recorded for the kth year are the sequence

$$X^{(0)}(k) = \left(x^{(0)}(1,k), x^{(0)}(2,k), ..., x^{(0)}(12,k)\right).$$

Now, the data records for the kth year jth month in terms of days are given in the sequence below:

$$X^{(0)}(j,k) = \left(x^{(0)}(1,j,k), x^{(0)}(2,j,k), ..., x^{(0)}(30,j,k)\right).$$

And, if the data are recorded with such details as to hour, day, month, and year, then the data for the kth year, jth month, and ith day are given in the following sequence:

$$X^{(0)}(i,j,k) = \left(x^{(0)}(1,i,j,k), x^{(0)}(2,i,j,k), ..., x^{(0)}(24,i,j,k)\right),$$

where $x^{(0)}(h,i,j,k)$, $h = 1, 2, ..., 24$, is the record for the hth hour of the kth year, jth month, and ith day.

Obviously, we have

$$x^{(0)}(k) = \sum_{j=1}^{12} x^{(0)}(j,k),$$

$$x^{(0)}(j,k) = \sum_{i=1}^{30} x^{(0)}(i,j,k),$$

$$x^{(0)}(i,j,k) = \sum_{h=1}^{24} x^{(0)}(h,i,j,k),$$

and

$$x^{(0)}(k) = \sum_{j=1}^{12}\sum_{i=1}^{30}\sum_{h=1}^{24} x^{(0)}(h,i,j,k),$$

This is called a laminated accumulation. It is not the same as what was defined as accumulating generation (or time accumulation).

Proposition 4.6.1. *Assume that* $X^{(0)}$ *is a non-negative sequence*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right),$$

where $x^{(0)}(k) \geq 0$ *and* $x^{(0)}(k) \in [a, b]$, $k = 1, 2, ..., n$. *If*

$$X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), ..., x^{(r)}(n)\right)$$

is the sequence of an r*th-order accumulating generation of* $X^{(0)}$, *then for any* $\varepsilon > 0$, *when* r *is sufficiently large, there exists an* N *such that for any* k *satisfying* $N < k \leq n$ *the following holds true.*

$$\frac{x^{(r)}(k)}{\sum_{i=1}^{k-1} x^{(r)}(i)} < \varepsilon.$$

That is to say, for a bounded non-negative sequence, after many applications of accumulating generations, the resultant sequence can be sufficiently smooth and the smooth ratio $\rho(k) \to 0$, as $k \to \infty$.

Proposition 4.6.2. *Let* $X^{(0)}$ *be the same as in Proposition 4.6.1. If*

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\right)$$

is the sequence of the first-order accumulating generation of $X^{(0)}$, *and*

$$Z^{(1)} = \left(z^{(1)}(2), z^{(1)}(3), ..., z^{(1)}(n)\right)$$

is the sequence of mean generation of consecutive neighbors of $X^{(1)}$, *then for any* $\varepsilon_1 \leq \varepsilon_2 \in [0, 1]$, *there exists a positive integer* $N = N(\varepsilon_1, \varepsilon_2)$ *such that for any* k *with* $N < k \leq n$, *the following holds true.*

$$\rho(k) = \frac{x^{(0)}(k)}{\sum_{i=1}^{k-1} x^{(0)}(i)} < \varepsilon_1, \quad \frac{x^{(0)}(k)}{z^{(1)}(k)} < \varepsilon_2.$$

Proof. It suffices for us to prove that

$$z^{(1)}(k) \geq \sum_{i=1}^{k-1} x^{(0)}(i).$$

From the definition of mean generation of consecutive neighbors, it follows that

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$$

for $k = 2, 3, ..., n$, and

$$\begin{aligned} \sum_{i=1}^{k-1} x^{(0)}(i) &= x^{(1)}(k-1) \\ &= 0.5x^{(1)}(k-1) + 0.5x^{(1)}(k-1) \end{aligned}$$

However,

$$\begin{aligned} x^{(1)}(k) &= \sum_{i=1}^{k} x^{(0)}(i) \\ &= \sum_{i=1}^{k-1} x^{(0)}(i) + x^{(0)}(k) \\ &= x^{(1)}(k-1) + x^{(0)}(k) \end{aligned}$$

and $x^{(0)}(k) \geq 0$, so

$$z^{(1)}(k) \geq \sum_{i=1}^{k-1} x^{(0)}(i). \ \square$$

Let us conclude this section with the following statement. Both accumulating generation and mean generation can increase the degree of smoothness of a sequence. Sometimes, after an application of accumulating generation, a mean generation can also be applied.

4.7 Randomness of Sequences of Accumulating Generations

In general, the randomness of a non-negative quasi-smooth sequence decreases if the accumulating generation procedure is applied. The smoother the original sequence is, the more clearly an exponential tendency in the generated sequence would show up. For example, for the retailing of bicycles in a certain city, we have the sequence of raw data as follows,

$$\begin{aligned} X^{(0)} &= \left\{x^{(0)}(k)\right\}_1^6 \\ &= (50810, 46110, 51177, 93775, 110574, 110524) \end{aligned}$$

whose sequence of the first-order accumulating generation is given by

$$\begin{aligned} X^{(1)} &= \left\{x^{(1)}(k)\right\}_1^6 \\ &= (50810, 96920, 148097, 241872, 352446, 462970). \end{aligned}$$

These sequences are plotted in Figures 4.4 and 4.5.

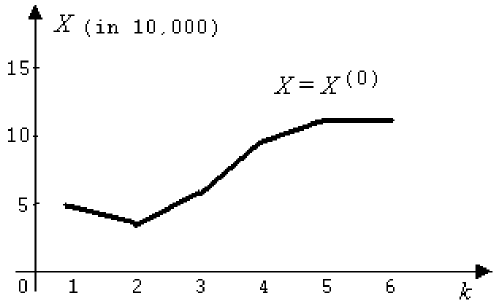

FIGURE 4.4. The curve of the sequence $X^{(0)}$

For the curve $X = X^{(0)}$ in Figure 4.4, it is very difficult to find an elementary curve to approximate it, whereas for the curve $X = X^{(1)}$ in Figure 4.5, it is quite like the curve of an exponential function so that it can be simulated with an exponential function.

Definition 4.7.1. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence of raw data, and

$$\alpha^{(1)}X^{(0)} = \left(\alpha^{(1)}x^{(0)}(1), \alpha^{(1)}x^{(0)}(2), ..., \alpha^{(1)}x^{(0)}(n)\right)$$

is the sequence generated by applying the inverse accumulating generator once on $X^{(0)}$.

1. The sequence $X^{(0)}$ *is said to be increasing at the* k*th step if*

$$\alpha^{(1)}x^{(0)}(k) = x^{(0)}(k) - x^{(0)}(k-1) > 0,$$

otherwise, $X^{(0)}$ *is said to be decreasing at step* k*;*

2. If for $k = 2, 3, ..., n$, $\alpha^{(1)}x^{(0)}(k) > 0$ *always holds true, then the sequence* $X^{(0)}$ *is said to be non-vibratingly increasing (or stepwise increasing);*

3. If for $k = 2, 3, ..., n$, $\alpha^{(1)}x^{(0)}(k) < 0$ *always holds true, then* $X^{(0)}$ *is said to be a non-vibratingly decreasing sequence; and*

4. If $\exists k_1, k_2 \geq 2$ *such that* $\alpha^{(1)}x^{(0)}(k_1) > 0$ *and* $\alpha^{(1)}x^{(0)}(k_2) < 0$*, then* $X^{(0)}$ *is called a stochastic sequence.*

Definition 4.7.2. *1. If* $X^{(0)}$ *is a non-vibrating sequence, and* $\alpha^{(1)}X^{(0)}$ *is a stochastic sequence, then* $X^{(0)}$ *is called a weak stochastic sequence of the first order.*

2. If for $i = 0, 1, 2, ..., r-1$, $\alpha^{(i)}X^{(0)}$ *is a non-vibrating sequence, and* $\alpha^{(r)}X^{(0)}$ *is a stochastic sequence, then* $X^{(0)}$ *is said to be a weak stochastic sequence of the* r*th order;*

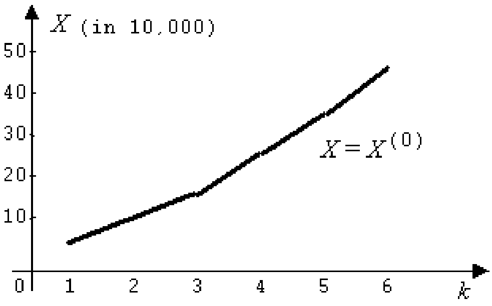

FIGURE 4.5. The curve of the sequence $X^{(1)}$

3. If for any $r \to \infty$, $\alpha^{(r)} X^{(0)}$ is a non-vibrating sequence, then $X^{(0)}$ is said to be a non-stochastic sequence.

Theorem 4.7.1. *Assume that $X^{(0)}$ is a positive sequence, that is, if*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right),$$

then $x^{(0)}(k) \geq 0, k = 1, 2, ..., n$, and that $X^{(r)}$ is a sequence generated by applying accumulating generation r times; then $X^{(r)}$ must be a weak stochastic sequence of the rth order.

4.8 Grey Exponentiality of Accumulating Generations

In this section, we look at some of the most applied results in the study of grey information.

Definition 4.8.1. *Assume that a continuous function in the form of*

$$X(t) = ce^{at} + b, c, a \neq 0,$$

is given, where a, b,and c are fixed constants.

1. When $b = 0, X(t)$ is called a homogeneous exponential function;

2. When $b \neq 0, X(t)$ is called a non-homogeneous exponential function.

Definition 4.8.2. *Assume that a sequence*

$$X = (x(1), x(2), ..., x(n))$$

is given.

1. If for $k = 1, 2, ..., n, x(k) = ce^{ak}$, $c \neq 0 \neq a$, then X is called a homogeneous exponential sequence; and

2. If for $k = 1, 2, ..., n, x(k) = ce^{ak} + b$ with $c, a, b \neq 0$, then X is called a non-homogeneous exponential sequence.

Theorem 4.8.1. *A sequence X is homogeneously exponential, if, and only if, for $k = 1, 2, \cdots, n$, the equation $\sigma(k) =$ a positive constant always holds true.*

Proof. $\Rightarrow$). Assume that for any $k, x(k) = ce^{ak}, c \neq 0 \neq a$. Then,

$$\sigma(k) = \frac{x(k)}{x(k-1)} = \frac{ce^{ak}}{ce^{a(k-1)}} = e^a = \text{a positive constant.}$$

$\Leftarrow$). Assume again that for any k, $\sigma(k) =$ a positive constant $= e^a$. Then,

$$x(k) = e^a x(k-1) = e^{2a} x(k-2) = \cdots = x(1)e^{a(k-1)}. \square$$

Definition 4.8.3. *Assume that a sequence*

$$X = (x(1),\ x(2),\ ...,\ x(n))$$

is given.

1. If for any k, $\sigma(k) \in (0, 1]$, then the sequence X is said to satisfy the law of negative grey exponent;

2. If for any k, $\sigma(k) \in (1, b]$, for some $b > 1$, then the sequence X is said to satisfy the law of positive grey exponent;

3. If for any k, $\sigma(k) \in [a, b]$ with $b - a = \delta$, then X is said to satisfy the law of grey exponent with the absolute degree of greyness δ;

4. When $\delta <$ 0.5, the sequence X is said to satisfy the law of quasi-exponent.

Theorem 4.8.2. *Assume that $X^{(0)}$ is a non-negative quasi-smooth sequence. Then the sequence $X^{(1)}$, generated by applying accumulating generation once on $X^{(0)}$, satisfies the law of quasi-exponent.*

Proof.

$$\begin{aligned} \sigma^{(1)}(k) &= \frac{x^{(1)}(k)}{x^{(1)}(k-1)} \\ &= \frac{x^{(0)}(k) + x^{(1)}(k-1)}{x^{(1)}(k-1)} \\ &= 1 + \rho(k). \end{aligned}$$

From the definition of quasi-smooth sequences, it follows that for each k, $\rho(k) < 0.5$. Hence,

$$\sigma^{(1)}(k) \in [1, 1.5),\ \ \delta < 0.5.$$

That is, $X^{(1)}$ satisfies the law of quasi-exponent. $\square$

Theorem 4.8.2 is the theoretical foundation for grey systems modeling. As a matter of fact, because all economic systems, ecological systems, agricultural systems, etc., can be seen as general energy systems, in which the accumulation and release of energies satisfy the law of exponent, the exponential modeling of grey systems theory has a very wide range of applications.

Theorem 4.8.3. *Assume that $X^{(0)}$ is a non-negative sequence. If $X^{(r)}$ satisfies a law of exponent and the stepwise ratio of $X^{(r)}$ is given by $\sigma^{(r)}(k) = \sigma$, then*

1. We have

$$\sigma^{(r+1)}(k) = \frac{1-\sigma^k}{1-\sigma^{k-1}};$$

2. When $\sigma \in (0,1)$,

$$\lim_{k\to\infty} \sigma^{(r+1)}(k) = 1,$$

and for each k,

$$\sigma^{(r+1)}(k) \in [1, 1+\sigma];$$

3. When $\sigma > 1$,

$$\lim_{k\to\infty} \sigma^{(r+1)}(k) = \sigma,$$

and for each k,

$$\sigma^{(r+1)}(k) \in (\sigma, 1+\sigma].$$

Proof. 1. Because $X^{(r)}$ satisfies a law of exponent and for each k,

$$\sigma^{(r)}(k) = \frac{x^{(r)}(k)}{x^{(r)}(k-1)} = \sigma,$$

for each k, we have that

$$\begin{aligned} x^{(r)}(k) &= \sigma x^{(r)}(k-1) \\ &= \sigma^2 x^{(r)}(k-2) \\ &= \cdots \\ &= \sigma^{k-1} x^{(r)}(1). \end{aligned}$$

So,

$$X^{(r)} = \left(x^{(r)}(1), \sigma x^{(r)}(1), \sigma^2 x^{(r)}(1), ..., \sigma^{n-1} x^{(r)}(1)\right)$$

and

$$\begin{aligned} X^{(r+1)} = &\left(x^{(r)}(1), (1+\sigma)x^{(r)}(1), (1+\sigma+\sigma^2)x^{(r)}(1), ..., \right. \\ &\left. (1+\sigma+\sigma^2+\cdots+\sigma^{n-1})x^{(r)}(1)\right). \end{aligned}$$

Therefore,

$$\begin{aligned} \sigma^{(r+1)}(k) &= \frac{x^{(r+1)}(k)}{x^{(r+1)}(k-1)} \\ &= \frac{(1+\sigma+\sigma^2+\cdots+\sigma^{k-1})x^{(r)}(1)}{(1+\sigma+\sigma^2+\cdots+\sigma^{k-2})x^{(r)}(1)}. \end{aligned}$$

That is,

$$\sigma^{(r+1)}(k) = \frac{1-\sigma^k}{1-\sigma} \div \frac{1-\sigma^{k-1}}{1-\sigma} = \frac{1-\sigma^k}{1-\sigma^{k-1}}.$$

2. When $0 < \sigma < 1$, $\sigma^{(r+1)}(k)$ decreases as k increases. When $k = 2$,

$$\sigma^{(r+1)}(2) = \frac{x^{(r+1)}(2)}{x^{(r+1)}(1)} = 1 + \sigma.$$

When $k \to \infty$,

$$\sigma^{(r+1)}(k) = \frac{1-\sigma^k}{1-\sigma^{k-1}} \to 1.$$

So, for each k,

$$\sigma^{(r+1)}(k) \in (1, 1+\sigma].$$

3. When $\sigma > 1$, $\sigma^{(r+1)}(k)$ decreases as k increases. When $k = 2$,

$$\sigma^{(r+1)}(2) = 1 + \sigma;$$

and when $k \to \infty$,

$$\sigma^{(r+1)}(k) = \frac{1-\sigma^k}{1-\sigma^{k-1}} \to \sigma.$$

So, for each k,

$$\sigma^{(r+1)}(k) \in (\sigma, 1+\sigma]. \ \square$$

This Theorem 4.8.3 implies that if a sequence of the rth accumulating generation of $X^{(0)}$ satisfies an obvious law of exponent, an additional application of *AGO* will destroy the pattern which has been obviously seen. It tells us that application of accumulating generation needs to be stopped when necessary. In practical applications, if a sequence of the rth accumulation generation of $X^{(0)}$ satisfies the law of quasi-exponent, we will generally no longer apply any further generation. From Theorem 4.8.2, it follows that only one application of accumulating generation is needed for a non-negative quasi-smooth sequence before establishing an exponential model.

5

Grey Incidence Analysis

5.1 Introduction

General abstract systems, similar to social systems, economic systems, agricultural systems, ecological systems, education systems, etc., involve many factors. Some mutual reactions among the factors determine the development situation and tendency of the systems. We often want to know:

- Which factors among the many are more important than others?
- Which factors have more effects on the future development of the systems than others?
- Which factors actually cause desirable changes in the systems so that these factors need to be strengthened?
- Which factors hinder desirable development of the systems so that they need to be controlled?

All these problems are commonly studied in the analysis of systems. For example, in a system of crop production, it is desired to increase the total production of food. However, there exist many factors, such as the area planted, irrigation facilities, fertilizers, soil quality, seeds, labor availability, weather conditions, farming technologies, and related government policies, affecting the desirable outcome. In order to achieve the goal of less input and more output with as great economic, social, and ecological benefits as possible, we must make use of the theory of systems analysis.

Many methods in statistics, such as regression analysis, variance analysis, and principal component analysis, are all commonly used in the analysis of systems. However, these methods have the following pitfalls.

1. A large amount of data is required. Otherwise it would be difficult to draw statistical conclusions with reasonable confidence and reliability.
2. It is required that all samples or populations satisfy certain typical probability distribution(s), that the relation between the main characteristic variable of the system and factor variables is roughly linear. These requirements are often difficult to satisfy in real-life practice.
3. Heavy-duty computations are often needed.
4. It often happens that quantitative conclusions may not agree with qualitative analysis results, causing misunderstandings about the systems under consideration.

In modern China, for example, due to historical reasons, most governmental and/or private sources have only limited amounts of statistical records. And, these limited amounts of records, at the same time, contain a great degree of uncertainty. Also, due to some human factors, many statistical records show fluctuations with large rises and falls without many typical distribution patterns. Therefore, applying statistical methods can hardly achieve many useful conclusions.

Now, the so-called grey incidence analysis remedies this defect found in existing statistics when applied in the content of systems analysis. It can be applied to cases of various sample sizes and distributions with a relatively small amount of computation. And, in general, each application of grey incidence analysis does not result in situations of disagreement between quantitative analysis and qualitative analysis.

The fundamental idea of grey incidence analysis is that the closeness of a relationship is judged based on the similarity level of the geometric patterns of sequence curves. The more similar the curves are, the higher the degree of incidence between sequences, and vice versa.

When analyzing an abstract system or a phenomenon, what's most important is to choose the right sequence of characteristic data to describe the system's behavior. This sequence of data is called a mapping quantity of the special system's behavior. For example, the number of average years of education received by citizens can be used to reflect a nation's level of literacy. Crime rates can be employed to represent the safety and social order of a community. The number of patient registrations of all hospitals in a community can be used as an indicator for the level of health of residents of the community. After the data for the system's behavioral characteristics and data for related factors are collected, we draw the graphs for all the data sets, from which some elementary analysis can be conducted. For

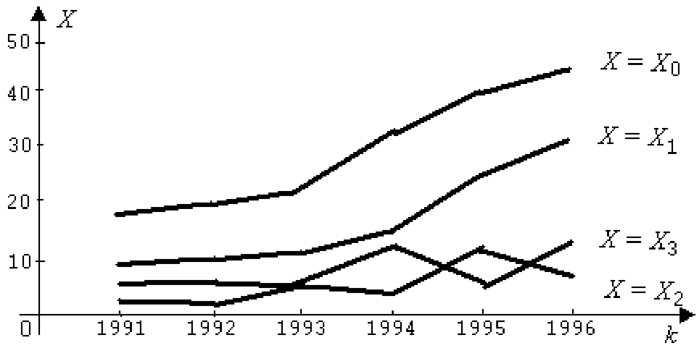

FIGURE 5.1. Graphs of the Xs sequences

example, we have collected the data of a region for its total agricultural production, denoted as X_0, for its total farming production, denoted as X_1, for its total livestock husbandry production, denoted as X_2, and for its total production of rural business enterprises, denoted as X_3, for the years from 1991 to 1996 as follows.

$$\begin{aligned} X_0 &= (18, 20, 22, 35, 41, 46), \\ X_1 &= (8, 11, 12, 17, 24, 29), \\ X_2 &= (3, 2, 7, 4, 11, 6,), \text{ and} \\ X_3 &= (5, 7, 7, 11, 5, 10). \end{aligned}$$

The graph of each sequence X_i, $i = 0, 1, 2, 3$, is given in Figure 5.1. Intuitively speaking, the curve representing farming production is very similar to that representing total agricultural production, whereas the curves for livestock and business enterprises have relatively obvious geometric differences. Therefore, it can be concluded that agriculture in this region is mainly farming, and that livestock husbandry and rural business enterprises are still not well developed.

5.2 Grey Incidence Factors and Set of Grey Incidence Operators

In systems analysis, after the mapping quantities, which describe systems' behaviors well, have been chosen, we need to clarify all the factors that effectively affect the systems' behaviors. If quantitative analysis is needed, it is necessary to process the mapping quantities and effective factors through sequence operators so that these quantities and factors might become non-dimensional with similar behaviors for negatively correlated factors and positively correlated factors.

Definition 5.2.1. *Assume that* X_i *is a systems' factor with the* k*th observation value being* $x_i(k)$, $k = 1, 2, \cdots, n$. *Then*

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is called a behavioral sequence of the factor X_i.

1. *If* k *stands for time, then* $x_i(k)$ *represents an observation of the factor* X_i *at the time moment* k, *and*

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is called a behavioral time sequence of the factor X_i.

2. *If* k *is an ordinality of some criteria and* $x_i(k)$ *is the observation of the factor* X_i *at the criterion* k, *then*

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is called a behavioral criterion sequence of the factor X_i; *and*

3. *If* k *is the ordinal number of the object observed and* $x_i(k)$ *stands for the observation of the factor* X_i *of the* k*th object, then*

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is called a behavioral horizontal sequence of the factor X_i.

For example, when X_i represents some economic factor, k time, and $x_i(k)$ the observation value of the factor X_i at the time moment k, the sequence

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is an economic behavioral time sequence. If k is the ordinality of a criterion, assuming that the set of all criteria applied is ordered, then the sequence

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is an economic behavioral criterion sequence. If k stands for the ordinality of different economic districts, assuming that the set of all economic districts studied has been ordered, then the sequence

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is an economic behavioral horizontal sequence.

No matter which sequence we have, being time, or criterion, or horizontal, we can always conduct the needed incidence analysis.

Definition 5.2.2. *Assume that*

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

is a behavioral sequence of a factor X_i, *and* D_1 *a sequence operator satisfying*

$$X_i D_1 = (x_i(1)d_1, x_i(2)d_1, ..., x_i(n)d_1),$$

where

$$x_i(k)d_1 = \frac{x_i(k)}{x_i(1)},$$

for $k = 1, 2, ..., n$. *Then* D_1 *is called an initialing operator with* X_i *as its preimage and* $X_i D_1$ *as its image, is called the initial image of* X_i.

Definition 5.2.3. *Let* X_i *be the same as in Definition 5.2.2 and* D_2 *a sequence operator such that*

$$X_i D_2 = (x_i(1)d_2, x_i(2)d_2, ..., x_i(n)d_2),$$

and

$$x_i(k)d_2 = \frac{x_i(k)}{\overline{X}_i},\ \overline{X}_i = \frac{1}{n}\sum_{i=1}^{n} x_i(k),$$

for $k = 1,\ 2,\ ..., n$. *Then* D_2 *is called an averaging operator with* $X_i D_2$ *as its image, called the average image of* X_i.

Definition5.2.4. *Let* X_i *be the same as in Definition 5.2.2 and* D_3 *a sequence operator such that*

$$X_i D_3 = (x_i(1)d_3, x_i(2)d_3, ..., x_i(n)d_3),$$

where

$$x_i(k)d_3 = \frac{x_i(k) - \min_k \{x_i(k)\}}{\max_k \{x_i(k)\} - \min_k \{x_i(k)\}},$$

for $k = 1, 2, ..., n$. *Then* D_3 *is called an interval operator with* $X_i D_3$ *as its image, called the interval image of* X_i.

Proposition 5.2.1. *Initialing operator* D_1, *averaging operator* D_2, *and interval operator* D_3 *can all transform a behavioral sequence of a system into a non-dimensional sequence.*

In general, the operators D_1, D_2, and D_3 should not be mixed in applications. And, when analyzing systems factors, a choice among D_1, D_2, and D_3 can be made based on the practical situation involved.

Definition 5.2.5. *Let* X_i *be the same as in Definition 5.2.2 satisfying*

$$x_i(k) \in [0, 1],$$

for $k = 1, 2, ..., n$, *and* D_4 *a sequence operator such that*

$$X_i D_4 = (x_i(1)d_4, x_i(2)d_4, ..., x_i(n)d_4),$$

where

$$x_i(k)d_4 = 1 - x_i(k),$$

for $k = 1, 2,..., n$. *Then* D_4 *is called a reversing operator with* X_iD_4 *as the image of* X_i, *called the reverse image of* X_i.

Proposition 5.2.2. *The interval image of any behavioral sequence has a reverse image.*

Proof. In fact, all data values in an interval image belong to $[0, 1]$, so a reversing operator can be defined. □

Definition 5.2.6. *Let* X_i *be the same as in Definition 5.2.2 and* D_5 *a sequence operator such that*

$$X_iD_5 = (x_i(1)d_5, x_i(2)d_5, ..., x_i(n)d_5)$$

where

$$x_i(k)d_5 = 1/x_i(k),$$

for $k = 1, 2, ..., n$. *Then* D_5 *is called a reciprocating operator with* X_iD_5 *as the reciprocal image of* X_i.

Proposition 5.2.3. *If there exists a negative correlation between a system factor* X_i *and a system behavior* X_0, *then the reverse image* X_iD_4 *and the reciprocal image* X_iD_5 *of the factor* X_i *have a positive correlation with* X_0.

Definition 5.2.7. *The following*

$$D = \{D_i | i = 1, 2, 3, 4, 5\}$$

is called the set of grey incidence operators.

Definition 5.2.8. *Assume that* X *is the set of all factors involved in a study of a system, and* D *the set of all grey incidence operators. Then* (X, D) *is called the space of grey incidence factors of the system.*

5.3 Metric Spaces

The space, consisting of system factors and grey incidence operators, forms a base for grey incidence analysis. On such a base, comparisons and evaluations can be done in order to study systems' behaviors of factors.

If each factor in a space of grey incidence factors is seen as a point in the space without size and volume, and each data value of the factor, observed at a different time moment, different index, or different object, is seen as the coordinate of the point, we will be able to study the relationship between factors or between factors and the system's characteristics in a special n-dimensional space. In this way, the relevant degree of grey incidence can be defined by using the distance function in the n-dimensional space.

Definition 5.3.1. *Assume that X, Y, and Z are points in the n-dimensional Euclidean space $\mathbf{R}^n$, and the real number $d(X, Y)$ satisfies the following.*

1. $d(X, Y) \geq 0$, $d(X, Y) = 0 \Longleftrightarrow X = Y$;

2. $d(X, Y) = d(Y, X)$;

3. $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

Then $d(X, Y)$ is called a distance in the n-dimensional Euclidean space $\mathbf{R}^n$ and d a distance function.

Proposition 5.3.1. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

and

$$Y = (y(1), y(2), ..., y(n))$$

are points in the n-dimensional Euclidean space $\mathbf{R}^n$. Define

$$d_1(X, Y) = |x(1) - y(1)| + |x(2) - y(2)| + \cdots + |x(n) - y(n)|;$$

$$d_2(X, Y) = [|x(1) - y(1)|^2 + |x(2) - y(2)|^2 + \cdots + |x(n) - y(n)|^2 \,]^{\frac{1}{2}};$$

$$d_3(X, Y) = \frac{d_1(X, Y)}{1 + d_1(X, Y)};$$

$$d_p(X, Y) = [|x(1) - y(1)|^p + |x(2) - y(2)|^p + \cdots + |x(n) - y(n)|^p \,]^{\frac{1}{p}};$$

and

$$d_\infty(X, Y) = \max \{|x(k) - y(k)|| k = 1, 2, ..., n\}.$$

Then $d_1(X, Y), d_2(X, Y), d_3(X, Y), d_p(X, Y)$, and $d_\infty(X, Y)$ are all distances in the n-dimensional Euclidean space $\mathbf{R}^n$.

Definition 5.3.2. *Let*

$$X = (x(1), x(2), ..., x(n))$$

be a point in the n-dimensional Euclidean space, and

$$O = (0, 0, ..., 0)$$

the origin. Then the distance $d(X, O)$ between X and O is called the norm of the point X, denoted $\|X\|$.

Related to the distances in Proposition 5.3.1, we have the following commonly used norms.

1. 1-norm: $||X||_1 = \sum\limits_{k=1}^{n} |x(k)|$;
2. 2-norm: $||X||_2 = \left[\sum\limits_{k=1}^{n} |x(k)|^2 \right]^{\frac{1}{2}}$;

3. p-norm: $||X||_p = \left[\sum_{k=1}^{n} |x(k)|^p \right]^{\frac{1}{p}}$; and

4. $\infty - norm: ||X||_\infty = \max\{|x(k)||k = 1, 2, ..., n\}$.

Definition 5.3.3. *Assume that $X(t), Y(t)$, and $Z(t)$ are continuous functions defined on a set I of numbers. If the real number $d(X(t), Y(t))$ satisfies the following,*

1. $d(X(t), Y(t)) \geq 0$, $d(X(t), Y(t)) = 0 \Longleftrightarrow \forall t, X(t) = Y(t)$;

2. $d(X(t), Y(t)) = d(Y(t), X(t))$;

3. $d(X(t), Z(t)) \leq d(X(t), Y(t)) + d(Y(t), Z(t))$,

then $d(X(t), Y(t))$ is called a distance in the space consisting of all continuous functions defined on I.

Here, we have treated the functions $X(t)$ and $Y(t)$ as two points in the functional space.

Proposition 5.3.2. *Assume that $X(t)$ and $Y(t)$ are continuous functions defined on interval $[a, b]$. Define*

$$d_1(X(t), Y(t)) = \int_a^b |X(t) - Y(t)| dt;$$

$$d_2(X(t), Y(t)) = [\int_a^b |X(t) - Y(t)|^2 dt\ ]^{\frac{1}{2}};$$

$$d_3(X(t), Y(t)) = \frac{d_1(X(t), Y(t))}{1 + d_1(X(t), Y(t))};$$

$$d_4(X(t), Y(t)) = \int_a^b \frac{|X(t) - Y(t)|}{1 + |X(t) - Y(t)|} dt;$$

$$d_p(X(t), Y(t)) = \left[\int_a^b |X(t) - Y(t)|^p dt \right]^{\frac{1}{p}};$$

and

$$d_\infty(X(t), Y(t)) = \max\{|X(t) - Y(t)||t \in [a, b]\}.$$

Then $d_1(X(t), Y(t))$, $d_2(X(t), Y(t))$, $d_3(X(t), Y(t))$, $d_4(X(t), Y(t))$, $d_p(X(t), Y(t))$, and $d_\infty(X(t), Y(t))$ are all distances in the functional space of all continuous functions defined on I.

Similar to notes 1 to 4 beneath Definition 5.3.2, we can define the norm on the functional space of all continuous functions using distance functions d_1 to d_∞ as defined in Proposition 5.3.2. More specifically, we have

Definition 5.3.4. *Assume that $X(t)$ is a continuous function defined on the interval $[a, b]$, and 0 the zero function defined on the interval $[a, b]$. Then $d(X(t), 0)$ is called the norm of the continuous function $X(t)$, denoted as*

$$||X(t)|| = d(X(t), 0).$$

Accordingly, we have the following commonly used norms.

1. $||X(t)||_1 = \int_a^b |X(t)|dt$;
2. $||X(t)||_2 = [\int_a^b |X(t)|^2 dt\ ]^{\frac{1}{2}}$;
3. $||X(t)||_p = [\int_a^b |X(t)|^p dt\ ]^{\frac{1}{p}}$;
4. $||X(t)||_\infty = \max\{|X(t)||t \in [a, b]\}$.

Besides,

$$||X(t)|| = \frac{\int_a^b |X(t)|dt}{1 + \int_a^b |X(t)|dt}$$

and

$$||X(t)|| = \int_a^b \frac{|X(t)|}{1 + |X(t)|} dt$$

are important for us to derive the definition of degree of grey incidences.

5.4 Degrees of Grey Incidences

First, let us idenfity each sequence of data with its graph.

Definition 5.4.1. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of data. Then

$$X = \quad \cup_{k=1}^{n-1}\{(t, x(k) + (t-k)(x(k+1) - x(k))) \mid t \in [k, k+1]\}$$

is called the zigzagged line corresponding to the sequence X.

Here, we have used the same symbol X to represent the original sequence and its zigzagged line. For the sake of convenience for our discussion, we do not always distinguish a sequence and its zigzagged line.

Proposition 5.4.1. *Assume that sequence X_0 of data, describing a system's characteristic behaviors, is increasing, and X_i is a sequence of relevant factors' behaviors. Then,*

1. *When X_i is increasing, X_i and X_0 are positively correlated; and*
2. *When X_i is decreasing, X_i and X_0 are negatively correlated.*

Because negatively correlated sequences can be transformed into positively correlated sequences by using a reversing operator or a reciprocating operator, as defined in Section 5.2, we put our emphasis on the study of positively correlated relationships.

Definition 5.4.2. *Let $X = (x(1), x(2), ..., x(n))$ be a sequence of data.*

1.The following

$$\alpha = x(k) - x(k-1),$$

for $k = 1, 2, ..., n$, *is called the slope of* X *on the interval* $[k-1, k]$.

2. The following

$$\alpha = \frac{x(s) - x(k)}{s - k}, s > k,$$

$k = 1, 2, ..., n$, *is called the mean slope of* X *on the interval* $[k, s]$. *And*

3. The following

$$\alpha = \frac{1}{n-1}[x(n) - x(1)]$$

is called the mean slope of X.

Theorem 5.4.1. *Assume that* X_i *and* X_j *are non-negative increasing sequences of data,* $X_j = X_i + c$, *where* c *is a nonzero constant, and* D_1 *is an initialing operator. Let*

$$Y_i = X_i D_1, Y_j = X_j D_1$$

be the initial images of X_i *and* X_j, α_i *and* α_j *the mean slopes of* X_i *and* X_j, *and* β_i *and* β_j *the mean slopes of* Y_i *and* Y_j, *respectively. Then, the following must hold true.*

1. $\alpha_i = \alpha_j$;

2. When $c < 0$, $\beta_i < \beta_j$. *And when* $c > 0$, $\beta_i > \beta_j$.

Proof. 1. Assume that

$$X_i = (x_i(1), x_i(2), ..., x_i(n)),$$

and

$$\begin{aligned} X_j &= (x_j(1), x_j(2), ..., x_j(n)) \\ &= (x_i(1) + c, x_i(2) + c, ..., x_i(n) + c). \end{aligned}$$

From Definition 5.4.2, it follows that

$$\alpha_i = \frac{1}{n-1}(x_i(n) - x_i(1))$$

and

$$\begin{aligned} \alpha_j &= \tfrac{1}{n-1}(x_j(n) - x_j(1)) \\ &= \tfrac{1}{n-1}(x_i(n) + c - x_i(1) - c) \\ &= \tfrac{1}{n-1}(x_i(n) - x_i(1)) = \alpha_i. \end{aligned}$$

2. From

$$Y_i = X_i D_1 = \left(\frac{x_i(1)}{x_i(1)}, \frac{x_i(2)}{x_i(1)}, ..., \frac{x_i(n)}{x_i(1)} \right)$$

and

$$\begin{aligned} Y_j &= X_j D_1 = \left(\frac{x_j(1)}{x_j(1)}, \frac{x_j(2)}{x_j(1)}, ..., \frac{x_j(n)}{x_j(1)} \right) \\ &= \left(\frac{x_i(1)+c}{x_i(1)+c}, \frac{x_i(2)+c}{x_i(1)+c}, ..., \frac{x_i(n)+c}{x_i(1)+c} \right), \end{aligned}$$

it follows that

$$\begin{aligned} \beta_i &= \frac{1}{n-1} \left(\frac{x_i(n)}{x_i(1)} - \frac{x_i(1)}{x_i(1)} \right) \\ &= \frac{1}{x_i(1)(n-1)} \left(x_i(n) - x_i(1) \right) = \frac{1}{x_i(1)} \alpha_i \end{aligned}$$

and

$$\begin{aligned} \beta_j &= \frac{1}{n-1} \left(\frac{x_i(n)+c}{x_i(1)+c} - \frac{x_i(1)+c}{x_i(1)+c} \right) \\ &= \frac{1}{(x_i(1)+c)(n-1)} \left(x_i(n) - x_i(1) \right) = \frac{1}{x_i(1)+c} \alpha_i. \end{aligned}$$

When $c < 0$, $x_i(1) > x_i(1) + c$, it follows that

$$\frac{1}{x_i(1)} < \frac{1}{x_i(1)+c}.$$

So, $\beta_i < \beta_j$. When $c > 0$, $x_i(1) < x_i(1) + c$, it follows that

$$\frac{1}{x_i(1)} > \frac{1}{x_i(1)+c}.$$

So, $\beta_i > \beta_j$. □

This theorem reflects the following characteristics of increasing sequences. When the absolute amounts of increase of two increasing sequences are the same, the sequence with a smaller initial value will increase faster than the sequence with a greater initial value. In order to keep a same relative rate of increase, the absolute amount of increase of the sequence with the greater initial value must be greater than that of the sequence with a smaller initial value.

Definition 5.4.3. *Assume that*

$$X_0 = \left(x_0(1),\ x_0(2), ...,\ x_0(n) \right)$$

is a sequence of data representing a system's characteristics, and

$$X_i \quad = (x_i(1), x_i(2), ..., x_i(n)), \ i = 1, 2, ..., n$$

are sequences of relevant factors. For a given real number $\gamma(x_0(k), x_i(k))$, *if the real number*

$$\gamma(X_0, X_i) = \frac{1}{n}\sum_{k=1}^{n}\gamma(x_0(k), x_i(k))$$

satisfies

1. The property of normality

$$0 < \gamma(X_0, X_i) \leq 1,$$

and

$$\gamma(X_0, X_i) = 1 \Longleftrightarrow X_0 = X_i.$$

2. The property of wholeness

$$\forall X_i, X_j \in X = \{X_s | s = 1, 2, \cdots m; m \geq 2\},$$

we have

$$\gamma(X_i, X_j) \neq \gamma(X_j, X_i), (i \neq j).$$

3. The property of pair symmetry. For $X_i, X_j \in X$, *then*

$$\gamma(X_i, X_j) = \gamma(X_j, X_i) \Longleftrightarrow X = \{X_i, X_j\}.$$

4. The property of closeness. The smaller

$$|x_0(k) - x_i(k)|$$

is, the larger

$$\gamma(x_0(k), x_i(k)) \ .$$

Then $\gamma(X_0, X_i)$ *is called the degree of grey incidence of* X_i *with respect to* X_0, *and* $\gamma(x_0(k), x_i(k))$ *the incidence coefficient of* X_i *with respect to* X_0 *at point* k.

In the axioms for grey incidences, that is, conditions 1 to 4 in Definition 5.4.3,

1. $\gamma(X_0, X_i) \in (0, 1]$ implies that any two behavioral sequences of a system cannot be absolutely not related;

2. The property of wholeness reflects the influence of the environment on comparisons of grey incidences. When the environment changes, the degree of grey incidences also changes accordingly;

3. The property of pair symmetry implies that when the set of grey incidence factors contain only two sequences, they satisfy the property of symmetry; and

4. The property of closeness is a requirement for the quantification of degree of incidence.

Theorem 5.4.2. *Assume that* $m+1$ *behavioral sequences of a system are given*

$$X_i = (x_i(1), x_i(2), ..., x_i(n)), i = 0, 1, 2, ..., m.$$

For $\zeta \in (0,1)$, *define*

$$\begin{aligned}\gamma_{0i} &= \gamma(x_0(k), x_i(k)) \\ &= \left[\min_i \min_k |x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|\right] \\ &\div \left[|x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|\right]\end{aligned}$$

and

$$\gamma(X_0, X_i) = \frac{1}{n}\sum_{k=1}^{n} \gamma(x_0(k), x_i(k)).$$

Then $\gamma(X_0, X_i)$ *satisfies the four axioms for grey incidences, where* ζ *is called the distinguishing coefficient.*

Proof. 1. The property of normality. If

$$|x_0(k) - x_i(k)| = \min_i \min_k |x_0(k) - x_i(k)|,$$

then

$$\gamma(x_0(k), x_i(k)) = 1.$$

If

$$|x_0(k) - x_i(k)| \neq \min_i \min_k |x_0(k) - x_i(k)|,$$

then

$$|x_0(k) - x_i(k)| > \min_i \min_k |x_0(k) - x_i(k)|.$$

Therefore,

$$\begin{aligned}&\min_i \min_k |x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)| \\ &< |x_0(k) - x_i(k)| + \zeta \max_i \max_k |x_0(k) - x_i(k)|.\end{aligned}$$

So,

$$\gamma(x_0(k), x_i(k)) < 1.$$

It is obvious that for any k, $\gamma\left(x_0\left(k\right), x_i\left(k\right)\right) > 0$. Hence,

$$0 < \gamma(X_0, X_i) < 1.$$

2. The property of wholeness. If

$$X = \{X_s | s = 0, 1, 2, ..., m; m \geq 2\},$$

then for any X_{s_1},$X_{s_2} \in X$, in general, we have

$$\max_i \max_k |x_{s_1}(k) - x_i(k)| \neq \max_i \max_k |x_{s_2}(k) - x_i(k)|.$$

So, the property of wholeness holds true.

3. The property of pair symmetry. If $X = \{X_0, X_1\}$, then

$$|x_0(k) - x_1(k)| = |x_1(k) - x_0(k)|$$

and

$$\max_i \max_k |x_0(k) - x_i(k)| = \max_i \max_k |x_1(k) - x_i(k)|,$$

where $i = 1$ in the left end, and $i = 0$ on the right end. So,

$$\gamma(X_0, X_1) = \gamma(X_1, X_0).$$

4. The property of closeness. It is obvious. □

The degree $\gamma\left(X_0, X_i\right)$ of grey incidence is often written as γ_{0i}, and the incidence coefficient $\gamma\left(x_0\left(k\right), x_i\left(k\right)\right)$ at point k as $\gamma_{0i}(k)$.

Based on the algorithm described in Theorem 5.4.2, each computation for the degree of grey incidences can be accomplished by going through the following steps.

Step 1: Find the initial image (or average image) of each sequence. Let

$$X_i' = \frac{X_i}{x_i(1)} = \left(x_i'(1), x_i'(2), ..., x_i'(n)\right), i = 0, 1, 2, ..., m.$$

Step 2: Find difference sequences. Denote

$$\Delta_i(k) = |x_0'(k) - x_i'(k)|, \text{ and}$$

$$\Delta_i = \left(\Delta_i(1), \Delta_i(2), ..., \Delta_i(n)\right), i = 0, 1, 2, ..., m.$$

Step 3: Find the maximum and minimum differences. And write

$$M = \max_i \max_k \Delta_i(k), m = \min_i \min_k \Delta_i(k).$$

Step 4: Find incidence coefficients.

$$\gamma_{0i}(k) = \frac{m + \zeta M}{\Delta_i(k) + \zeta M},$$

for $\zeta \in (0,1)$, $k = 1,2,...,n; i = 1,2,...,m$.

Step 5: Compute the degree of incidences.

$$\gamma_{0i} = \frac{1}{n}\sum_{k=1}^{n}\gamma_{0i}(k), i = 0,1,2,...,m.$$

Example 5.4.1. In a study, some behavioral data for areas in industry, agriculture, transportation, and business are provided as follows.

In industry:

$$X_1 \quad = (x_1(1), x_1(2), x_1(3), x_1(4)) = (45.8, 43.4, 42.3, 41.9).$$

In agriculture:

$$X_2 \quad = (x_2(1), x_2(2), x_2(3), x_2(4)) = (39.1, 41.6, 43.9, 44.9).$$

In transportation:

$$X_3 \quad = (x_3(1), x_3(2), x_3(3), x_3(4)) = (3.4, 3.3, 3.5, 3.5).$$

And in business:

$$X_4 \quad = (x_4(1), x_4(2), x_4(3), x_4(4)) = (6.7, 6.8, 5.4, 4.7).$$

Compute the degree of grey incidence by using X_1 and X_2 as a system's characteristic sequences.

Solution. 1. First, we compute the degree of incidence for the case that X is treated as a system's characteristic sequence.

Step 1: Compute the initial image. From

$$X_i^{'} = X_i/x_i(1) = \left(x_i^{'}(1), x_i^{'}(2), x_i^{'}(3), x_i^{'}(4)\right),$$

$i = 1,2,3,4$, it follows that

$$X_1^{'} \quad = (1, 0.9475, 0.9235, 0.9138),$$

$$X_2^{'} \quad = (1, 1.063, 1.1227, 1.1483),$$

$$X_3^{'} \quad = (1, 0.97, 1.0294, 1.0294), \text{ and}$$

$$X_4^{'} \quad = (1, 1.0149, 0.805, 0.7).$$

Step 2: Compute difference sequences. From

$$\Delta_i(k) = |x_1'(k) - x_i'(k)|,$$

$i = 2, 3, 4$, it follows that

$$\begin{aligned}
\Delta_2 &= (0, 0.1155, 0.1992, 0.2335), \\
\Delta_3 &= (0, 0.0225, 0.1059, 0.1146), \text{ and} \\
\Delta_4 &= (0, 0.0674, 0.1185, 0.2148).
\end{aligned}$$

Step 3: Compute the difference between the two extremes.

$$M = \max_i \max_k \Delta_i(k) = 0.2335, \text{and } m = \min_i \min_k \Delta_i(k) = 0.$$

Step 4: Compute the incidence coefficients. Take $\zeta = 0.5$. Then we have

$$\gamma_{1i} = \frac{0.11675}{\Delta_i(k) + 0.11675},$$

$i = 2, 3, 4$. Therefore,

$$\begin{aligned}
&\gamma_{12}(1) = 1, \quad \gamma_{12}(2) = 0.503, \quad \gamma_{12}(3) = 0.3695, \quad \gamma_{12}(4) = 0.3333; \\
&\gamma_{13}(1) = 1, \quad \gamma_{13}(2) = 0.8384, \quad \gamma_{13}(3) = 0.5244, \quad \gamma_{13}(4) = 0.504; \\
&\gamma_{14}(1) = 1, \quad \gamma_{14}(2) = 0.634, \quad \gamma_{14}(3) = 0.4963, \quad \gamma_{14}(4) = 0.352.
\end{aligned}$$

Step 5: Compute the degree of grey incidences.

$$\begin{aligned}
\gamma_{12} &= \tfrac{1}{4}\sum_{k=1}^{4} \gamma_{12}(k) = 0.551, \\
\gamma_{13} &= \tfrac{1}{4}\sum_{k=1}^{4} \gamma_{13}(k) = 0.717, \text{ and} \\
\gamma_{14} &= \tfrac{1}{4}\sum_{k=1}^{4} \gamma_{14}(k) = 0.621.
\end{aligned}$$

2. For the case when X_2 is seen as the system's characteristic: from

$$\Delta_i(k) = |x_2'(k) - x_i'(k)|,$$

$i = 1, 3, 4$, it follows that

$$\begin{aligned}
\Delta_1 &= (0, 0.1155, 0.1992, 0.2335), \\
\Delta_3 &= (0, 0.093, 0.0933, 0.1189), \\
\Delta_4 &= (0, 0.0481, 0.3177, 0.4483).
\end{aligned}$$

So,

$$M = \max_i \max_k \Delta_i(k) = 0.4483,$$

$$m = \min_i \min_k \Delta_i(k) = 0.$$

Now, take $\zeta = 0.5$; we obtain that

$$\gamma_{2i}(k) = \frac{0.22415}{\Delta_i(k) + 0.22415}.$$

Hence,

$$\gamma_{21}(1) = 1, \quad \gamma_{21}(2) = 0.66, \quad \gamma_{21}(3) = 0.53, \quad \gamma_{21}(4) = 0.489;$$

$$\gamma_{23}(1) = 1, \quad \gamma_{23}(2) = 0.706, \quad \gamma_{23}(3) = 0.706, \quad \gamma_{23}(4) = 0.653;$$

$$\gamma_{24}(1) = 1, \quad \gamma_{24}(2) = 0.823, \quad \gamma_{24}(3) = 0.415, \quad \gamma_{24}(4) = 0.333;$$

and

$$\gamma_{21} = \tfrac{1}{4}\sum_{k=1}^{4} \gamma_{21}(k) = 0.670,$$

$$\gamma_{23} = \tfrac{1}{4}\sum_{k=1}^{4} \gamma_{23}(k) = 0.766,$$

$$\gamma_{24} = \tfrac{1}{4}\sum_{k=1}^{4} \gamma_{24}(k) = 0.643.$$

Combining the result $\gamma_{12} = 0.551$ in 1, it is obvious that $\gamma_{12} \neq \gamma_{21}$, which is exactly the property of wholeness in Definition 5.4.3.

5.5 Absolute Degree of Grey Incidence

The first result in this section establishes a relationship between behaviors of a sequence and its zigzagged line.

Proposition 5.5.1. *Assume that*

$$X_i = (x_i(1), x_i(2), ..., x_i(n)),$$

stands for a behavioral sequence of data, and that the zigzagged line

$$(x_i(1) - x_i(1), x_i(2) - x_i(1), ..., x_i(n) - x_i(1))$$

is denoted as $X_i - x_i(1)$. *Let*

$$s_i = \int_1^n (X_i - x_i(1))\, dt.$$

Then

1. *When* X_i *is an increasing sequence,* $s_i \geq 0$;
2. *When* X_i *is a decreasing sequence,* $s_i \leq 0$; *and*
3. *When* X_i *is a vibrating sequence, the sign of* s_i *is not fixed.*

The proof of this proposition is based on definitions of increasing, decreasing, and vibrating sequences and properties of integrations. Here, all the details are omitted.

Definition 5.5.1. *Let* $X_i = (x_i(1), x_i(2), ..., x_i(n))$ *be the same as in Proposition 5.5.1 and* D *a sequence operator*

$$X_i D = (x_i(1)d, x_i(2)d, ..., x_i(n)d)\,,\ k = 1, 2, ..., n.$$

where

$$x_i(k)d = x_i(k) - x_i(1).$$

Then D *is called a zero starting point operator with* $X_i D$ *as the image of zero starting point of* X_i*, denoted*

$$X_i^0 = \left(x_i^0(1), x_i^0(2), ..., x_i^0(n)\right).$$

Proposition 5.5.2. *Assume that the images of the zero starting point of two behavioral sequences*

$$X_i = (x_i(1), x_i(2), ..., x_i(n)) \text{ and } X_j = (x_j(1), x_j(2), ..., x_j(n))$$

are

$$X_i^0 = \left(x_i^0(1), x_i^0(2), \cdots, x_i^0(n)\right) \text{ and } X_j^0 = \left(x_j^0(1), x_j^0(2), \cdots, x_j^0(n)\right),$$

respectively. Let

$$s_i - s_j = \int_1^n \left(X_i^0 - X_j^0\right) dt.$$

Then the following hold true.

1. *If* X_i^0 *is always above* X_j^0*, then* $s_i - s_j \geq 0$;
2. *If* X_i^0 *is always underneath* X_j^0*, then* $s_i - s_j \leq 0$; *and*
3. *If* X_i^0 *and* X_j^0 *alternate their positions, the sign of* $s_i - s_j$ *is not fixed.*

Definition 5.5.2. *The sum of time intervals between consecutive observations of a sequence* X_i *is called the length of* X_i.

It should be noted that two sequences of the same length may not have the same number of observations. For example,

$$X_1 = (x_1(1), x_1(3), x_1(6))\,,$$

$$X_2 = (x_2(1), x_2(3), x_2(5), x_2(6))\,,$$

$$X_3 = (x_3(1), x_3(2), x_3(3), x_3(4), x_3(5), x_3(6))\,.$$

Even though these sequences all have 6 as their length, they have different numbers of observation values.

Definition 5.5.3. *Assume that two sequences X_i and X_j are of the same length, and s_i and s_j are defined as in Proposition 5.5.1. Then*

$$\varepsilon_{ij} = \frac{1 + |s_i| + |s_j|}{1 + |s_i| + |s_j| + |s_i - s_j|}$$

is called the absolute degree of grey incidence of X_i and X_j, or absolute degree of incidence for short.

At this point, we have only introduced the definition of absolute degree of grey incidence of same length sequences. As for sequences of different lengths, several methods can be used to define this concept. For example, one can either delete the extra values of the longer sequence, or employ the grey modeling method GM(1, 1) (see Chapter 7 for more details) developed for predictions to prolong the shorter sequence to the length of the longer sequence so that the absolute degree of grey incidence can be defined. However, all these methods would lead to different values of absolute degree of grey incidence.

Theorem 5.5.1. *The absolute degree of grey incidence*

$$\varepsilon_{ij} = \frac{1 + |s_i| + |s_j|}{1 + |s_i| + |s_j| + |s_i - s_j|}$$

satisfies the properties of normality, pair symmetry, and closeness, but not wholeness.

Proof. 1. The property of normality. It is obvious that $\varepsilon_{ij} > 0$ and $|s_i - s_j| > 0$. So, $\varepsilon_{ij} \leq 1$.

2. The property of pair symmetry. From the fact that

$$|s_i - s_j| = |s_j - s_i|,$$

it follows that $\varepsilon_{ij} = \varepsilon_{ji}$.

3. The property of closeness: It is obvious.

4. Because the absolute degree of grey incidence is only a measure for the correlation between the sequences X_i and X_j without involving other considered factors, there is no problem of wholeness here. □

Proposition 5.5.3. *Assume that the lengths of two sequences X_i and X_j are the same. Let*

$$X_i' = X_i - a, \; X_j' = X_j - b,$$

where a and b are constants. If the absolute degree of grey incidences of X_i' and X_j' is ε_{ij}', then $\varepsilon_{ij}' = \varepsilon_{ij}$.

In fact, when X_i and X_j are moved horizontally, the values of s_i, s_j, and $s_i - s_j$ are not changed. So, ε_{ij} is not changed.

Definition 5.5.4. *If the time intervals of any two consecutive observation values of a sequence X have the same length, then X is called an equal time interval sequence.*

Lemma 5.5.1. *Assume that X is an equal-time interval sequence. If the time interval length $\ell \neq 1$, then the time axis transformation*

$$t: \quad T \Rightarrow T, t \mapsto t/\ell$$

can transform X into an 1-time-interval sequence.

Proof. Under the transformation defined here,

$$X_i = (x_i(\ell), x_i(2\ell), ..., x_i(n\ell))$$

is transformed to the following with t changed to t/ℓ and $k\ell$ changed to $k\ell/\ell = k$.

$$X_i' = (x_i(1), x_i(2), ..., x_i(n)).$$

X_i' is an 1-time-interval sequence. □

Lemma 5.5.2. *Assume that X_i and X_j are 1-time-interval sequences of the same length, and*

$$X_i^0 = \left(x_i^0(1), x_i^0(2), ..., x_i^0(n)\right) \text{ and } X_j^0 = \left(x_j^0(1), x_j^0(2), ..., x_j^0(n)\right)$$

are the zero images of X_i and X_j. Then,

$$|s_i| = \left|\sum_{k=2}^{n-1} x_i^0(k) + \frac{1}{2} x_i^0(n)\right|,$$

$$|s_j| = \left|\sum_{k=2}^{n-1} x_j^0(k) + \frac{1}{2} x_j^0(n)\right|,$$

and

$$|s_i - s_j| = \left|\sum_{k=2}^{n-1} [x_i^0(k) - x_j^0(k)] + \frac{1}{2}[x_i^0(n) - x_j^0(n)]\right|.$$

Proof. We complete our argument with three situations.

1. X_i and X_j are either both increasing or both decreasing sequences, and the zigzagged lines X_i^0 and X_j^0 do not intersect.

In this case, $|s_i|, |s_j|$, and $|s_i - s_j|$ are determined by areas of the following triangles with curvilinear sides, respectively.

$$X = 0, \; X = X_i^0, \text{ and } t = n;$$

$$X = 0, \; X = X_j^0, \text{ and } t = n;$$

and

$$X = X_i^0,\ X = X_j^0,\ \text{and } t = n.$$

They are sums of little areas of $n-1$ small trapezoids of height 1. The little trapezoids for $|s_i|$ have their base lengths as follows.

$$0, |x_i^0(2)|, |x_i^0(3)|, ..., |x_i^0(n)|.$$

The little trapezoids for $|s_j|$ have their base lengths as follows.

$$0, |x_j^0(2)|, |x_j^0(3)|, ..., |x_j^0(n)|.$$

And the little trapezoids for $|s_i - s_j|$ have their base lengths as follows:

$$0, |x_i^0(2) - x_j^0(2)|, |x_i^0(3) - x_j^0(3)|, \cdots, |x_i^0(n) - x_j^0(n)|.$$

So,

$$\begin{aligned}
|s_i| &= \tfrac{1}{2}|x_i^0(2)| + \tfrac{1}{2}[|x_i^0(2)| + |x_i^0(3)|] + \cdots + \tfrac{1}{2}[|x_i^0(n-1)| + |x_i^0(n)|] \\
&= \sum_{k=2}^{n-1} |x_i^0(k)| + \tfrac{1}{2}|x_i^0(n)|, \\
|s_j| &= \tfrac{1}{2}|x_j^0(2)| + \tfrac{1}{2}[|x_j^0(2)| + |x_j^0(3)|] + \cdots + \tfrac{1}{2}[|x_j^0(n-1)| + |x_j^0(n)|] \\
&= \sum_{k=2}^{n-1} |x_j^0(k)| + \tfrac{1}{2}|x_j^0(n)|,
\end{aligned}$$

and

$$\begin{aligned}
|s_i - s_j| &= \tfrac{1}{2}|x_i^0(2) - x_j^0(2)| + \tfrac{1}{2}[|x_i^0(2) - x_j^0(2)| + |x_i^0(3) - x_j^0(3)|] \\
&\quad + \cdots + \tfrac{1}{2}[|x_i^0(n-1) - x_j^0(n-1)| + |x_i^0(n) - x_j^0(n)|] \\
&= \sum_{k=2}^{n-1} |x_i^0(k) - x_j^0(k)| + \tfrac{1}{2}|x_i^0(n) - x_j^0(n)|.
\end{aligned}$$

From the assumption that X_i and X_j are either both increasing or both decreasing and that the zigzagged lines X_i^0 and X_j^0 do not intersect, it follows that for $k = 2, 3, ..., n$, all $x_i^0(k)$s have the same sign, all $x_j^0(k)$s have the same sign and $x_i^0(k) - x_j^0(k)$s have the same sign. Therefore, we have

$$\begin{aligned}
|s_i| &= \sum_{k=2}^{n-1} |x_i^0(k)| + \tfrac{1}{2}|x_i^0(n)| = \left| \sum_{k=2}^{n-1} x_i^0(k) + \tfrac{1}{2}x_i^0(n) \right|, \\
|s_j| &= \sum_{k=2}^{n-1} |x_j^0(k)| + \tfrac{1}{2}|x_j^0(n)| = \left| \sum_{k=2}^{n-1} x_j^0(k) + \tfrac{1}{2}x_j^0(n) \right|,
\end{aligned}$$

and

$$\begin{aligned} |s_i - s_j| &= \sum_{k=2}^{n-1} |x_i^0(k) - x_j^0(k)| + \tfrac{1}{2}|x_i^0(n) - x_j^0(n)| \\ &= \left| \sum_{k=2}^{n-1} [x_i^0(k) - x_j^0(k)] + \tfrac{1}{2}[x_i^0(n) - x_j^0(n)] \right| . \end{aligned}$$

2. Both X_i and X_j are vibrating sequences and the zigzagged lines X_i^0 and X_j^0 do not intersect. Because X_i^0 and X_j^0 do not intersect, from the discussion in 1 it follows that

$$|s_i - s_j| = \left| \sum_{k=2}^{n-1} [x_i^0(k) - x_j^0(k)] + \frac{1}{2}[x_i^0(n) - x_j^0(n)] \right| .$$

In the following, we look at $|s_i|$ and $|s_j|$.

Because X_i is a vibrating sequence, s_i equals the algebraic sum of various parts, bounded by $X = 0, X = X_i^0$, and $t = n$ while taking the parts above $X = 0$ positive and the parts beneath $X = 0$ negative.

Assume that $x_i^0(k), k = 2, 3, ..., n$, changes signs only once, and

$$(x_i^0(m), x_i^0(m+1))$$

is the only pair of points where the change of signs occurred. Assume that $x_i^0(m) > 0$ and $x_i^0(m+1) < 0$. Then

for $k = 2, 3, ..., m-1$, $x_i^0(k) > 0$;

for $k = m+1, m+2, ..., n$, $x_i^0(k) < 0$.

Denote

$$s_{im} = \int_m^{m+1} X_i^0 dt.$$

Then

$$\begin{aligned} s_i &= \tfrac{1}{2}|x_i^0(2)| + \tfrac{1}{2}[|x_i^0(2)| + |x_i^0(3)|] \\ &+ \cdots + \tfrac{1}{2}[|x_i^0(m-1)| + |x_i^0(m)|] + s_{im} \\ &- \tfrac{1}{2}[|x_i^0(m+1)| + |x_i^0(m+2)|] \\ &- \cdots - \tfrac{1}{2}[|x_i^0(n-1)| + |x_i^0(n)|]. \end{aligned}$$

We now compute s_{im}. As shown in Figure 5.2, the equation of the straight-line AB is

$$X = x_i^0(m) + (t - m)(x_i^0(m+1) - x_i^0(m)).$$

So, the intersection of AB and $X = 0$ is

$$C\left(m + \frac{x_i^0(m)}{x_i^0(m+1) - x_i^0(m)}, 0\right).$$

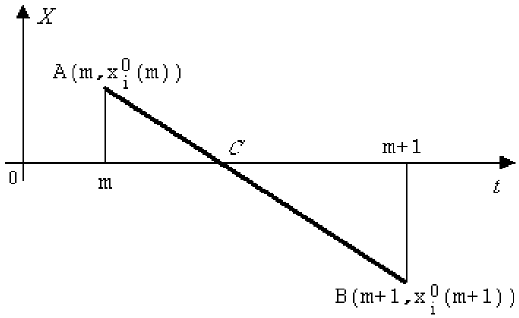

FIGURE 5.2. Graph needed for the computation of s_{im}

Therefore,

$$\begin{aligned} s_{im} &= \tfrac{1}{2}|x_i^0(m)| \cdot \left| \frac{x_i^0(m)}{x_i^0(m+1) - x_i^0(m)} \right| \\ &\quad - \tfrac{1}{2}|x_i^0(m+1)| \cdot \left[1 - \left| \frac{x_i^0(m)}{x_i^0(m+1) - x_i^0(m)} \right| \right] \\ &= \tfrac{1}{2}[|x_i^0(m)| + |x_i^0(m+1)|] \cdot \left| \frac{x_i^0(m)}{x_i^0(m+1) - x_i^0(m)} \right| \\ &\quad - \tfrac{1}{2}|x_i^0(m+1)|. \end{aligned}$$

Because $x_i^0(m) > 0$ and $x_i^0(m+1) < 0$, we have that

$$\begin{aligned} \left| \frac{x_i^0(m)}{x_i^0(m+1) - x_i^0(m)} \right| &= \frac{|x_i^0(m)|}{|x_i^0(m+1) - x_i^0(m)|} \\ &= \frac{|x_i^0(m)|}{|x_i^0(m+1)| + |x_i^0(m)|}. \end{aligned}$$

Hence,

$$s_{im} = \frac{1}{2}|x_i^0(m)| - \frac{1}{2}|x_i^0(m+1)|.$$

Now, by considering

$$\left| x_i^0(k) \right| = x_i^0(k),$$

for $k = 2, 3, ..., m$, and

$$\left| x_i^0(k) \right| = -x_i^0(k),$$

for $k = m+1, m+2, ..., n$, it follows that

$$|s_i| = \left| \sum_{k=2}^{n-1} x_i^0(k) + \frac{1}{2} x_i^0(n) \right|.$$

As for the case with $x_i^0(m) < 0$ and $x_i^0(m+1) > 0$, a similar argument can be given.

As for the case with several pairs of points of sign changes, we can consider each pair individually, and obtain the following,

$$|s_i| = \left|\sum_{k=2}^{n-1} x_i^0(k) + \frac{1}{2}x_i^0(n)\right|.$$

Similarly, it can be proven that when X_j is a vibrating sequence, we also have

$$|s_j| = \left|\sum_{k=2}^{n-1} x_j^0(k) + \frac{1}{2}x_j^0(n)\right|.$$

3. X_i and X_j are vibrating sequences with intersecting X_i^0 and X_j^0. From 2 we have already had

$$|s_i| = \left|\sum_{k=2}^{n-1} x_i^0(k) + \frac{1}{2}x_i^0(n)\right|,$$

and

$$|s_j| = \left|\sum_{k=2}^{n-1} x_j^0(k) + \frac{1}{2}x_j^0(n)\right|.$$

Now, $s_i - s_j$ equals the algebraic sum of various parts bounded by $X = X_i^0, X = X_j^0$, and $t = n$, taking the parts with X_i^0 on the top of X_j^0 positive and the other parts negative. Similar to 2, it can be proven that

$$|s_{i-}s_j| = \left|\sum_{k=2}^{n-1} [x_i^0(k) - x_j^0(k)] + \frac{1}{2}[x_i^0(n) - x_j^0(n)]\right|. \square$$

Theorem 5.5.2. *Assume that X_i and X_j are two sequences of the same length, same time distances, and equal time interval. Then,*

$$\begin{aligned}\varepsilon_{ij} \quad = \quad & 1 + \left|\sum_{k=2}^{n-1} x_i^0(k) + \tfrac{1}{2}x_i^0(n)\right| + \left|\sum_{k=2}^{n-1} x_j^0(k) + \tfrac{1}{2}x_j^0(n)\right| \\ & \div \left\{1 + \left|\sum_{k=2}^{n-1} x_i^0(k) + \tfrac{1}{2}x_i^0(n)\right| + \left|\sum_{k=2}^{n-1} x_j^0(k) + \tfrac{1}{2}x_j^0(n)\right| \right. \\ & \left. + \left|\sum_{k=2}^{n-1} [x_i^0(k) - x_j^0(k)] + \tfrac{1}{2}[x_i^0(n) - x_j^0(n)]\right| \right\}.\end{aligned}$$

Proof. From Lemma 5.5.1, it can be assumed that X_i and X_j are all 1-time interval sequences. So, from Lemma 5.5.2 and Definition 5.5.3, the result follows. $\square$

Theorem 5.5.3. *Assume that two sequences X_i and X_j have the same length, and that they have different lengths of time intervals or at least one of them is a non-equal-time-interval sequence. If the method of mean generation is used to fill in relevant blanks so that the sequences become sequences with the same relevant time steps and equal-time-intervals, then the absolute degree ε_{ij} of grey incidence is unchanged.*

Proof. This argument is completed with discussions of several cases.

Case 1. X_i and X_j are sequences with the same corresponding time steps and non-equal time intervals. Without loss of generality, we can assume that there is only one pair of points with 2 as their (time) distance and that all other intervals of consecutive entries have length 1. Assume that the zero starting point images of X_i and X_j are

$$X_i^0 = \left(x_i^0(1), x_i^0(2), ..., x_i^0(m), x_i^0(m+2), ..., x_i^0(n)\right)$$

and

$$X_j^0 = \left(x_j^0(1), x_j^0(2), ..., x_j^0(m), x_j^0(m+2), ..., x_j^0(n)\right),$$

respectively. Now, we only need to fill in the gaps $x_i^0(m+1)$ and $x_j^0(m+1)$. Define

$$x_i^0(m+1) = \frac{1}{2}\left[x_i^0(m) + x_i^0(m+2)\right],$$

and

$$x_j^0(m+1) = \frac{1}{2}\left[x_j^0(m) + x_j^0(m+2)\right].$$

Now X_i^0 and X_j^0 are all equal-time-interval sequences. Let

$$s_{im} = \int_m^{m+2} X_i^0 dt, s_{jm} = \int_m^{m+2} X_j^0 dt$$

and

$$s_{im} - s_{jm} = \int_m^{m+2} (X_i^0 - X_j^0)dt.$$

Before $x_i^0(m+1)$ and $x_j^0(m+1)$ are placed in the sequences, we have

$$s_{im} = x_i^0(m) + x_i^0(m+2),$$
$$s_{jm} = x_j^0(m) + x_j^0(m+2),$$

and

$$s_{im} - s_{jm} = [x_i^0(m) - x_j^0(m)] + [x_i^0(m+2) - x_j^0(m+2)].$$

After $x_i^0(m+1)$ and $x_j^0(m+1)$ are placed in the sequences, we have

$$\begin{aligned} s_{im} &= \tfrac{1}{2}[x_i^0(m) + x_i^0(m+1)] + \tfrac{1}{2}[x_i^0(m+1) + x_i^0(m+2)] \\ &= x_i^0(m+1) + \tfrac{1}{2}[x_i^0(m) + x_i^0(m+2)] \\ &= \tfrac{1}{2}[x_i^0(m) + x_i^0(m+2)] + \tfrac{1}{2}[x_i^0(m) + x_i^0(m+2)] \\ &= x_i^0(m) + x_i^0(m+2). \end{aligned}$$

Similarly, we can prove that the values for s_{jm} and $s_{im} - s_{jm}$ do not change either. Therefore, $|s_i|$, $|s_j|$, and $|s_i - s_j|$ do not change, nor does ε_{ij}.

Similar to Lemma 5.5.2, it is not difficult to reason that no matter whether the pairs of points

$$\left(x_i^0(m), x_i^0(m+2)\right),$$

$$\left(x_j^0(m), x_j^0(m+2)\right),$$

and

$$\left(x_i^0(m) - x_j^0(m), x_i^0(m+2) - x_j^0(m+2)\right)$$

represent sign changes, the conclusion above holds true.

Case 2. X_i and X_j are equal-time-interval sequences of different lengths with unequal corresponding time intervals. Without loss of generality, we may assume that X_i is a 1-time-interval sequence and X_j a 2-time-interval sequence. Let the zero starting point images of X_i and X_j be respectively given as

$$X_i^0 = \left(x_i^0(1), x_i^0(2), x_i^0(3), ..., x_i^0(2n+1)\right)$$

and

$$X_j^0 = \left(x_j^0(1), x_j^0(3), x_j^0(5), ..., x_j^0(2n+1)\right).$$

We now only need to fill in blanks in X_j^0 with

$$x_j^0(2k) = \frac{1}{2}\left[x_j^0(2k-1) + \ x_j^0(2k+1)\right],$$

$k = 1, 2, ..., n$, to transform X_i^0 and X_j into equal-time-interval sequences with corresponding time intervals the same. The rest of the proof is the same as in Case 1 and is omitted here.

Case 3. X_i and X_j have different time intervals and at least one of them is a non-equal-time-interval sequence. In this case, the method of mean generations can be used as mentioned in Case 2 to fill in the blanks in the sequences so that X_i^0 and X_j^0 are transformed into sequences with the corresponding intervals the same. Now, as in Case 1, the method of mean generations can be applied to transform X_i^0 and X_j^0 into equal-time-interval sequences. As for the argument of an unchanging ε_{ij}, it can be given in a way similar to that in Case 1.

As for the situation of multiple blanks between two neighboring entries in the sequences, we can fill in the blanks one after another applying mean generations. All details are omitted here. □

Example 5.5.1. Assume that sequences

$$X_0 \quad = (x_0(1), x_0(2), x_0(3), x_0(4), x_0(5), x_0(7)) = (10, 9, 15, 14, 14, 16)$$

and

$$X_1 = (x_1(1), x_1(3), x_1(7)) = (46, 70, 98)$$

are given. Find the absolute degree ε_{01} of incidence of the sequences X_0 and X_1.

Solution. 1. Transform X_1 into a sequence with the same corresponding time intervals as X_0. Let

$$x_1(5) = \frac{1}{2}[x_1(3) + x_1(7)] = \frac{1}{2}[70 + 98] = 84,$$

$$x_1(2) = \frac{1}{2}[x_1(1) + x_1(3)] = \frac{1}{2}[46 + 70] = 58,$$

and

$$x_1(4) = \frac{1}{2}[x_1(3) + x_1(5)] = \frac{1}{2}[70 + 84] = 77.$$

So, we have a new sequence X_1 in the place of the original X_i:

$$\begin{aligned} X_1 &= (x_1(1), x_1(2), x_1(3), x_1(4), x_1(5), x_1(7)) \\ &= (46, 58, 70, 77, 84, 98). \end{aligned}$$

2. Transform X_0 and X_1 into equal-time-interval sequences. Let

$$x_0(6) = \frac{1}{2}[x_0(5) + x_0(7)] = \frac{1}{2}[14 + 16] = 15$$

and

$$x_1(6) = \frac{1}{2}[x_1(5) + x_1(7)] = \frac{1}{2}[84 + 98] = 91.$$

Then the new sequences X_0 and X_1 look as follows.

$$\begin{aligned} X_0 &= (x_0(1), x_0(2), x_0(3), x_0(4), x_0(5), x_0(6), x_0(7)) \\ &= (10, 9, 15, 14, 14, 15, 16) \end{aligned}$$

and

$$\begin{aligned} X_1 &= (x_1(1), x_1(2), x_1(3), x_1(4), x_1(5), x_1(6), x_1(7)) \\ &= (46, 58, 70, 77, 84, 91, 98). \end{aligned}$$

And, they are all 1-time-interval sequences.

3. Computing zero starting point images of X_0 and X_1 gives that

$$\begin{aligned} X_0^0 &= \left(x_0^0(1), x_0^0(2), x_0^0(3), x_0^0(4), x_0^0(5), x_0^0(6), x_0^0(7)\right) \\ &= (0, -1, 5, 4, 4, 5, 6) \end{aligned}$$

and

$$\begin{aligned} X_1^0 &= \left(x_1^0(1), x_1^0(2), x_1^0(3), x_1^0(4), x_1^0(5), x_1^0(6), x_1^0(7)\right) \\ &= (0, 12, 24, 31, 38, 45, 52)\,. \end{aligned}$$

4. Find $|s_0|$, $|s_1|$, and $|s_1 - s_0|$.

$$\begin{aligned} |s_0| &= \left|\sum_{k=2}^{6} x_0^0(k) + \tfrac{1}{2}x_0^0(7)\right| \\ &= |(-1) + 5 + 4 + 4 + 5 + \tfrac{1}{2}\cdot 6| = 20, \\ |s_1| &= \left|\sum_{k=2}^{6} x_1^0(k) + \tfrac{1}{2}x_1^0(7)\right| \\ &= |12 + 24 + 31 + 38 + 45 + \tfrac{1}{2}\cdot 52| = 176, \end{aligned}$$

and

$$\begin{aligned} |s_1 - s_0| &= \left|\sum_{k=2}^{6} [x_1^0(k) - x_0^0(k)] + \tfrac{1}{2}[[x_1^0(7) - x_0^0(7)]\right| \\ &= |13 + 19 + 27 + 34 + 40 + \tfrac{1}{2}\cdot 46| = 156. \end{aligned}$$

5. Compute the absolute degree of grey incidences.

$$\varepsilon_{01} = \frac{1 + |s_0| + |s_1|}{1 + |s_0| + |s_1| + |s_1 - s_0|} = \frac{197}{353} \approx 0.5581.$$

Theorem 5.5.4. *The absolute degree ε_{ij} of grey incidences satisfies the following conditions.*

1. $0 < \varepsilon_{ij} \leq 1$;

2. ε_{ij} is only related to the geometric shapes of X_i and X_j, and has nothing to do with the spatial positions of X_i and X_j. Or in other words, moving horizontally does not change the value of the absolute degree of grey incidences;

3. Any two sequences are not absolutely unrelated. That is, ε_{ij} never equals zero;

4. The more X_i and X_j are geometrically similar, the greater ε_{ij};

5. When X_i and X_j are parallel, or X_j^0 is vibrating around X_i^0 with the area of the parts with X_j^0 on top of X_i^0 being equal to that of the parts with X_j^0 beneath X_i^0, $\varepsilon_{ij} = 1$;

6. When any one of the data values in X_i or X_j changes, ε_{ij} also changes accordingly;

7. When the lengths of X_i and X_j change, ε_{ij} also changes accordingly;

8. $\varepsilon_{ii} = 1, \varepsilon_{jj} = 1$; and

9. $\varepsilon_{ij} = \varepsilon_{ji}$.

5.6 Relative Degree of Grey Incidence

When the idea of rate of change is considered, we have the following:

Definition 5.6.1. *Assume that X_i and X_j are two sequences of the same length with the initial values being zero, and $X_i^{'}$ and $X_j^{'}$ the initial images of X_i and X_j, respectively. Then, the absolute degree of grey incidence of $X_i^{'}$ and $X_j^{'}$ is called the relative degree of grey incidence, or relative degree of incidence for short, of X_i and X_j, denoted r_{ij}.*

The concept of relative degree of grey incidences of sequences X_i and X_j is a quantitative representation of the rates of change of X_i and X_j relative to their starting points. The closer the rates of change of X_i and X_j are, the greater r_{ij} is, and vice versa.

Proposition 5.6.1. *Assume that X_i and X_j are two sequences of the same length with non-zero initial values. If $X_i = cX_j$, where $c > 0$ is a constant, then $r_{ij} = 1$.*

Proof. Assume that

$$X_j = (x_j(1), x_j(2), ..., x_j(n)).$$

Then

$$X_i \quad = (x_i(1), x_i(2), ..., x_i(n)) = (cx_j(1), cx_j(2), ..., cx_j(n)).$$

The initial images of X_i and X_j are, respectively,

$$X_j^{'} = X_j/x_j(1) = \left(\frac{x_j(1)}{x_j(1)}, \frac{x_j(2)}{x_j(1)}, ..., \frac{x_j(n)}{x_j(1)}\right)$$

and

$$\begin{aligned} X_i^{'} = X_i/x_i(1) \quad &= \left(\frac{x_i(1)}{x_i(1)}, \frac{x_i(2)}{x_i(1)}, ..., \frac{x_i(n)}{x_i(1)}\right) \\ &= \left(\frac{cx_j(1)}{cx_j(1)}, \frac{cx_j(2)}{cx_j(1)}, ..., \frac{cx_j(n)}{cx_j(1)}\right) \\ &= \left(\frac{x_j(1)}{x_j(1)}, \frac{x_j(2)}{x_j(1)}, ..., \frac{x_j(n)}{x_j(1)}\right). \end{aligned}$$

So, $X_i^{'} = X_j^{'}$, which implies that their absolute degree of incidence is 1. Therefore, the relative degree of incidence of X_i and X_j is $r_{ij} = 1$. □

Proposition 5.6.2. *Assume that X_i and X_j are two sequences of the same length with non-zero initial values. Then, the relative degree r_{ij} of incidence and the absolute degree ε_{ij} of incidence of X_i and X_j do not have to have any connections. When ε_{ij} is relatively large, r_{ij} can be very small. When ε_{ij} is very small, r_{ij} can also be relatively large.*

Proof. Assume that

$$X_i = (x_i(1), x_i(2), ..., x_i(n))$$

and

$$X_j = (x_j(1), x_j(2), ..., x_j(n)).$$

If $\varepsilon_{ij} \approx 1$, then $X_i \approx X_j + c$, for some fixed constant c. That is,

$$X_i \approx (x_j(1) + c, x_j(2) + c, ..., x_j(n) + c).$$

The initial images of X_i and X_j are, respectively,

$$X_i^{'} \approx \left(\frac{x_j(1)+c}{x_j(1)+c}, \frac{x_j(2)+c}{x_j(1)+c}, ..., \frac{x_j(n)+c}{x_j(1)+c}\right)$$

and

$$X_j^{'} = \left(\frac{x_j(1)}{x_j(1)}, \frac{x_j(2)}{x_j(1)}, ..., \frac{x_j(n)}{x_j(1)}\right).$$

The value $\left|s_i^{'} - s_j^{'}\right|$ relative to $X_i^{'}$ and $X_j^{'}$ is

$$\begin{aligned} |s_i^{'} - s_j^{'}| &\approx \left|\sum_{k=2}^{n-1}\left[\frac{x_j(k)}{x_j(1)} - \frac{x_j(k)+c}{x_j(1)+c}\right]\right. \\ &\quad \left. + \frac{1}{2}\left[\frac{x_j(n)}{x_j(1)} - \frac{x_j(n)+c}{x_j(1)+c}\right]\right| \\ &= \left|\sum_{k=2}^{n-1}\frac{c\cdot[x_j(k)-x_j(1)]}{x_j(1)[x_j(1)+c]} + \frac{1}{2}\frac{c\cdot[x_j(n)-x_j(1)]}{x_j(1)[x_j(1)+c]}\right|. \end{aligned}$$

So, it is obvious that as long as the absolute values of c and

$$x_j(k) - x_j(1),$$

$k = 2, 3, ..., n$, are sufficiently large, $\left|s_i^{'} - s_j^{'}\right|$ will be sufficiently large. Therefore, the absolute degree r_{ij} of incidence of $X_i^{'}$ and $X_j^{'}$, that is the relative degree of incidence of X_i and X_j, can be sufficiently small.

If $r_{ij} \approx 1$, then $X_i \approx c \cdot X_j$ for some fixed constant c. That is,

$$X_i \approx (c \cdot x_j(1), c \cdot x_j(2), ..., c \cdot x_j(n)),$$

and the value $|s_i - s_j|$ relative to X_i and X_j is

$$\begin{aligned} |s_i - s_j| &\approx \left|\sum_{k=2}^{n-1}\{[x_j(k)-x_j(1)] - c\cdot[x_j(k)-x_j(1)]\}\right. \\ &\quad \left. +\tfrac{1}{2}\{[x_j(n)-x_j(1)] - c\cdot[x_j(n)-x_j(1)]\}\right| \\ &= |1-c| \cdot \left|\sum_{k=2}^{n-1}[x_j(k)-x_j(1)] + \tfrac{1}{2}[x_j(n)-x_j(1)]\right|. \end{aligned}$$

As long as the absolute values of c and

$$x_j(k) - x_j(1),$$

$k = 2, 3, ..., n$, are sufficiently large, $|s_i - s_j|$ will be sufficiently large. Therefore, the absolute degree ε_{ij} of X_i and X_j can be sufficiently small. □

Example 5.6.1. Compute the relative degree of incidence for X_0 and X_1 in Example 5.5.1.

Solution. 1. Transformation into same corresponding time intervals. Let

$$x_1(5) = \frac{1}{2}[x_1(3) + x_1(7)] = \frac{1}{2}[70 + 98] = 84,$$

$$x_1(2) = \frac{1}{2}[x_1(1) + x_1(3)] = \frac{1}{2}[46 + 70] = 58,$$

and

$$x_1(4) = \frac{1}{2}[x_1(3) + x_1(5)] = \frac{1}{2}[70 + 84] = 77.$$

So, we have

$$\begin{aligned} X_1 &= (x_1(1), x_1(2), x_1(3), x_1(4), x_1(5), x_1(7)) \\ &= (46, 58, 70, 77, 84, 98). \end{aligned}$$

2. Transformation into equal-time-intervals. Let

$$x_0(6) = \frac{1}{2}[x_0(5) + x_0(7)] = \frac{1}{2}[14 + 16] = 15$$

and

$$x_1(6) = \frac{1}{2}[x_1(5) + x_1(7)] = \frac{1}{2}[84 + 98] = 91.$$

Then we have the sequences

$$\begin{aligned} X_0 &= (x_0(1), x_0(2), x_0(3), x_0(4), x_0(5), x_0(6), x_0(7)) \\ &= (10, 9, 15, 14, 14, 15, 16) \end{aligned}$$

and

$$\begin{aligned} X_1 &= (x_1(1), x_1(2), x_1(3), x_1(4), x_1(5), x_1(6), x_1(7)) \\ &= (46, 58, 70, 77, 84, 91, 98). \end{aligned}$$

are all 1-time-interval sequences.

3. Compute the initial images of X_0 and X_1.

$$\begin{aligned} X_0' &= \left(\frac{x_0(1)}{x_0(1)}, \frac{x_0(2)}{x_0(1)}, \frac{x_0(3)}{x_0(1)}, \frac{x_0(4)}{x_0(1)}, \frac{x_0(5)}{x_0(1)}, \frac{x_0(6)}{x_0(1)}, \frac{x_0(7)}{x_0(1)}\right) \\ &= (1, 0.9, 1.5, 1.4, 1.4, 1.5, 1.6), \end{aligned}$$

and

$$X_1' = \left(\frac{x_1(1)}{x_1(1)}, \frac{x_1(2)}{x_1(1)}, \frac{x_1(3)}{x_1(1)}, \frac{x_1(4)}{x_1(1)}, \frac{x_1(5)}{x_1(1)}, \frac{x_1(6)}{x_1(1)}, \frac{x_1(7)}{x_1(1)}\right)$$
$$= (1, 1.26, 1.52, 1.67, 1.83, 1.98, 2.13).$$

4. Compute the images of zero starting points of X_0' and X_1'.

$$X_0'^0 = \left(x_0'^0(1), x_0'^0(2), x_0'^0(3), x_0'^0(4), x_0'^0(5), x_0'^0(6), x_0'^0(7)\right)$$
$$= (0, -0.1, 0.5, 0.4, 0.4, 0.5, 0.6),$$

and

$$X_1'^0 = \left(x_1'^0(1), x_1'^0(2), x_1'^0(3), x_1'^0(4), x_1'^0(5), x_1'^0(6), x_1'^0(7)\right)$$
$$= (0, 0.26, 0.52, 0.67, 0.83, 0.98, 1.13).$$

5. Compute $\left|s_0'\right|$, $\left|s_1'\right|$ and $\left|s_1' - s_0'\right|$.

$$|s_0'| = \left|\sum_{k=2}^{6} x_0'^0(k) + \tfrac{1}{2} x_0'^0(7)\right|$$
$$= \left|(-0.1) + 0.5 + 0.4 + 0.4 + 0.5 + \tfrac{1}{2} \cdot 0.6\right| = 2,$$

$$|s_1'| = \left|\sum_{k=2}^{6} x_1'^0(k) + \tfrac{1}{2} x_1'^0(7)\right|$$
$$= \left|0.26 + 0.52 + 0.67 + 0.83 + 0.98 + \tfrac{1}{2} \cdot 1.13\right| = 3.825,$$

and

$$|s_1' - s_0'| = \left|\sum_{k=2}^{6} [x_1'^0(k) - x_0'^0(k)] + \tfrac{1}{2}[x_1'^0(7) - x_0'^0(7)]\right|$$
$$= \left|0.36 + 0.02 + 0.37 + 0.43 + 0.48 + \tfrac{1}{2} \cdot 0.53\right| = 1.925.$$

6. Compute the relative degree of incidence.

$$r_{01} = \frac{1 + |s_0'| + |s_1'|}{1 + |s_0'| + |s_1'| + |s_1' - s_0'|} = \frac{6.825}{8.75} = 0.78.$$

Proposition 5.6.3. *Assume that X_i and X_j are sequences of the same length with non-zero initial entries, a and b non-zero constants, and the relative degree of incidence of aX_i and bX_j is r_{ij}'. Then $r_{ij}' = r_{ij}$. Or in* *other words, scalar multiplication does not change the relative degree of incidence.*

In fact, the initial images of aX_i and bX_j are respectively equal to those of X_i and X_j. So, scalar multiplication does not act in any way under the function of initialing operators. Hence, $r'_{ij} = r_{ij}$.

Proposition 5.6.4. *Each relative degree r_{ij} of grey incidence satisfies the following properties.*

1. $0 < r_{ij} \leq 1$;

2. r_{ij} *is only related to the rate of change of the initial entries of* X_i *and* X_j, *and has nothing to do with the magnitudes of other entries. Or, scalar multiplication does not change the relative degree of incidences;*

3. There always exists some relationship between the rates of change of any two sequences. That is, r_{ij} *never equals zero;*

4. The closer the individual rates of change of X_i *and* X_j *with respect to their initial points, the greater* r_{ij} *is;*

5. When the rates of change of X_i *and* X_j *with respect to their initial points are the same, that is,* $X_i = aX_j$, *or when the images of zero initial points of the initial images of* X_i *and* X_j *satisfy* $X_j^{\prime 0}$ *waves around* $X_i^{\prime 0}$, *and the area of the parts with* $X_j^{\prime 0}$ *above* $X_i^{\prime 0}$ *equals that of the parts with* $X_j^{\prime 0}$ *underneath* $X_i^{\prime 0}$, $r_{ij} = 1$;

6. When an entry in X_i *or* X_j *is changed,* r_{ij} *will change accordingly;*

7. When the lengths of sequences change, r_{ij} *will also change;*

8. $r_{ii} = 1$, $r_{jj} = 1$; *and*

9. $r_{ij} = r_{ji}$.

5.7 Synthetic Degree of Grey Incidence

When the overall relationship of closeness between sequences is considered, we have the following.

Definition 5.7.1. *Assume that* X_i *and* X_j *are sequences of the same length with non-zero initial entries, that* ε_{ij} *and* r_{ij} *are the absolute degree and the relative degree of grey incidence of* X_i *and* X_j, *and that* $\theta \in [0, 1]$. *Then*

$$\rho_{ij} = \theta\varepsilon_{ij} + (1 - \theta)r_{ij}$$

is called the synthetic degree of (grey) incidence between X_i *and* X_j.

The concept of synthetic degree of grey incidence is a numerical index that well describes the overall relationship of closeness between sequences. For example, it reflects the similarity between the zigzagged lines X_i and X_j, and also depicts the degree of closeness of the individual rates of change of X_i and X_j with respect to their initial points. In general, we can take $\theta = 0.5$. If we are more interested in the relationship between some absolute quantities, some greater value can be used as θ. If we are putting more emphasis on rates of change, some smaller value can be employed for θ.

Example 5.7.1. Find the synthetic degree of grey incidence of X_0 and X_1 in Example 5.5.1.

Solution: From Examples 5.5.1 and 5.6.1, it follows that $\varepsilon_{01} = 0.5581$ and $r_{01} = 0.78$. Take $\theta = 0.5$; we have

$$\begin{aligned} \rho_{01} &= \theta\varepsilon_{01} + (1-\theta)r_{01} \\ &= 0.5 \cdot 0.5581 + 0.5 \cdot 0.78 \approx 0.669. \end{aligned}$$

Proposition 5.7.1. *The synthetic degree ρ_{ij} of grey incidence satisfies the following properties.*

1. $0 < \rho_{ij} \le 1$;

2. ρ_{ij} is related to not only each observation value in the sequences X_i and X_j, and also the rate of change of each data value with respect to its initial point;

3. ρ_{ij} never equals zero;

4. If an entry value in X_i or X_j is changed, ρ_{ij} also changes accordingly;

5. If the length of X_i or X_j changes, ρ_{ij} also changes accordingly;

6. For different θ value, ρ_{ij} is also different;

7. When $\theta = 1$, then $\rho_{ij} = \varepsilon_{ij}$; and when $\theta = 0$, then $\rho_{ij} = r_{ij}$;

8. $\rho_{ii} = 1, \rho_{jj} = 1$; and

9. $\rho_{ij} = \rho_{ji}$.

5.8 Order of Grey Incidences

All the different degrees of grey incidence, which we have discussed earlier, are numerical characteristics for the relationship of closeness between two sequences. For a chosen operator of grey incidences (see Section 5.,2 for more details), the values of the degree of grey incidence, the absolute degree of grey incidence and the relative degree of incidence are all unique. When an operator of grey incidence and a θ value are all chosen, the synthetic degree of grey incidence is also unique. This kind of conditional uniqueness does not affect our analysis of problems of interest. When analyzing systems, and studying relationships between systems' characteristic behaviors and relevant factors' behaviors, we are mainly interested in the ordering of the degrees of incidence between the systems' characteristic behaviors and each relevant factor's behavioral sequence. So, the importance of magnitudes of the degrees of incidence is relative.

Definition 5.8.1. *Assume that X is an arbitrary set, and that "$\geq$" is a binary relation defined on the set X. If "$\geq$" satisfies reflexivity, antisymmetry, and transitivity, that is,*

1. Reflexivity: For any $X_i \in X, X_i \geq X_i$;

2. Antisymmetry: For any $X_i, X_j \in X$, if $X_i \geq X_j$, and $X_j \geq X_i$, then $X_i = X_j$;

3. Transitivity: For any $X_i, X_j, X_k \in X$, if $X_i \geq X_j$ and $X_j \geq X_k$, then $X_i \geq X_k$,

then "$\geq$" is called a partial order (relation) on the set X. A set X with a partial ordering "$\geq$" is called a partially ordered set and is denoted as $(X, \geq)$.

Definition 5.8.2. *Assume that $(X, \geq)$ is a partially ordered set. If the ordering relation "$\geq$" satisfies the following, for any $X_i, X_j \in X$, one of $X_i \geq X_j$ or $X_j \geq X_i$ must be true, then "$\geq$" is called a linear order(ing) on the set X, and the set with a linear order relation $\geq$ is called an ordered set.*

Definition 5.8.3. *Assume that X_0 is a sequence of a system's characteristic behaviors, that X_i and X_j are sequences of two relevant factors' behaviors, and that γ is the degree of grey incidence. If $\gamma_{0i} \geq \gamma_{0j}$, then the factor X_i is said to be more favorable than the factor X_j, denoted as $X_i \succ X_j$. The relation "$\succ$" is called the grey incidence order or the order of grey incidence, induced by the degree of grey incidence. Accordingly, the orders of incidence, induced by generalized degrees of grey incidence, are called generalized orders of grey incidence, where the generalized orders include absolute order of grey incidence, relative order of grey incidence, and synthetic order of grey incidence.*

Theorem 5.8.1. *Assume that X_0 is a sequence for a system's characteristic behaviors, and $X_1, X_2, ..., X_m$ sequences of relevant factors's behaviors. Let*

$$X = \{X_1, X_2, ..., X_m\}.$$

Then the order of grey incidences, absolute order of grey incidence, relative order of grey incidence, and the synthetic order of grey incidence are all partial orderings on the set X.

Proof. It suffices for us to show the case of absolute order of grey incidence. All other cases can be shown similarly.

1. Reflexivity. For any $X_i \in X$, if X_i and X_0 have the same length, then the absolute degree ε_{0i} of grey incidence is well defined. From

$$\varepsilon_{0i} = \varepsilon_{0i}$$

it follows that $X_i \succ X_i$.

2. Antisymmetry. Assume that $X_i \succ X_j$ and $X_j \succ X_i$. Then $\varepsilon_{0i} \geq \varepsilon_{0j}$ and $\varepsilon_{0j} \geq \varepsilon_{0i}$. So, $\varepsilon_{0i} = \varepsilon_{0j}$. That is, $X_i = X_j$.

3. Transitivity. Assume that $X_i \succ X_j$ and $X_j \succ X_k$, then $\varepsilon_{0i} \geq \varepsilon_{0j}$ and $\varepsilon_{0j} \geq \varepsilon_{0k}$. So, $\varepsilon_{0i} \geq \varepsilon_{0k}$. That is, $X_i \succ X_k$. □

Theorem 5.8.2. *Assume that X_0 is a sequence of a system's characteristic behaviors, and $X_1, X_2, ..., X_m$ behavioral sequences of some factors with the same length as X_0. Let $X = \{X_1, X_2, ..., X_m\}$. Then,*

1. The order of grey incidence and absolute order of grey incidence are linear orders on the set X;

2. If the initial entries of $X_0, X_1, X_2, ..., X_m$ are non-zero, then the relative order of grey incidence and the synthetic order of grey incidence are also linear orders on the set X.

Proof. 1. It is obvious to see that the order of grey incidence is linear. Because $X_1, X_2, ..., X_m$ all have the same length as X_0, for any $X_i, X_j \in X$, ε_{0i} and ε_{0j} are well defined, and one of the following inequalities $\varepsilon_{0i} \geq \varepsilon_{0j}$ and $\varepsilon_{0j} \geq \varepsilon_{0i}$ must hold true, one of $X_i \succ X_j$ and $X_j \succ X_i$ must also hold true. So, the absolute order of incidence is a linear order.

2. If the initial entries of $X_0, X_1, X_2, ..., X_m$ are non-zero, then for any $X_i, X_j \in X$, r_{0i}, r_{0j} and ρ_{0i}, ρ_{0j} are well-defined. So, under either the relative order of grey incidence or the synthetic order of grey incidence, we can obtain either $X_i \succ X_j$ or $X_j \succ X_i$. □

5.9 Preference Analysis

The idea of matrices is needed, when two sets of data sequences are involved.

Definition 5.9.1. *Assume that $Y_1, Y_2, ..., Y_s$ are sequences of a system's characteristic behaviors, and $X_1, X_2, ..., X_m$ are behavioral sequences of relevant factors. If the sequences $Y_1, Y_2, ..., Y_s$; X_1, X_2,..., X_m have the same length, γ_{ij}, $i = 1, 2, ..., s; j = 1, 2, ..., m$, is the degree of grey incidence of Y_i and X_j, then*

$$\Gamma = \left[\gamma_{ij}\right] = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{s1} & \gamma_{s2} & \cdots & \gamma_{sm} \end{bmatrix}$$

is called the matrix of grey incidences.

In the matrix of grey incidences, the entries in the ith row, $i = 1, 2, ..., s$, are the degrees of grey incidence of the sequence Y_i of the system's characteristic behaviors and the sequences X_1, X_2, ..., X_m of the relevant factors. And the entries in the jth column, $j = 1, 2, ..., m$, are the degrees of the sequences Y_1, Y_2, ..., Y_s of the system's characteristic behaviors and X_j.

Similarly, we can introduce matrices for various generalized incidences. For example, the absolute matrix of grey incidences

$$A = [\varepsilon_{ij}]_{s\times m} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \varepsilon_{s1} & \varepsilon_{s2} & \cdots & \varepsilon_{sm} \end{bmatrix}_{s\times m},$$

the relative matrix of grey incidences

$$B = [r_{ij}]_{s\times m} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{s1} & r_{s2} & \cdots & r_{sm} \end{bmatrix}_{s\times m},$$

and the synthetic matrix of grey incidence

$$C = [\rho_{ij}]_{s\times m} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{s1} & \rho_{s2} & \cdots & \rho_{sm} \end{bmatrix}_{s\times m}$$

are well defined. By making use of the various matrices of grey incidences, we can conduct preference analysis for a system's behaviors or relevant factors.

Definition 5.9.2. *Let Y_i, i= 1, 2, ..., s, and X_j, j = 1, 2, ..., m be the same as in Definition 5.9.1 and*

$$\Gamma = [\gamma_{ij}]_{s\times m} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{s1} & \gamma_{s2} & \cdots & \gamma_{sm} \end{bmatrix}_{s\times m}$$

the matrix of grey incidence. If there exist k and $i \in \{1, 2, ..., s\}$ satisfying

$$\gamma_{kj} \geq \gamma_{ij},$$

for $j = 1, 2, ..., m$, then we say that the system's characteristic Y_k is more favorable than the system's characteristic Y_i, denoted as $Y_k \succ Y_i$. If for any i = 1, 2, ..., s, with $i \neq k$, we always have $Y_k \succ Y_i$, then Y_k is said to be the most favorable characteristic.

Definition 5.9.3. *Let Y_i, i = 1, 2, ..., s, X_j, j = 1, 2, ..., m, and*

$$\Gamma = [\gamma_{ij}]_{s\times m} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{s1} & \gamma_{s2} & \cdots & \gamma_{sm} \end{bmatrix}_{s\times m}$$

be the same as in Definition 5.9.1. If there exist ℓ and $j \in \{1, 2, ..., m\}$ satisfying

$$\gamma_{i\ell} \geq \gamma_{ij},$$

for $i = 1, 2, ..., s$, then the factor X_ℓ is said to be more favorable than the factor X_j, denoted as $X_\ell \succ X_j$. If for any j = 1, 2, ..., m, $j \neq \ell$, we always have $X_\ell \succ X_j$, then X_ℓ is called the most favorable factor.

Definition 5.9.4. *Assume that*

$$\Gamma = \left[\gamma_{ij}\right] = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{s1} & \gamma_{s2} & \cdots & \gamma_{sm} \end{bmatrix}$$

is the matrix of grey incidence.

1. If there exist k, $i \in \{1, 2, ..., s\}$, satisfying

$$\sum_{j=1}^{m} \gamma_{kj} \geq \sum_{j=1}^{m} \gamma_{ij},$$

then the system's characteristic Y_k is said to be more quasi-favorable than the characteristic Y_i, denoted $Y_k \succcurlyeq Y_i$.

2. If there exist ℓ, $j \in \{1, 2, ..., m\}$ satisfying

$$\sum_{i=1}^{s} \gamma_{i\ell} \geq \sum_{i=1}^{s} \gamma_{ij},$$

then the factor X_ℓ is said to be more quasi-favorable than the factor X_j, denoted $X_\ell \succcurlyeq X_j$.

Definition 5.9.5. *1. If there exists a $k \in \{1, 2, ..., s\}$ such that for any i = 1, 2, ..., s, $Y_k \succcurlyeq Y_i$, then Y_k is called a quasi-preferred characteristic of the system.*

2. If there exists an $\ell \in \{1, 2, ..., m\}$ such that for any j = 1, 2, ..., m, $X_\ell \succcurlyeq X_j$, then the factor X_ℓ is said to be a quasi-preferred factor.

Proposition 5.9.1. *In a system with s characteristics and m relevant factors, there may not exist the most favorable characteristic and the most favorable factor, but there must be quasi-preferred characteristics and factors.*

Example 5.9.1. Assume that

$$Y_1 = (170, 174, 197, 216.4, 235.8),$$

$$Y_2 = (57.55, 70.74, 76.8, 80.7, 89.85),$$

and

$$Y_3 = (68.56, 70, 85.38, 99.83, 103.4)$$

are sequences of a system's characteristic behaviors,

$$\begin{aligned} X_1 &= (308.58, 310, 295, 346, 367)\,, \\ X_2 &= (195.4, 189.9, 187.2, 205, 222.7)\,, \\ X_3 &= (24.6, 21, 12.2, 15.1, 14.57)\,, \\ X_4 &= (20, 25.6, 23.3, 29.2, 30)\,, \end{aligned}$$

and

$$X_5 = (18.98, 19, 22.3, 23.5, 27.655)$$

behavioral sequences of relevant factors. Do a preference analysis for the system given.

Solution. 1. Compute the absolute matrix of incidence. For each behavioral sequence, we compute its image of zeroing starting points,

$$\begin{aligned} Y_1^0 &= (0, 4, 27, 46.4, 65.8)\,, \\ Y_2^0 &= (0, 13.19, 19.25, 23.15, 32.3)\,, \\ Y_3^0 &= (0, 1.44, 16.82, 31.27, 34.84)\,; \end{aligned}$$

$$\begin{aligned} X_1^0 &= (0, 1.42, -13.58, 37.42, 58.42)\,, \\ X_2^0 &= (0, -5.5, -8.2, 9.6, 27.3)\,, \\ X_3^0 &= (0, -3.6, -12.4, -9.5, -10.03)\,, \\ X_4^0 &= (0, 5.6, 3.3, 9.2, 10)\,, \end{aligned}$$

and

$$X_5^0 = (0, 0.02, 3.32, 4.52, 8.675)\,.$$

Corresponding to the system's characteristic Y_1, we have

$$\begin{aligned} |Y_{s_1}| &= \left| \sum_{k=2}^{4} y_1^0(k) + \tfrac{1}{2} y_1^0(5) \right| \\ &= \left| 4 + 27 + 46.4 + \tfrac{1}{2} \cdot 65.8 \right| = 110.3, \end{aligned}$$

$$\begin{aligned} |X_{s_1}| &= \left| \sum_{k=2}^{4} x_1^0(k) + \tfrac{1}{2} x_1^0(5) \right| \\ &= \left| 1.42 + (-13.58) + 37.42 + \tfrac{1}{2} \cdot 58.42 \right| = 54.47, \end{aligned}$$

$$\begin{aligned}|X_{s_1}-Y_{s_1}| &= \left|\sum_{k=2}^{4}[x_1^0(k)-y_1^0(k)]+\tfrac{1}{2}[x_1^0(5)-y_1^0(5)]\right| \\ &= |[1.42-4]+[-13.58-27]+[37.42-46.4] \\ &\quad +\frac{1}{2}\cdot[58.42-65.8]| = 55.9,\end{aligned}$$

$$\begin{aligned}\varepsilon_{11} &= \frac{1+|Y_{s_1}|+|X_{s_1}|}{1+|Y_{s_1}|+|X_{s_1}|+|X_{s_1}-Y_{s_1}|} \\ &= \frac{1+110.3+54.47}{1+110.3+54.47+55.9}=\frac{165.77}{221.67}=0.748;\end{aligned}$$

$$\begin{aligned}|X_{s_2}| &= \left|\sum_{k=2}^{4}x_2^0(k)+\tfrac{1}{2}x_2^0(5)\right| \\ &= \left|(-5.5)+(-8.2)+9.6+\tfrac{1}{2}\cdot 27.3\right| = 9.55,\end{aligned}$$

$$\begin{aligned}|X_{s_2}-Y_{s_1}| &= \left|\sum_{k=2}^{4}[x_2^0(k)-y_1^0(k)]+\tfrac{1}{2}[x_2^0(5)-y_1^0(5)]\right| \\ &= |[-5.5-4]+[8.2-27]+[9.6-46.4] \\ &\quad +\frac{1}{2}[27.3-65.8]| = 100.75,\end{aligned}$$

$$\begin{aligned}\varepsilon_{12} &= \frac{1+|Y_{s_1}|+|X_{s_2}|}{1+|Y_{s_1}|+|X_{s_2}|+|X_{s_2}-Y_{s_1}|} \\ &= \frac{1+110.3+9.55}{1+110.3+9.55+100.75}=\frac{120.85}{221.6}=0.545;\end{aligned}$$

$$\begin{aligned}|X_{s_3}| &= \left|\sum_{k=2}^{4}x_3^0(k)+\tfrac{1}{2}x_3^0(5)\right| \\ &= |-3.6-12.4-9.5-\tfrac{1}{2}\cdot 10.03| = 30.515,\end{aligned}$$

$$\begin{aligned}|X_{s_3}-Y_{s_1}| &= \left|\sum_{k=2}^{4}[x_3^0(k)-y_1^0(k)]+\tfrac{1}{2}[x_3^0(5)-y_1^0(5)]\right| \\ &= |[-3.6-4]+[-12.4-27]+[-9.5-46.4] \\ &\quad +\frac{1}{2}[-10.03-65.8]| = 140.815,\end{aligned}$$

$$\begin{aligned}\varepsilon_{13} &= \frac{1+|Y_{s_1}|+|X_{s_3}|}{1+|Y_{s_1}|+|X_{s_3}|+|X_{s_3}-Y_{s_1}|} \\ &= \frac{1+110.3+30.515}{1+110.3+30.515+140.815} = \frac{141.815}{282.63} = 0.502;\end{aligned}$$

$$\begin{aligned}|X_{s_4}| &= \left|\sum_{k=2}^{4} x_4^0(k) + \tfrac{1}{2}x_4^0(5)\right| \\ &= |5.6+3.3+9.2+\tfrac{1}{2}\cdot 10| = 23.1,\end{aligned}$$

$$\begin{aligned}|X_{s_4}-Y_{s_1}| &= \left|\sum_{k=2}^{4}[x_4^0(k)-y_1^0(k)] + \tfrac{1}{2}[x_4^0(5)-y_1^0(5)]\right| \\ &= |[5.6-4]+[3.3-27]+[9.2-46.4] \\ &\quad +\frac{1}{2}[10-65.8]| = 87.2,\end{aligned}$$

$$\begin{aligned}\varepsilon_{14} &= \frac{1+|Y_{s_1}|+|X_{s_4}|}{1+|Y_{s_1}|+|X_{s_4}|+|X_{s_4}-Y_{s_1}|} \\ &= \frac{1+110.3+23.1}{1+110.3+23.1+87.2} = \frac{134.4}{221.6} = 0.606;\end{aligned}$$

$$\begin{aligned}|X_{s_5}| &= \left|\sum_{k=2}^{4} x_5^0(k) + \tfrac{1}{2}x_5^0(5)\right| \\ &= |0.02+3.32+4.52+\tfrac{1}{2}\cdot 8.675| = 12.1975,\end{aligned}$$

$$\begin{aligned}|X_{s_5}-Y_{s_1}| &= \left|\sum_{k=2}^{4}[x_5^0(k)-y_1^0(k)] + \tfrac{1}{2}[x_5^0(5)-y_1^0(5)]\right| \\ &= |[0.02-4]+[3.32-27]+[4.52-46.4] \\ &\quad +\frac{1}{2}\cdot[8.675-65.8]| = 98.1025,\end{aligned}$$

$$\begin{aligned}\varepsilon_{15} &= \frac{1+|Y_{s_1}|+|X_{s_5}|}{1+|Y_{s_1}|+|X_{s_5}|+|X_{s_5}-Y_{s_1}|} \\ &= \frac{1+110.3+12.1975}{1+110.3+12.1975+98.1025} = \frac{123.4975}{221.6} = 0.557.\end{aligned}$$

Corresponding to the system's characteristic Y_2, we have

$$\begin{aligned}
|Y_{s_2}| &= \left|\sum_{k=2}^{4} y_2^0(k) + \tfrac{1}{2}y_2^0(5)\right| \\
&= |13.19 + 19.25 + 23.15 + \tfrac{1}{2}\cdot 32.3| = 71.74,
\end{aligned}$$

$$\begin{aligned}
|X_{s_1} - Y_{s_2}| &= \left|\sum_{k=2}^{4} [x_1^0(k) - y_2^0(k)] + \tfrac{1}{2}[x_1^0(5) - y_2^0(5)]\right| \\
&= |[1.42 - 13.19] + [-13.58 - 19.25] + [37.42 - 23.15] \\
&\quad + \frac{1}{2}\cdot[58.42 - 32.3]| = 17.27,
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{21} &= \frac{1 + |Y_{s_2}| + |X_{s_1}|}{1 + |Y_{s_2}| + |X_{s_1}| + |X_{s_1} - Y_{s_2}|} \\
&= \frac{1 + 71.74 + 54.47}{1 + 71.74 + 54.47 + 17.27} = \frac{127.21}{144.48} = 0.88;
\end{aligned}$$

$$\begin{aligned}
|X_{s_2} - Y_{s_2}| &= \left|\sum_{k=2}^{4} [x_2^0(k) - y_2^0(k)] + \tfrac{1}{2}[x_2^0(5) - y_2^0(5)]\right| \\
&= |[-5.5 - 13.19] + [-8.2 - 19.25] + [9.6 - 23.15] \\
&\quad + \frac{1}{2}\cdot[27.3 - 32.3]| = 62.19,
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{22} &= \frac{1 + |Y_{s_2}| + |X_{s_2}|}{1 + |Y_{s_2}| + |X_{s_2}| + |X_{s_2} - Y_{s_2}|} \\
&= \frac{1 + 71.74 + 9.55}{1 + 71.74 + 9.55 + 62.19} = \frac{82.29}{144.48} = 0.57;
\end{aligned}$$

$$\begin{aligned}
|X_{s_3} - Y_{s_2}| &= \left|\sum_{k=2}^{4} [x_3^0(k) - y_2^0(k)] + \tfrac{1}{2}[x_3^0(5) - y_2^0(5)]\right| \\
&= |[-3.6 - 13.19] + [-12.4 - 19.25] \\
&\quad + [-9.5 - 23.15] + \frac{1}{2}\cdot[-10.03 - 32.3]| = 102.255,
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{23} &= \frac{1 + |Y_{s_2}| + |X_{s_3}|}{1 + |Y_{s_2}| + |X_{s_3}| + |X_{s_3} - Y_{s_2}|} \\
&= \frac{1 + 71.74 + 30.515}{1 + 71.74 + 30.515 + 102.255} = \frac{103.255}{205.51} = 0.502;
\end{aligned}$$

$$|X_{s_4} - Y_{s_2}| = \left|\sum_{k=2}^{4}[x_4^0(k) - y_2^0(k)] + \tfrac{1}{2}[x_4^0(5) - y_2^0(5)]\right|$$

$$= |[5.6 - 13.19] + [3.3 - 19.25] + [9.2 - 23.15]$$

$$+\tfrac{1}{2}\cdot[10 - 32.3]| = 48.64,$$

$$\varepsilon_{24} = \frac{1 + |Y_{s_2}| + |X_{s_4}|}{1 + |Y_{s_2}| + |X_{s_4}| + |X_{s_4} - Y_{s_2}|}$$

$$= \frac{1 + 71.74 + 23.1}{1 + 71.74 + 23.1 + 48.64} = \frac{95.84}{144.48} = 0.663;$$

$$|X_{s_5} - Y_{s_2}| = \left|\sum_{k=2}^{4}[x_5^0(k) - y_2^0(k)] + \tfrac{1}{2}[x_5^0(5) - y_2^0(5)]\right|$$

$$= |[0.02 - 13.19] + [3.32 - 19.25] + [4.52 - 23.15]$$

$$+\tfrac{1}{2}\cdot[8.675 - 32.3]| = 59.5425,$$

$$\varepsilon_{25} = \frac{1 + |Y_{s_2}| + |X_{s_5}|}{1 + |Y_{s_2}| + |X_{s_5}| + |X_{s_5} - Y_{s_2}|}$$

$$= \frac{1 + 71.74 + 12.1975}{1 + 71.74 + 12.1975 + 59.5425} = \frac{84.9375}{144.48} = 0.588;$$

Corresponding to the system's characteristic Y_3, we can obtain similarly

$$\varepsilon_{31} = 0.907, \varepsilon_{32} = 0.574, \varepsilon_{33} = 0.503,$$

$$\varepsilon_{34} = 0.675, \varepsilon_{35} = 0.594.$$

Therefore, the absolute matrix of incidences is

$$A = [\varepsilon_{ij}]_{3\times 5} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} & \varepsilon_{14} & \varepsilon_{15} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} & \varepsilon_{24} & \varepsilon_{25} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} & \varepsilon_{34} & \varepsilon_{35} \end{bmatrix}$$

$$= \begin{bmatrix} 0.748 & 0.545 & 0.502 & 0.606 & 0.557 \\ 0.880 & 0.570 & 0.502 & 0.663 & 0.588 \\ 0.907 & 0.574 & 0.503 & 0.675 & 0.594 \end{bmatrix}.$$

2. Compute the relative matrix of incidence.

The initial images of the system's characteristic behaviors Y_i, $i = 1, 2, 3$, and the behavioral sequences of relevant factors X_j, $j = 1, 2, 3, 4, 5$, are

$$Y_1^{'} = (1, 1.0235, 1.1588, 1.2729, 1.3871)\,,$$

$$Y_2^{'} = (1, 1.2292, 1.3345, 1.4023, 1.5613)\,,$$

$$Y_3^{'} = (1, 1.0210, 1.2398, 1.4561, 1.5082)\,;$$

$$X_1^{'} = (1, 1.0046, 0.9560, 1.1213, 1.1893)\,,$$

$$X_2^{'} = (1, 0.9719, 0.9580, 1.0491, 1.1397)\,,$$

$$X_3^{'} = (1, 0.8537, 0.4959, 0.6138, 0.5923)\,,$$

$$X_4^{'} = (1, 1.28, 1.165, 1.46, 1.5)\,,$$

and

$$X_5^{'} = (1, 1.0011, 1.1749, 1.2381, 1.4571)\,.$$

The images of zero starting points of all $Y_i^{'}$, $i = 1, 2, 3$, and $X_j^{'}$, $j = 1, 2, 3, 4, 5$, are

$$\begin{aligned} Y_1^{'0} &= \left(y_1^{'0}(1), y_1^{'0}(2), y_1^{'0}(3), y_1^{'0}(4), y_1^{'0}(5)\right) \\ &= (0, 0.0235, 0.1588, 0.2729, 0.3871)\,, \\ Y_2^{'0} &= \left(y_2^{'0}(1), y_2^{'0}(2), y_2^{'0}(3), y_2^{'0}(4), y_2^{'0}(5)\right) \\ &= (0, 0.2292, 0.3345, 0.4023, 0.5613)\,, \\ Y_3^{'0} &= \left(y_3^{'0}(1), y_3^{'0}(2), y_3^{'0}(3), y_3^{'0}(4), y_3^{'0}(5)\right) \\ &= (0, 0.0210, 0.2398, 0.4561, 0.5082)\,; \\ X_1^{'0} &= \left(x_1^{'0}(1), x_1^{'0}(2), x_1^{'0}(3), x_1^{'0}(4), x_1^{'0}(5)\right) \\ &= (0, 0.0046, -0.044, 0.1213, 0.1893)\,, \\ X_2^{'0} &= \left(x_2^{'0}(1), x_2^{'0}(2), x_2^{'0}(3), x_2^{'0}(4), x_2^{'0}(5)\right) \\ &= (0, -0.0281, -0.042, 0.0491, 0.1397)\,, \end{aligned}$$

$$\begin{aligned} X_3^{'0} &= \left(x_3^{'0}(1), x_3^{'0}(2), x_3^{'0}(3), x_3^{'0}(4), x_3^{'0}(5)\right) \\ &= (0, -0.1463, -0.5041, -0.3862, -0.4077)\,, \end{aligned}$$

$$X_4^{'0} = \left(x_4^{'0}(1), x_4^{'0}(2), x_4^{'0}(3), x_4^{'0}(4), x_4^{'0}(5)\right)$$
$$= (0, 0.28, 0.165, 0.46, 0.5),$$

and

$$X_5^{'0} = \left(x_5^{'0}(1), x_5^{'0}(2), x_5^{'0}(3), x_5^{'0}(4), x_5^{'0}(5)\right)$$
$$= (0, 0.0011, 0.1749, 0.2381, 0.4571).$$

From

$$|Y_{s_i}^{'}| = \left|\sum_{k=2}^{4} y_i^{'0}(k) + \frac{1}{2} \cdot y_i^{'0}(5)\right|, i = 1, 2, 3;$$

$$|X_{s_j}^{'}| = \left|\sum_{k=2}^{4} x_j^{'0}(k) + \frac{1}{2} \cdot x_j^{'0}(5)\right|, j = 1, 2, 3, 4, 5,$$

$$|X_{s_j}^{'} - Y_{s_i}^{'}| = \left|\sum_{k=2}^{4} [x_j^{'0}(k) - y_i^{'0}(k)] + \frac{1}{2} \cdot [x_j^{'0}(5) - y_i^{'0}(5)]\right|$$

$i = 1, 2, 3$; $j = 1, 2, 3, 4, 5$, and

$$r_{ij} = \frac{1 + |Y_{s_i}^{'}| + |X_{s_j}^{'}|}{1 + |Y_{s_i}^{'}| + |X_{s_j}^{'}| + |X_{s_j}^{'} - Y_{s_i}^{'}|},$$

$i = 1, 2, 3$; $j = 1, 2, 3, 4, 5$, it follows that

$$r_{11} = 0.7945, r_{12} = 0.7389, r_{13} = 0.6046,$$
$$r_{14} = 0.8471, r_{15} = 0.9973;$$

$$r_{21} = 0.6937, r_{22} = 0.6571, r_{23} = 0.5837,$$
$$r_{24} = 0.9738, r_{25} = 0.8271;$$

$$r_{31} = 0.7300, \quad r_{32} = 0.6866, \quad r_{33} = 0.6101,$$
$$r_{34} = 0.9444, \quad r_{35} = 0.8884.$$

So, the relative matrix of incidence is obtained as follows:

$$B = [r_{ij}]_{3\times 5} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \end{bmatrix}$$

$$= \begin{bmatrix} 0.7945 & 0.7389 & 0.6046 & 0.8471 & 0.9973 \\ 0.6937 & 0.6571 & 0.5837 & 0.9738 & 0.8271 \\ 0.7300 & 0.6866 & 0.6101 & 0.9444 & 0.8884 \end{bmatrix}.$$

3. Compute the synthetic matrix C of incidence. Take $\theta = 0.5$, then

$$C = \theta A + (1-\theta)B = [\theta\varepsilon_{ij} + (1-\theta)r_{ij}]$$

$$= [\rho_{ij}] = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} & \rho_{25} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} & \rho_{35} \end{bmatrix}$$

$$= \begin{bmatrix} 0.7713 & 0.6420 & 0.5533 & 0.7266 & 0.7772 \\ 0.7869 & 0.6136 & 0.5429 & 0.8184 & 0.7076 \\ 0.8185 & 0.6303 & 0.5566 & 0.8097 & 0.7412 \end{bmatrix}.$$

4. Analysis. From the absolute matrix of incidence, it can be seen that because the rows of A satisfy

$$\varepsilon_{3j} > \varepsilon_{2j} \geq \varepsilon_{1j},$$

$j = 1, 2, 3, 4, 5$, we have

$$Y_3 \succ Y_2 \succ Y_1.$$

That is, Y_3 is the most favorable characteristic, Y_2 the second favorable, and Y_1 the last. All columns of A satisfy

$$\varepsilon_{i1} > \varepsilon_{i4} > \varepsilon_{i5} > \varepsilon_{i2} > \varepsilon_{i3},$$

$i = 1, 2, 3$. So, we have

$$X_1 \succ X_4 \succ X_5 \succ X_2 \succ X_3 .$$

That is, X_1 is the most favorable factor, X_4 the second, X_5 the third, X_2 the fourth, and X_3 the last.

From the relative matrix B of incidence, it can be seen that the elements in B satisfy

$$r_{i4} > r_{i1} > r_{i2} > r_{i3},$$

$i = 1, 2, 3$, and

$$r_{i5} > r_{i1} > r_{i2} > r_{i3},$$

$i = 1, 2, 3$. So, we conclude that

$$X_4 \succ X_1 \succ X_2 \succ X_3$$

and

$$X_5 \succ X_1 \succ X_2 \succ X_3.$$

Hence, X_3 is the most unfavorable factor of the system. Now let us further consider

$$\sum_{j=1}^{5} r_{1j} = 3.9824 > \sum_{j=1}^{5} r_{3j} = 3.8595 > \sum_{j=1}^{5} r_{2j} = 3.7354$$

So, we can conclude that

$$Y_1 \succcurlyeq Y_3 \succcurlyeq Y_2.$$

That is, Y_1 is the quasi-preferred characteristic. Because

$$\sum_{i=1}^{3} r_{i4} = 2.7653 > \sum_{i=1}^{3} r_{i5} = 2.7128 > \sum_{i=1}^{3} r_{i1} = 2.2182$$

$$> \sum_{i=1}^{3} r_{i2} = 2.0826 > \sum_{i=1}^{3} r_{i3} = 1.7984,$$

we have that

$$X_4 \succcurlyeq X_5 \succcurlyeq X_1 \succcurlyeq X_2 \succcurlyeq X_3.$$

That is, X_4 is the quasi-preferred factor, X_5 the next, and X_3 the most unfavorable factor.

From the synthetic matrix C of incidences, it can be seen that entries of C satisfy

$$\rho_{i1} > \rho_{i2} > \rho_{i3}, \rho_{i4} > \rho_{i2} > \rho_{i3}$$

and

$$\rho_{i5} > \rho_{i2} > \rho_{i3},$$

$i = 1, 2, 3$. So, we have

$$X_1 \succ X_2 \succ X_3, X_4 \succ X_2 \succ X_3$$

and

$$X_5 \succ X_2 \succ X_3.$$

That is, X_3 is the least preferred. We now further consider

$$\sum_{j=1}^{5}\rho_{3j} = 3.5563 > \sum_{j=1}^{5}\rho_{1j} = 3.4704 > \sum_{j=1}^{5}\rho_{2j} = 3.4694,$$

So,

$$Y_3 \succcurlyeq Y_1 \succcurlyeq Y_2.$$

That is, Y_3 is the quasi-preferred characteristic.

From

$$\sum_{i=1}^{3}\rho_{i1} = 2.3767 > \sum_{i=1}^{3}\rho_{i4} = 2.3547 > \sum_{i=1}^{3}\rho_{i5} = 2.2260$$

$$> \sum_{i=1}^{3}\rho_{i2} = 1.8859 > \sum_{i=1}^{3}\rho_{i3} = 1.6528,$$

it follows that

$$X_1 \succcurlyeq X_4 \succcurlyeq X_5 \succcurlyeq X_2 \succcurlyeq X_3.$$

So, X_1 is the quasi-preferred factor, X_4 the next, X_5 is more favorable than X_2, and X_3 is the most unfavorable.

The reason why the conclusions of the three incidence analyses do not agree with each other is because the absolute order of incidences looks at relationships from the angle of absolute magnitudes, the relative order of incidences from the angle of rates of change at each moment of the observation data with respect to their initial points, and the synthetic order of incidences from the combined angle of both absolute magnitudes and rates of change. In practical applications, one of these orders can be chosen based on the specific information and circumstances given. For the sake of convenience, when the system's behavioral sequence and the relevant factors' sequences are worked on by a specified operator of grey incidences, it will be enough to consider the absolute order of incidences.

5.10 Practical Applications

This chapter ends with two of the many real-life projects in which we have been involved.

Example 5.10.1. In this project, we look at a grey incidence analysis of such an economy that consists of (non-governmental) enterprises owned individually and collectively, at Change County, Henan Province, The People's Republic of China.

In recent years, (non-governmental) enterprises owned individually and collectively at Change County have developed rapidly. From 1983 to 1986, for example, the average annual growth of these non-governmental enterprises was 51.6%. These enterprises occupied an important position in the overall picture of the region's economic development. In 1986, the revenue of these enterprises reached 35,388 (10,000 yuan), accounting for 60% of the total industrial and agricultural revenue of the county. So, it became a common concern of the county on how to effectively speed up the development of these non-governmental enterprises in order to help the region's economy to take off from its historical ground. Based on relevant analysis, it was known that these enterprises were mainly dominated by four factors: fixed capital, circulating capital, labor forces, and after-tax profits. The sequences of the production revenue and the relevant factors of this county's non-governmental enterprises are given in the following Table 5.1.

Table 5.1. Recorded values of variables for the years 1983 to 1986

	1983	1984	1985	1986
X_0 (production revenue)	10,155	12,588	23,408	35,388
X_1 (fixed capitals)	3,799	3,605	5,460	6,982
X_2 (circulating capitals)	1,752	2,160	2,213	4,753
X_3 (labor forces: person)	24,186	45,590	57,685	85,540
X_4 (after-tax profits)	1,164	1,788	3,134	4,478

with 10,000 yuan as the unit.

1. Compute the absolute degree of incidence. Let

$$\begin{aligned} X_i^0 &= (x_i(1) - x_i(1), x_i(2) - x_i(1), x_i(3) - x_i(1), x_i(4) - x_i(1)) \\ &= \left(x_i^0(1), x_i^0(2), x_i^0(3), x_i^0(4)\right), \end{aligned}$$

$i = 0, 1, 2, 3, 4$; then

$$X_0^0 = \left(x_0^0(1), x_0^0(2), x_0^0(3), x_0^0(4)\right) = (0, 2433, 13325, 25233),$$

$$X_1^0 = \left(x_1^0(1), x_1^0(2), x_1^0(3), x_1^0(4)\right) = (0, -194, 1661, 3183),$$

$$X_2^0 = \left(x_2^0(1), x_2^0(2), x_2^0(3), x_2^0(4)\right) = (0, 408, 461, 3001),$$

$$X_3^0 = \left(x_3^0(1), x_3^0(2), x_3^0(3), x_3^0(4)\right) = (0, 21404, 33499, 61354),$$

$$X_4^0 = \left(x_4^0(1), x_4^0(2), x_4^0(3), x_4^0(4)\right) = (0, 624, 2030, 3314).$$

From

$$|s_i| = \left| \sum_{k=2}^{3} x_i^0(k) + \frac{1}{2} x_i^0(4) \right|,$$

$i = 1, 2, 3, 4$, it follows that

$$|s_0| = \left|2433 + 13325 + \tfrac{1}{2} \cdot 25233\right| = 28374.5,$$

$$|s_1| = \left|-194 + 1661 + \tfrac{1}{2} \cdot 3183\right| = 3058.5,$$

$$|s_2| = \left|408 + 461 + \tfrac{1}{2} \cdot 3001\right| = 2369.5,$$

$$|s_3| = \left|21404 + 33499 + \tfrac{1}{2} \cdot 61354\right| = 85580,$$

and

$$|s_4| = \left|624 + 2030 + \tfrac{1}{2} \cdot 3314\right| = 4311.$$

From

$$|s_i - s_0| = \left|\sum_{k=2}^{3}[x_i^0(k) - x_0^0(k)] + \frac{1}{2}[x_i^0(4) - x_0^0(4)]\right|,$$

$i = 1, 2, 3, 4$, it follows that

$$|s_1 - s_0| = 25316, |s_2 - s_0| = 26005,$$

$$|s_3 - s_0| = 57205.5, |s_4 - s_0| = 24063.5.$$

From

$$\varepsilon_{0i} = \frac{1 + |s_0| + |s_i|}{1 + |s_0| + |s_i| + |s_i - s_0|},$$

$i = 1, 2, 3, 4$, it follows that

$$\varepsilon_{01} = 0.554,\ \varepsilon_{02} = 0.542,\ \varepsilon_{03} = 0.666,\ \varepsilon_{04} = 0.576.$$

2. Compute the relative degree of incidence. We first compute the initial images of X_i, $i = 0, 1, 2, 3, 4$. From

$$X_i' = \left(\frac{x_i(1)}{x_i(1)}, \frac{x_i(2)}{x_i(1)}, \frac{x_i(3)}{x_i(1)}, \frac{x_i(4)}{x_i(1)}\right)$$

$$= \left(x_i'(1), x_i'(2), x_i'(3), x_i'(4)\right),$$

$i = 1, 2, 3, 4$, it follows that

$$X_0' = \left(x_0'(1), x_0'(2), x_0'(3), x_0'(4)\right)$$

$$= (1, 1.2396, 2.3051, 3.4848),$$

$$X_1' = \left(x_1'(1), x_1'(2), x_1'(3), x_1'(4)\right) = (1, 0.9489, 1.4372, 1.8379),$$

$$X_2' = \left(x_2'(1), x_2'(2), x_2'(3), x_2'(4)\right) = (1, 1.2329, 1.2631, 2.7129),$$

$$X_3^{'} = \left(x_3^{'}(1), x_3^{'}(2), x_3^{'}(3), x_3^{'}(4)\right) = (1, 1.8850, 2.3851, 3.5368)\,,$$

and

$$X_4^{'} = \left(x_4^{'}(1), x_4^{'}(2), x_4^{'}(3), x_4^{'}(4)\right) = (1, 1.5361, 2.6924, 3.8471)\,.$$

The images of X_i, i =0, 1, 2, 3, 4, of zero starting points are given as follows.

$$X_i^{'0} = \left(x_i^{'}(1) - x_i^{'}(1), x_i^{'}(2) - x_i^{'}(1), x_i^{'}(3) - x_i^{'}(1), x_i^{'}(4) - x_i^{'}(1)\right)$$

$$= \left(x_i^{'0}(1), x_i^{'0}(2), x_i^{'0}(3), x_i^{'0}(4)\right),$$

$i = 1, 2, 3, 4$, and

$$X_0^{'0} = \left(x_0^{'0}(1), x_0^{'0}(2), x_0^{'0}(3), x_0^{'0}(4)\right) = (0, 0.2396, 1.3051, 2.4848)\,,$$

$$X_1^{'0} = \left(x_1^{'0}(1), x_1^{'0}(2), x_1^{'0}(3), x_1^{'0}(4)\right) = (0, -0.0511, 0.4372, 0.8379)\,,$$

$$X_2^{'0} = \left(x_2^{'0}(1), x_2^{'0}(2), x_2^{'0}(3), x_2^{'0}(4)\right) = (0, 0.2329, 0.2631, 1.7129)\,,$$

$$X_3^{'0} = \left(x_3^{'0}(1), x_3^{'0}(2), x_3^{'0}(3), x_3^{'0}(4)\right) = (0, 0.8850, 1.3851, 2.5368)\,,$$

and

$$X_4^{'0} = \left(x_4^{'0}(1), x_4^{'0}(2), x_4^{'0}(3), x_4^{'0}(4)\right) = (0, 0.5361, 1.6924, 2.8471)\,.$$

From

$$|s_i^{'}| = \left|\sum_{k=2}^{3} x_i^{'0}(k) + \frac{1}{2} \cdot x_i^{'0}(4)\right|,$$

$i = 0, 1, 2, 3, 4$, and

$$|s_i^{'} - s_0^{'}| = \left|\sum_{k=2}^{3}[x_i^{'0}(k) - x_0^{'0}(k)] + \frac{1}{2} \cdot [x_i^{'0}(4) - x_0^{'0}(4)]\right|,$$

$i = 1, 2, 3, 4$, it follows that

$$|s_0^{'}| = 2.7871, |s_1^{'}| = 0.80505, |s_2^{'}| = 1.35245,$$

$$|s_3^{'}| = 3.5385, |s_4^{'}| = 3.65205;$$

and

$$|s_1^{'} - s_0^{'}| = 1.98205, \quad |s_2^{'} - s_0^{'}| = 1.43465,$$

$$|s_3^{'} - s_0^{'}| = 0.7514, \quad |s_4^{'} - s_0^{'}| = 0.86495.$$

From

$$r_{0i} = \frac{1 + |s_0^{'}| + |s_i^{'}|}{1 + |s_0^{'}| + |s_i^{'}| + |s_i^{'} - s_0^{'}|},$$

$i = 1, 2, 3, 4$, it follows that

$$r_{01} = 0.6985, r_{02} = 0.7818, r_{03} = 0.9070, r_{04} = 0.8958.$$

3. Compute the synthetic degree of incidence. Take θ= 0.5. So, from

$$\rho_{0i} = \theta\varepsilon_{0i} + (1 - \theta)r_{0i},$$

$i = 1, 2, 3, 4$, it follows that

$$\rho_{01} = 0.6263, \rho_{02} = 0.6619, \rho_{03} = 0.7865, \rho_{04} = 0.7359.$$

4. Final analysis. From

$$\rho_{03} > \rho_{04} > \rho_{02} > \rho_{01},$$

it can be known that

$$X_3 \succ X_4 \succ X_2 \succ X_1,$$

with X_3 being the most favorable factor, X_4 the second, X_2 the third, and X_1 the last. That is to say, the labor forces have the greatest effect on production revenue of the county, after-tax profits have the second greatest effect, and the fixed capital has the least effect on the revenue. This result agrees very well with the actual situation in the region, where the non-governmental enterprises have been mainly (human) labor intensive types so that production growth has been mainly realized through increases of labor forces. In the countryside of China, there is an unlimited source of labor surplus. How to sufficiently and effectively make use of this supply of labor is the only way for China, with its current and special circumstances, to develop its commodity production and to bring about a prosperous economy. Therefore, actively developing businesses requiring intensive labor is the main direction for the near future development of Chinese non-governmental enterprises. As for after-tax profits, they have mainly been used for improving employees' fringe benefits and for innovations of technology. This end has been, stimulating the employees' enthusiasm for efficiency and more working hours, and on the other hand, increasing production qualities of businesses.

Example 5.10.2. In this example, we look at a grey incidence analysis for technological innovations in industry in Henan Province, The People's Republic of China.

From the years 1986 to 1987, the subcommittee for policies of the committee on science and technology of Henan Province had organized special forces to check into the then-current situations in technological innovations in state-owned industrial operations in Henan Province. Here, in this example, we conduct a grey incidence analysis based on the 89 replies from the related industrial operations.

We have seven characteristic variables for the system:

1. Y_1 – total industrial production output;
2. Y_2 – total net industrial production output;
3. Y_3 – total sales revenue;
4. Y_4 – total volume of profits;
5. Y_5 – net amount of production growth caused by technological innovations;
6. Y_6 – increased sales revenue caused by technological innovations;
7. Y_7 – increase in tax and profits caused by technological innovations;

and nine relevant factors:

1. X_1 – year-end value of fixed capital without considering depreciation;
2. X_2 – total value of circulating capital;
3. X_3 – total number of employees;
4. X_4 – numbers of engineers and technicians;
5. X_5 – number of people involved in technological innovations;
6. X_6 – investment in technological innovations;
7. X_7 – investment in fixed capital of the current year;
8. X_8 – investment used for new products, new equipment, and technology;
9. X_9 – the ratio of X_8 and X_6.

The relevant degrees of incidence between various Y_i and X_j are given in the following matrix.

$$\Gamma = \left[\gamma_{ij}\right]_{7\times 9} =$$

$$\begin{bmatrix} 0.867 & 0.912 & 0.885 & 0.909 & 0.577 & 0.744 & 0.708 & 0.760 & 0.828 \\ 0.833 & 0.877 & 0.926 & 0.872 & 0.566 & 0.732 & 0.688 & 0.734 & 0.832 \\ 0.898 & 0.942 & 0.854 & 0.947 & 0.578 & 0.737 & 0.724 & 0.778 & 0.799 \\ 0.612 & 0.624 & 0.701 & 0.626 & 0.501 & 0.567 & 0.574 & 0.569 & 0.749 \\ 0.487 & 0.478 & 0.576 & 0.480 & 0.397 & 0.539 & 0.549 & 0.922 & 0.431 \\ 0.368 & 0.365 & 0.355 & 0.367 & 0.423 & 0.375 & 0.395 & 0.376 & 0.352 \\ 0.470 & 0.458 & 0.426 & 0.463 & 0.424 & 0.557 & 0.519 & 0.518 & 0.435 \end{bmatrix}.$$

From

$$\sum_{i=1}^{4}\gamma_{i3} = 3.366 > \sum_{i=1}^{4}\gamma_{i2} = 3.355 > \sum_{i=1}^{4}\gamma_{i4} = 3.354$$

$$> \sum_{i=1}^{4}\gamma_{i1} = 3.210 > \sum_{i=1}^{4}\gamma_{i9} = 3.208 > \sum_{i=1}^{4}\gamma_{i8} = 2.841$$

$$> \sum_{i=1}^{4}\gamma_{i6} = 2.780 > \sum_{i=1}^{4}\gamma_{i7} = 2.669 > \sum_{i=1}^{4}\gamma_{i5} = 2.222,$$

and from the angle of influence on the entire industrial economy, it follows that

$$X_3 \succcurlyeq X_2 \succcurlyeq X_4 \succcurlyeq X_1 \succcurlyeq X_9 \succcurlyeq X_8 \succcurlyeq X_6 \succcurlyeq X_7 \succcurlyeq X_5.$$

That is, the number X_3 of employees is the most favorable factor, and the next are the circulating capital X_2 and the number X_4 of engineers and technicians.

Now, from

$$\sum_{i=5}^{7}\gamma_{i8} = 1.816 > \sum_{i=5}^{7}\gamma_{i7} = 1.488 > \sum_{i=5}^{7}\gamma_{i6} = 1.471$$

$$> \sum_{i=5}^{7}\gamma_{i3} = 1.36 > \sum_{i=5}^{7}\gamma_{i1} = 1.325 > \sum_{i=5}^{7}\gamma_{i4} = 1.31$$

$$> \sum_{i=5}^{7}\gamma_{i2} = 1.301 > \sum_{i=5}^{7}\gamma_{i5} = 1.244 > \sum_{i=5}^{7}\gamma_{i9} = 1.218,$$

and from the angle of influence on effects of technological innovations in industry, we have that

$$X_8 \succcurlyeq X_7 \succcurlyeq X_6 \succcurlyeq X_3 \succcurlyeq X_1 \succcurlyeq X_4 \succcurlyeq X_2 \succcurlyeq X_5 \succcurlyeq X_9.$$

That is, the investment used in introducing new products, new equipment, and technology is the most favorable factor, and investments X_7 and X_6 for fixed capital and technological innovations are the second most favorable factors.

6

Grey Clusters and Grey Statistical Evaluations

6.1 Introduction

Grey cluster is a method, based on matrices of grey incidences or whitenization weight functions of grey numbers, to classify observation indices or observational objects into definable classes. A cluster can be seen as a set of all observational objects arranged in the same class. In practical applications, very often each observational object possesses many characteristic features, which causes difficulties in accurate classification of the object. For example, *yin cai shi jiao*, meaning to teach students in accordance with their aptitude, has been debated in the Chinese community of teachers for many years. Because it is difficult or impossible to classify individual students into the right professions where they would potentially succeed, many of the results of the debate have not been implemented in the actual education practice. That is why the classical method of forming classes of groups of students based on their various different talents and interests has still been widely used in today's education in China, even though it has been proven that this method of teaching ruins talent and causes many inconveniences for most intelligent people. In business practices, similarly, due to the difficulty of accurately classifying human talents, many opportunities have been lost.

Grey cluster, in terms of the objects to be clustered, can be divided into a cluster of grey incidences and a cluster of whitenization weight functions of grey classes. The cluster of grey incidences is mainly used to classify factors of the same type in order to simplify complicated systems. Through clustering of grey incidences, we can check to see whether there exist some factors among many with close connections so that we can make use of a combined index (or criterion) or one of these factors to represent its group of factors, and at the same time, no information is seriously damaged. It belongs to the study of the problem of deleting variables of a system without altering its fundamental characteristics. Before a large-scale research or investigation is started, through the use of clustering of grey incidences of data values collected traditionally, the number of variables considered in the study can be greatly decreased so that the relevant costs of research will also be reduced. The cluster of whitenization weight functions of grey classes is mainly applied to check whether an observational object belongs to a predefined class. Otherwise the object could be treated differently. In practice, it is more complicated to use the cluster of whitenization weight functions of grey classes than the cluster of grey incidences.

Grey statistical evaluation is a method available for the researcher to check through the whole of the system under consideration and to determine which predefined classes a set of same-class observational objects belong to, based on a comprehensive evaluation of the objects. It has been mainly applied to synthetic evaluations of, for example, plans for production investments, divisions of agricultural economic districts, community planning, teaching schedules, etc., or to the final determination of an optimal or satisfactory plan after statistical methods have been applied to the relevant data.

6.2 Clusters of Grey Incidences

Assume that there exist n observational objects and that m characteristic data values for each of these objects have been collected. So, we have the sequences

$$X_i = (x_i(1), x_i(2), ..., x_i(n)), \ i = 1, 2, ..., m.$$

For all $i \leq j$, i, $j = 1, 2, ..., m$, we calculate the absolute degree ε_{ij} of incidence of X_i and X_j, and obtain the following upper triangular matrix A

$$A = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdot & \cdot & \cdot & \varepsilon_{1m} \\ & \varepsilon_{22} & \cdot & \cdot & \cdot & \varepsilon_{23} \\ & & \cdot & & & \cdot \\ & & & \cdot & & \cdot \\ & & & & \cdot & \cdot \\ & & & & & \varepsilon_{mm} \end{bmatrix},$$

where $\varepsilon_{ii} = 1$, $i = 1, 2, ..., m$.

Definition 6.2.1. *The previous matrix A is called the incidence matrix of the characteristic variables.*

Take a fixed critical value $r \in [0, 1]$ with the general requirement that $r > 0.5$. When $\varepsilon_{ij} \geq r$, $i \neq j$, the variables X_i and X_j are treated as those of the same characteristics.

Definition 6.2.2. *The classification of characteristic variables under a fixed critical value r is called a $cluster\ of\ r\ grey\ incidences$.*

The value of r can be chosen based on the practical needs involved in the study. The closer to 1 the value of r is, the finer the classification is with fewer variables in each class. On the other hand, the smaller the value of r is, the coarser the classification is with relatively more variables in each class.

Example 6.2.1. For the qualifications for a certain office, the search committee has proposed 15 criteria:

1. Impression on application package;
2. Academic abilities;
3. Likability by others;
4. Level of self-confidence;
5. Intelligence;
6. Honesty;
7. Ability to sell;
8. Experience;
9. Motivation;
10. Ambition;
11. Physical appearance;
12. Ability to comprehend;
13. Potential for future growth;
14. Interpersonal skills;
15. Adaptability.

Members of the committee admit that some of these 15 criteria could be well related or mixed and hope that through the study of some sample of a few data points, these 15 criteria can be classified into fewer categories. By using the method of scoring to quantify the criteria, 9 observational objects have been scored according to each of the criteria. The following matrix gives the scores, where O_i stands for the ith object, i = 1, 2, ..., 9:

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9
X_1	6	9	7	5	6	7	9	9	9
X_2	2	5	3	8	8	7	8	9	7
X_3	5	8	6	5	8	6	8	8	8
X_4	8	10	9	6	4	8	8	9	8
X_5	7	9	8	5	4	7	8	9	8
X_6	8	9	9	9	9	10	8	8	8
X_7	8	10	7	2	2	5	8	8	5
X_8	3	5	4	8	8	9	10	10	9
X_9	8	9	9	4	5	6	8	9	8
X_{10}	9	9	9	5	5	5	10	10	9
X_{11}	7	10	8	6	8	7	9	9	9
X_{12}	7	8	8	8	8	8	8	9	8
X_{13}	5	8	6	7	8	6	9	9	8
X_{14}	7	8	8	6	7	6	8	9	9
X_{15}	10	10	10	5	7	6	10	10	10

(6.1)

For all $i \leq j$, i, $j = 1, 2, ..., m$, we calculated the absolute degree ε_{ij} of incidence of X_i and X_j, and produced the following upper triangular matrix A

$$A = [\varepsilon_{ij}]_{15\times 15} =$$

$$\begin{bmatrix}
1 & .66 & .88 & .52 & .58 & .77 & .51 & .66 & .51 & .51 & .90 & .88 & .80 & .67 & .51 \\
 & 1.0 & .72 & .51 & .53 & .59 & .50 & .99 & .51 & .51 & .63 & .62 & .77 & .55 & .51 \\
 & & 1.0 & .51 & .56 & .70 & .51 & .72 & .51 & .51 & .80 & .78 & .90 & .63 & .51 \\
 & & & 1.0 & .56 & .53 & .58 & .51 & .69 & .62 & .52 & .52 & .51 & .54 & .60 \\
 & & & & 1.0 & .63 & .51 & .53 & .53 & .52 & .61 & .61 & .55 & .75 & .52 \\
 & & & & & 1.0 & .51 & .59 & .52 & .52 & .84 & .86 & .66 & .81 & .51 \\
 & & & & & & 1.0 & .50 & .70 & .83 & .51 & .51 & .51 & .51 & .89 \\
 & & & & & & & 1.0 & .51 & .51 & .63 & .62 & .77 & .55 & .51 \\
 & & & & & & & & 1.0 & .81 & .52 & .52 & .51 & .53 & .76 \\
 & & & & & & & & & 1.0 & .51 & .51 & .51 & .52 & .92 \\
 & & & & & & & & & & 1.0 & .97 & .74 & .71 & .51 \\
 & & & & & & & & & & & 1.0 & .73 & .72 & .51 \\
 & & & & & & & & & & & & 1.0 & .60 & .51 \\
 & & & & & & & & & & & & & 1.0 & .52 \\
 & & & & & & & & & & & & & & 1.0
\end{bmatrix}.$$

Applying this matrix can give us a cluster for the criteria, where the value for r can be different based on the requirements involved.

For example, let us take $r = 1$. Then, all the 15 criteria above belong to their own classes with each in its own class.

If we take $r = 0.80$, let us pick out all ε_{ij} values greater than 0.80 in the matrix A. So, we have

$$\varepsilon_{13} = 0.88,\ \varepsilon_{1,11} = 0.90,\ \varepsilon_{1,12} = 0.88,$$

$$\varepsilon_{1,13} = 0.80;\ \varepsilon_{28} = 0.99;\ \varepsilon_{3,11} = 0.80,$$

$$\varepsilon_{3,13} = 0.90;\ \varepsilon_{6,11} = 0.84,\ \varepsilon_{6,12} = 0.86,$$

$$\varepsilon_{6,14} = 0.81;\ \varepsilon_{7,10} = 0.83,\ \varepsilon_{7,15} = 0.89;$$

$$\varepsilon_{9,10} = 0.81;\ \varepsilon_{10,15} = 0.92;\ \varepsilon_{11,12} = 0.97.$$

So, we know that

- X_3, X_{11}, X_{12}, and X_{13} belong to the same class as X_1;
- X_8 belong to the same class as X_2;
- X_{11} and X_{13} belong to the same class as X_3;
- X_{11}, X_{12}, and X_{14} belong to the same class as X_6;
- X_{10} and X_{15} belong to the same class as X_7;
- X_{10} belong to the same class as X_9;
- X_{15} belong to the same class as X_{10};
- X_{12} belong to the same class as X_{11}.

Let each class be represented with the criterion with the minimum index contained in the class, and combine the classes, containing X_6 and X_{11}, respectively, with the class, containing X_1, put X_9 and X_{10} into the class, containing X_7, and treat X_4 and X_5 as individual classes. Then, we have obtained a cluster for our shortened list of criteria as follows,

$$\{X_1, X_3, X_6, X_{11}, X_{12}, X_{13}, X_{14}\},$$

$$\{X_2, X_8\}; \{X_4\}; \{X_5\}, \text{ and } \{X_7, X_9, X_{10}, X_{15}\}.$$

If further combination is needed, we can write out the incidence matrix for the 5 representative criteria, obtained after 0.80 clustering. This matrix is shown as follows.

$$A' = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{14} & \varepsilon_{15} & \varepsilon_{17} \\ & \varepsilon_{22} & \varepsilon_{24} & \varepsilon_{25} & \varepsilon_{27} \\ & & \varepsilon_{44} & \varepsilon_{45} & \varepsilon_{47} \\ & & & \varepsilon_{55} & \varepsilon_{57} \\ & & & & \varepsilon_{77} \end{bmatrix} = \begin{bmatrix} 1.0 & .66 & .52 & .58 & .51 \\ & 1.0 & .51 & .53 & .50 \\ & & 1.0 & .56 & .58 \\ & & & 1.0 & .51 \\ & & & & 1.0 \end{bmatrix}.$$

If we take $r = 0.58$, then in the previous clustering, the classes, containing X_2 and X_5, respectively, can be combined into the class of X_1, and X_4 can be combined into the same class with X_7. So, we obtain the following relatively coarser clustering,

$$\{X_1, X_2, X_3, X_5, X_6, X_8, X_{11}, X_{12}, X_{13}, X_{14}\}$$

and

$$\{X_4, X_7, X_9, X_{10}, X_{15}\},$$

where the class containing X_1 reflects the applicants' ability in character, personality, experience, etc. It includes the following aspects: committee members' impressions on application packages, academic ability, likability by others, intelligence, honesty, experiences, physical appearance, ability to comprehend, potential, and interpersonal skills. The class containing X_4 represents applicants' extrinsic ability, including creativity with five aspects: degree of self-confidence, ability to sell, motivation, ambition, and adaptability.

6.3 Clusters with Variable Weights

In this section, let us consider a more general situation than that considered in the previous section.

Definition 6.3.1. *Assume that there exist n objects to be clustered according to m cluster criteria into s different grey classes. The clustering method based on the observational value of the ith object, $i = 1, 2, \cdots, n$, at the jth criterion, $j = 1, 2, \cdots, m$, the ith object is classified into the kth grey class, $1 \le k \le s$, is called a grey clustering.*

Definition 6.3.2. *All the s grey classes formed by the n objects, defined by their observational values at criterion j, are called the j-criterion subclasses.*

The whitenization weight function of the kth subclass of the j-criterion is denoted $f_j^k(\cdot)$.

Definition 6.3.3. *Assume that the whitenization weight function $f_j^k(\cdot)$ of a j-criterion kth subclass is shown in Figure 6.1. Then the points $x_j^k(1)$, $x_j^k(2)$, $x_j^k(3)$, and $x_j^k(4)$ are called turning points of $f_j^k(\cdot)$.*

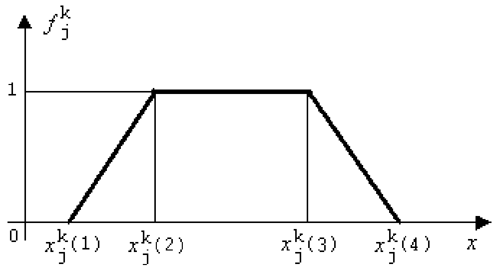

FIGURE 6.1. A typical whitenization function

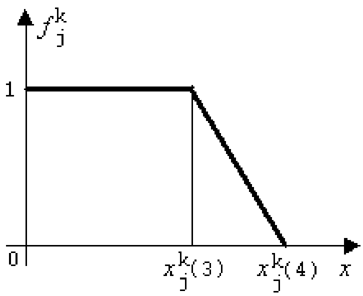

FIGURE 6.2. A whitenization weight function of lower measure

Definition 6.3.4. *1. If the whitenization weight function $f_j^k(\cdot)$ above does not have the first and the second turning points $x_j^k(1)$ and $x_j^k(2)$, as shown in Figure 6.2, then $f_j^k(\cdot)$ is called a whitenization weight function of lower measure.*

2. If the second $x_j^k(2)$ and the third $x_j^k(3)$ turning points of the whitenization weight function $f_j^k(\cdot)$ as in Figure 6.1 coincide, as shown in Figure 6.3, then $f_j^k(\cdot)$ is called a whitenization function of moderate (or middle) measure.

3. If the whitenization weight function $f_j^k(\cdot)$, as in Figure 6.1, does not have the third and fourth turning points $x_j^k(3)$ and $x_j^k(4)$, as shown in Figure 6.4, then $f_j^k(\cdot)$ is called a whitenization weight function of upper measure.

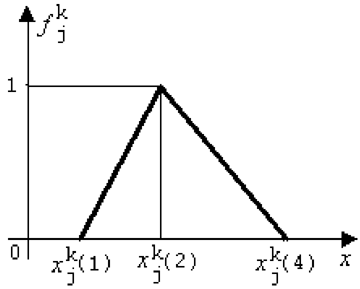

FIGURE 6.3. A whitenization weight function of middle measure

Proposition 6.3.1. *1. The typical whitenization weight function as shown in Figure 6.1 is given by*

$$f_j^k(x) = \begin{cases} 0, & x \notin [x_j^k(1), x_j^k(4)] \\ \dfrac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, & x \in [x_j^k(1), x_j^k(2)] \\ 1, & x \in [x_j^k(2), x_j^k(3)] \\ \dfrac{x_j^k(4) - x}{x_j^k(4) - x_j^k(3)}, & x \in [x_j^k(3), x_j^k(4)]. \end{cases}$$

2. The whitenization weight function of lower measure as shown in Figure 6.2 is given by

$$f_j^k(x) = \begin{cases} 0, & x \notin [0, x_j^k(4)] \\ 1, & x \in [0, x_j^k(3)] \\ \dfrac{x_j^k(4) - x}{x_j^k(4) - x_j^k(3)}, & x \in [x_j^k(3), x_j^k(4)]. \end{cases}$$

3. The whitenization weight function of moderate measure as shown in Figure 6.3 is given by

$$f_j^k(x) = \begin{cases} 0, & x \notin [x_j^k(1), x_j^k(4)] \\ \dfrac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, & x \in [x_j^k(1), x_j^k(2)] \\ 1, & x = x_j^k(2) \\ \dfrac{x_j^k(4) - x}{x_j^k(4) - x_j^k(2)}, & x \in [x_j^k(2), x_j^k(4)]. \end{cases}$$

4. The whitenization weight function of upper measure as shown in Figure (6.4) is given by

$$f_j^k(x) = \begin{cases} 0, & x < x_j^k(1) \\ \dfrac{x - x_j^k(1)}{x_j^k(2) - x_j^k(1)}, & x \in [x_j^k(1), x_j^k(2)] \\ 1, & x \geq x_j^k(2). \end{cases}$$

Definition 6.3.5. *1. For the whitenization weight function of the kth subclass of the j-criterion, as shown in Figure 6.1, define*

$$\lambda_j^k = \frac{1}{2}[x_j^k(2) + x_j^k(3)].$$

2. For the whitenization weight function of the kth subclass of the j-criterion as shown in Figure 6.2, let

$$\lambda_j^k = x_j^k(3).$$

3. For the whitenization weight functions of the kth subclass of the j-criterion as shown in Figures 6.3 and 6.4, let

$$\lambda_j^k = x_j^k(2).$$

Then λ_j^k is called the critical value for the kth subclass of the j-criterion.

Definition 6.3.6. *Assume that λ_j^k is the critical value for the kth subclass of the j-criterion. Then*

$$\eta_j^k = \frac{\lambda_j^k}{\sum_{j=1}^{m} \lambda_j^k}$$

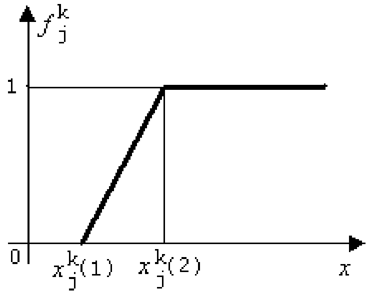

FIGURE 6.4. A whitenization weight function of upper measure

is called the weight of the j*-criterion with respect to the* k*th subclass.*

Definition 6.3.7. *Assume that* x_{ij} *is the observational value of object* i *with respect to criterion* j*,* $f_j^k(\cdot)$ *the whitenization weight function of the* k*th subclass of the* j*-criterion, and* η_j^k *the weight of the* j*-criterion with respect to the* k*th subclass. Then,*

$$\sigma_i^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \cdot \eta_j^k$$

is said to be the cluster coefficient of variable weight for object i *to belong to the* k*th grey class.*

Definition 6.3.8. *1. The following*

$$\begin{aligned}\sigma_i &= \left(\sigma_i^1, \sigma_i^2, \cdots, \sigma_i^s\right) \\ &= \left(\sum_{j=1}^{m} f_j^1(x_{ij}) \cdot \eta_j^1, \sum_{j=1}^{m} f_j^2(x_{ij}) \cdot \eta_j^2, \cdots, \sum_{j=1}^{m} f_j^s(x_{ij}) \cdot \eta_j^s\right)\end{aligned}$$

is called the cluster coefficient vector of object i*.*

2. The following matrix

$$\sum = \left[\sigma_i^k\right]_{n \times s}$$

is called the cluster coefficient matrix.

Definition 6.3.9. *If*

$$\sigma_i^{k^*} = \max_{1 \le k \le s} \left\{\sigma_i^k\right\},$$

then we say that object i *belongs to the grey class* k^**.*

Variable weight clustering is useful to cases of criteria with the same meanings and dimensions. When the meanings and dimensions of the criteria are different and the numbers of observational values of individual

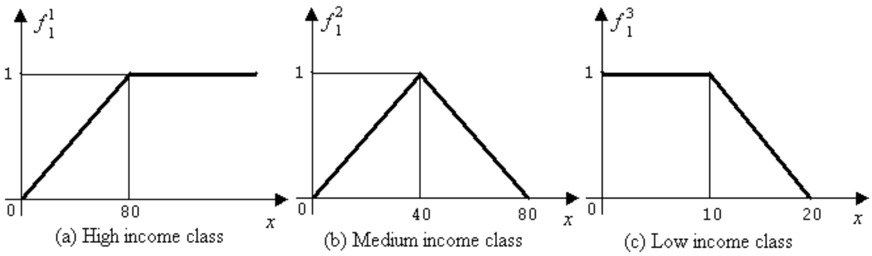

FIGURE 6.5. The whitenization weight function for the revenue from farming

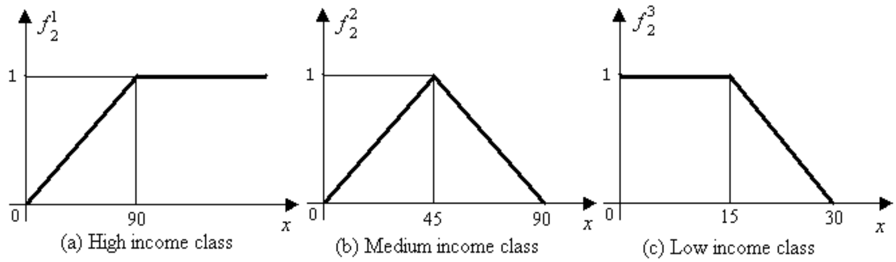

FIGURE 6.6. The whitenization weight function for the revenue from livestock husbandry

criteria are greatly different from each other, we should not use this method of clustering.

Example 6.3.1. Assume that we are interested in the study of three economic districts with the following three cluster criteria: revenue from farming, revenue from livestock husbandry, and revenue from industry. The observational values $x_{ij,}$ $i = 1, 2, 3$; $j = 1, 2, 3$, of the ith economic district with respect to the jth criterion is given in the following matrix A.

$$A = [x_{ij}]_{3\times 3} = \begin{bmatrix} 80 & 20 & 100 \\ 40 & 30 & 30 \\ 10 & 90 & 60 \end{bmatrix}.$$

Let us now perform a synthetic clustering based on high, medium, and low incomes.

Solution. Assume that the whitenization weight functions $f_j^k(\cdot)$, $j = 1, 2, 3$; $k = 1, 2, 3$, of the criteria: the revenues of farming, livestock husbandry and industry are as shown in Figures 6.5, 6.6 and 6.7.

From these figures and Proposition 6.3.1, it follows that

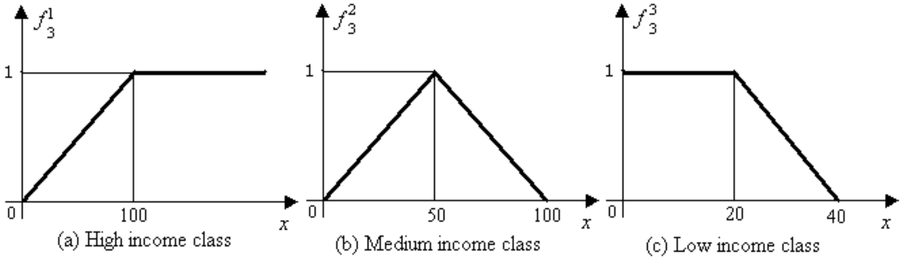

FIGURE 6.7. The whitenization weight function for the revenue from industry

$$f_1^1(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{80}, & 0 \le x < 80 \\ 1, & x \ge 80 \end{cases}, \quad f_1^2(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{40}, & 0 \le x \le 40 \\ \dfrac{80 - x}{40}, & 40 < x \le 80 \\ 0, & x > 80, \end{cases}$$

$$f_1^3(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 10 \\ \dfrac{20 - x}{10}, & 10 < x \le 20 \\ 0, & x > 20, \end{cases} \quad f_2^1(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{90}, & 0 \le x < 90 \\ 1, & x \ge 90, \end{cases}$$

$$f_2^2(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{45}, & 0 \le x \le 45 \\ \dfrac{90 - x}{45}, & 45 < x \le 90 \\ 0, & x > 90, \end{cases} \quad f_2^3(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 15 \\ \dfrac{30 - x}{15}, & 15 < x \le 30 \\ 0, & x > 30, \end{cases}$$

$$f_3^1(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{100}, & 0 \le x < 100 \\ 1, & x \ge 100, \end{cases} \quad f_3^2(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{50}, & 0 \le x \le 50 \\ \dfrac{100 - x}{50}, & 50 < x \le 100 \\ 0, & x > 100, \end{cases}$$

and

$$f_3^3(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 20 \\ \dfrac{40 - x}{20}, & 20 < x \le 40 \\ 0, & x > 40. \end{cases}$$

Therefore,

$$\lambda_1^1 = 80,\ \lambda_2^1 = 90,\ \lambda_3^1 = 100;$$

$$\lambda_1^2 = 40,\ \lambda_2^2 = 45,\ \lambda_3^2 = 50;$$

$$\lambda_1^3 = 10,\ \lambda_2^3 = 15,\ \lambda_3^3 = 20.$$

So, from

$$\eta_j^k = \frac{\lambda_j^k}{\sum_{j=1}^{3} \lambda_j^k},$$

we obtain that

$$\eta_1^1 = \frac{\lambda_1^1}{\sum_{j=1}^{3} \lambda_j^1} = \frac{80}{270},\ \eta_2^1 = \frac{\lambda_2^1}{\sum_{j=1}^{3} \lambda_j^1} = \frac{90}{270},\ \eta_3^1 = \frac{\lambda_3^1}{\sum_{j=1}^{3} \lambda_j^1} = \frac{100}{270};$$

$$\eta_1^2 = \frac{\lambda_1^2}{\sum_{j=1}^{3} \lambda_j^2} = \frac{40}{135},\ \eta_2^2 = \frac{\lambda_2^2}{\sum_{j=1}^{3} \lambda_j^2} = \frac{45}{135},\ \eta_3^2 = \frac{\lambda_3^2}{\sum_{j=1}^{3} \lambda_j^2} = \frac{50}{135};$$

$$\eta_1^3 = \frac{\lambda_1^3}{\sum_{j=1}^{3} \lambda_j^3} = \frac{10}{45},\ \eta_2^3 = \frac{\lambda_2^3}{\sum_{j=1}^{3} \lambda_j^3} = \frac{15}{45},\ \eta_3^3 = \frac{\lambda_3^3}{\sum_{j=1}^{3} \lambda_j^3} = \frac{20}{45}.$$

And from

$$\sigma_i^k = \sum_{j=1}^{3} f_j^k(x_{ij}) \cdot \eta_j^k,$$

when $i = 1$, we have

$$\sigma_1^1 = \sum_{j=1}^{3} f_j^1(x_{1j}) \cdot \eta_j^1 = 0.74, \ \sigma_1^2 = \sum_{j=1}^{3} f_j^2(x_{1j}) \cdot \eta_j^2 = 0.15,$$

$$\sigma_1^3 \quad = \sum_{j=1}^{3} f_j^3(x_{1j}) \cdot \eta_j^3 = 0.22.$$

That is,

$$\sigma_1 = \left(\sigma_1^1, \sigma_1^2, \sigma_1^3\right) = (0.74, 0.15, 0.22)\,.$$

When $i = 2$, we can calculate similarly and obtain the following

$$\sigma_2^1 = \sum_{j=1}^{3} f_j^1(x_{2j}) \cdot \eta_j^1 = 0.37, \sigma_2^2 = \sum_{j=1}^{3} f_j^2(x_{2j}) \cdot \eta_j^2 = 0.74,$$

$$\sigma_2^3 = \sum_{j=1}^{3} f_j^3(x_{2j}) \cdot \eta_j^3 = 0.22,$$

so,

$$\sigma_2 = \left(\sigma_2^1, \sigma_2^2, \sigma_2^3\right) = (0.37, 0.74, 0.22)\,.$$

When $i = 3$, we obtain similarly the following

$$\sigma_3^1 = \sum_{j=1}^{3} f_j^1(x_{3j}) \cdot \eta_j^1 = 0.59, \sigma_3^2 = \sum_{j=1}^{3} f_j^2(x_{3j}) \cdot \eta_j^2 = 0.15,$$

$$\sigma_3^3 = \sum_{j=1}^{3} f_j^3(x_{3j}) \cdot \eta_j^3 = 0.22.$$

So,

$$\sigma_3 = \left(\sigma_3^1, \sigma_3^2, \sigma_3^3\right) = (0.59, 0.15, 0.22)\,.$$

Combining all the results obtained earlier, we have the coefficient matrix of grey cluster as follows.

$$\sum = \left[\sigma_i^k\right]_{3\times 3} = \begin{bmatrix} 0.74 & 0.15 & 0.22 \\ 0.37 & 0.74 & 0.22 \\ 0.59 & 0.15 & 0.22 \end{bmatrix}.$$

From

$$\max_{1\le k\le 3}\left\{\sigma_1^k\right\} = \sigma_1^1 = 0.74, \max_{1\le k\le 3}\left\{\sigma_2^k\right\} = \sigma_2^2 = 0.74,$$

and

$$\max_{1\leq k\leq 3}\left\{\sigma_3^k\right\}=\sigma_3^1=0.59,$$

it follows that the second economic district belongs to the medium income grey class, and the first and the third economic districts belong to the high income grey class. Furthermore, from the cluster coefficients $\sigma_1^1=0.74$ and $\sigma_3^1=0.59$, it follows that there still exist some differences between the first and the third districts, even though they all belong to the high income grey class. If the income grey classes are refined, say, we use the five grey classes: high income, mid-high income, medium income, mid-low income, and low income, then different results can be obtained.

Besides, in general, the whitenization weight function of the j-criterion k subclasses is determined based on experience. When resolving practical problems, one can determine the whitenization weight functions from either the angle of the objects of the clustering, or looking at all the same type objects in the whole system, not just the ones involved in the clustering. For example, in Example 6.3.1, we could determine the whitenization weight functions not only from the three economic districts in question, but also from the same level economic districts in a city, a province, or from around the nation. Therefore, the results of grey cluster evaluations can only be applied to a certain range, which is the same as the one used in the determination of relevant whitenization weight functions.

6.4 Clusters with Fixed Weights

When the criteria for clustering have different meanings, dimensions, and sizes of observational data, applying variable weight clusterings may lead to the problem that some criteria participate in the clustering process very weakly. There are two ways to resolve this problem: one is to first transform the sample of data values of the criteria into non-dimensional values by using either the initiating operator or averaging operator, then cluster the resultant criteria. In this way, all the clustering criteria will be treated equally in the clustering process. The other way is to define a weight for each individual criterion before starting the clustering process. In this section, we emphasize this second method.

Definition 6.4.1. *Assume that x_{ij} $(i=1,2,...,n;\ j=1,2,\cdots,m)$ is the observational value of object i with respect to criterion j, and $f_j^k(\cdot)$ $(j=1,\ 2,...,\ m;\ k=1,\ 2,...,\ s)$ the whitenization weight function of the kth subclass of the j-criterion. If the weight η_j^k of the j-criterion with respect to the kth subclass has nothing to do with k, $j=1,\ 2,...,\ m$; $k=1,$ $2,\ ...,\ s$, that is, for any k_1 and $k_2\in\{1,\ 2,\ ...,\ s\}$, one always has that $\eta_j^{k_1}=\eta_j^{k_2}$, then the superscript k in the symbol η_j^k will be omitted and η_j^k*

written as η_j instead, $j = 1, 2, ..., m$. And

$$\sigma_i^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \cdot \eta_j$$

is called the fixed weight cluster coefficient for the object i to belong to the kth grey class.

Definition 6.4.2. *Assume that x_{ij} $(i = 1, 2, ..., n; j = 1, 2, ..., m)$ stands for the observational value of the object i with respect to criterion j, and $f_j^k(\cdot)$ $(j$ = 1, 2, ..., m; k = 1, 2, ..., $s)$ the whitenization weight function of the kth subclass of the j-criterion. If for any j = 1, 2, ..., m, $\eta_j = \frac{1}{m}$ always holds true, then*

$$\sigma_i^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \cdot \eta_j = \frac{1}{m} \sum_{j=1}^{m} f_j^k(x_{ij})$$

is called the equal weight cluster coefficient for the object i to belong to the kth grey class.

Fixed weight clustering can be performed according to the following steps:

Step 1: Determine the whitenization weight function $f_j^k(\cdot)$, j = 1, 2, ..., m; k = 1, 2, ..., s.

Step 2: Give a cluster weight η_j to each criterion, j = 1, 2, ..., m, based on either prior experience or results from a qualitative analysis.

Step 3: Compute all fixed weight cluster coefficients

$$\sigma_i^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \cdot \eta_j, \ i = 1, 2, ..., n; k = 1, 2, ..., s.$$

from the whitenization weight functions $f_j^k(\cdot)$ $(j$ = 1, 2, ..., m; k = 1, 2, ..., $s)$, obtained in step 1, cluster weights η_j, j = 1, 2, ..., m, obtained in step 2, and the observational values x_{ij} of object i with respect to criterion j, $(i$ = 1, 2, ..., n; j = 1, 2, ..., $m)$.

Step 4: If $\sigma_i^{k^*} = \max_{1 \le k \le s} \left\{ \sigma_i^k \right\}$, then one can decide that object i belongs to the k^*th grey class.

Example 6.4.1. Let us perform a grey clustering for the ecological adaptation of major strains of trees commercially used in China.

China is a huge country with a very complicated ecological environment, within which different strains of trees require some obviously different growing conditions. The area where the trees of a certain strain have been currently growing to a certain degree reflects the adaptability of the strain to the ecological environment. We now classify ecological environmental conditions into four main quantification criteria:

1. Geographical ecological measure;

2. Temperature ecological measure;

3. Precipitation ecological measure; and

4. Arid ecological measure.

Here, geographical ecological measure is an index representing the geographical width of the region in which the strain of trees grow. The numerical value of this measure is given by the product of differences of longitudes in the directions of east and west and latitudes in the directions of south and north. The temperature ecological measure indicates the adaptability of the strain of trees to various temperatures. Its numerical value is computed by using the difference of annual average temperatures of the southern and the northern bounds of the growing region. The precipitation ecological measure is the characteristic for the adaptability of the trees to precipitation conditions. Its numerical value is recorded as the difference of the maximum annual average precipitation and the minimum annual average precipitation of all areas in the growing region. The arid ecological measure is selected to describe a strain's adaptability to arid conditions in the atmosphere. Its value is the difference of the maximum and the minimum annual average aridities[1] in different areas located in the growing region.

Some statistics of the four measures for the 17 main strains of trees planted in China are given in the following Table 6.1,

[1] Aridity is the ratio of the maximum possible evaporation amount and precipitation.

Table 6.1. Statistics of the 17 main strains of trees

No.	geo. eco. measure	temp. eco measure	prec. eco measure	arid eco. measure
1	22.50	4	0	0
2	79.37	6	600	0.75
3	144.00	7	300	0.75
4	300.00	6.1	189	12.0
5	456.00	12	250	12.0
6	189.00	8	700	1.50
7	369.00	8	1300	2.25
8	1127.11	16.2	550	3.0
9	260.00	11	600	1.0
10	200.00	8	600	1.25
11	475.00	10	1000	0.75
12	314.10	8	900	0.75
13	282.80	7.4	1300	0.50
14	240.00	8	1200	0.50
15	160.00	5	1000	0.25
16	270.00	8	1200	0.25
17	900	1	200	0

where the trees are coded as follows.

Table 6.2. Codes of the trees studied

No.	Tree	No.	Tree
1	Camphor pine	10	Chinese white poplar
2	Korean pine	11	Oak
3	Northeast China ash	12	Huashan pine
4	Diversiform-leaved poplar	13	Masson pine
5	Sacsaoul	14	China fir
6	Chinese pine	15	Bamboo
7	Oriental arborvitae	16	Camphor tree
8	White elm	17	Southern Asian pine
9	Dryland willow		

Let us do a grey clustering based on wide adaptability, medium adaptability, and narrow adaptability.

Solution. Because the meanings of the criteria are different and there exist great differences among the values observed, we apply the method of fixed weight clustering.

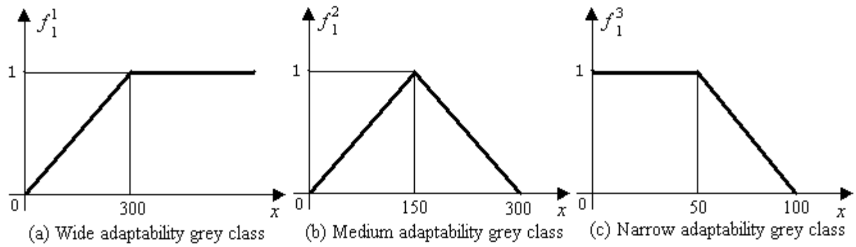

FIGURE 6.8. The whitenization weight function for geographical ecological measure

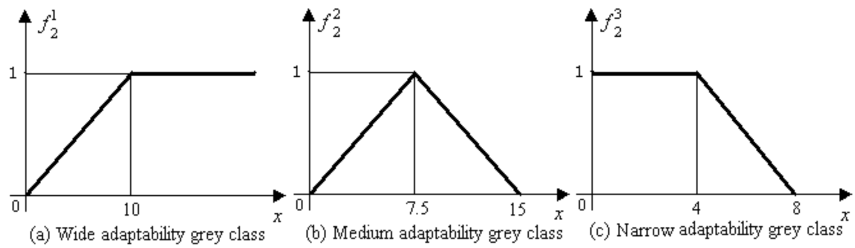

FIGURE 6.9. The whitenization weight function for temperature ecological measure

Step 1: Code the criteria and relative grey classes. Assume that the whitenization weight function $f_j^k(\cdot)$ of the kth subclass of the j-criterion, $j = 1, 2, 3, 4$, and $k = 1, 2, 3$, are as shown in Figures 6.8 to 6.11.

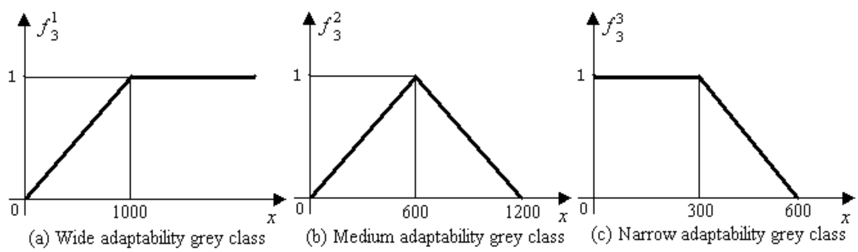

FIGURE 6.10. The whitenization weight function for precipitation ecological measure

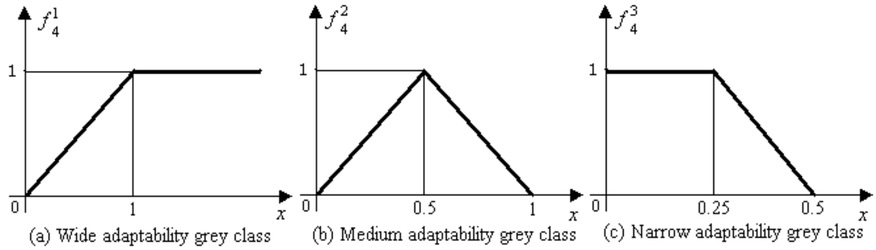

FIGURE 6.11. The whitenization weight function for aridity ecological measure

From Proposition 6.3.1 and Figures 6.8 to 6.11, it follows that

$$f_1^1(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{300}, & 0 \le x \le 300 \\ 1, & x > 300, \end{cases}$$

$$f_1^2(x) = \begin{cases} 0, & x < 0 \text{ or } x > 300 \\ \dfrac{x}{150}, & 0 \le x \le 150 \\ \dfrac{300 - x}{150}, & 150 < x \le 300, \end{cases}$$

$$f_1^3(x) = \begin{cases} 0, & x < 0 \text{ or } x > 100 \\ 1, & 0 \le x \le 50 \\ \dfrac{100 - x}{50}, & 50 < x \le 100, \end{cases} \qquad f_2^1(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{10}, & 0 \le x \le 10 \\ 1, & x > 10, \end{cases}$$

$$f_2^2(x) = \begin{cases} 0, & x < 0 \text{ or } x > 15 \\ \dfrac{x}{7.5}, & 0 \le x \le 7.5 \\ \dfrac{15 - x}{7.5}, & 7.5 < x \le 15, \end{cases}$$

$$f_2^3(x) = \begin{cases} 0, & x < 0 \text{ or } x > 8 \\ 1, & 0 \leq x \leq 4 \\ \dfrac{8-x}{4}, & 4 < x \leq 8, \end{cases}$$

$$f_3^1(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{1000}, & 0 \leq x \leq 1000 \\ 1, & x > 1000, \end{cases}$$

$$f_3^2(x) = \begin{cases} 0, & x < 0 \text{ or } x > 1200 \\ \dfrac{x}{600}, & 0 \leq x \leq 600 \\ \dfrac{1200-x}{600}, & 600 < x \leq 1200, \end{cases}$$

$$f_3^3(x) = \begin{cases} 0, & x < 0 \text{ or } x > 600 \\ 1, & 0 \leq x \leq 300 \\ \dfrac{600-x}{300}, & 300 < x \leq 600, \end{cases} \quad f_4^1(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1, \end{cases}$$

$$f_4^2(x) = \begin{cases} 0, & x < 0 \text{ or } x > 1 \\ \dfrac{x}{0.5}, & 0 \leq x \leq 0.5 \\ \dfrac{1-x}{0.5}, & 0.5 < x \leq 1, \end{cases}$$

$$f_4^3(x) = \begin{cases} 0, & x < 0 \text{ or } x > 0.5 \\ 1, & 0 \leq x \leq 0.25 \\ \dfrac{0.5-x}{0.25}, & 0.25 < x \leq 0.5. \end{cases}$$

Step 2: Let the weights for the geographical ecological, temperature ecological, precipitation ecological, and aridity ecological measures be

$$\eta_1 = 0.3, \eta_2 = 0.25, \eta_3 = 0.25, \eta_4 = 0.2.$$

Step 3: From

$$\sigma_i^k = \sum_{j=1}^{4} f_j^k(x_{ij}) \cdot \eta_j, \ i = 1, 2, \cdots, 17; k = 1, 2, 3,$$

Table 6.1, and the results in the previous two steps, it follows that:
When $i = 1$,

$$\sigma_1^1 = \sum_{j=1}^{4} f_j^1(x_{1j}) \cdot \eta_j = 0.12, \sigma_1^2 = \sum_{j=1}^{4} f_j^2(x_{1j}) \cdot \eta_j = 0.18,$$

$$\sigma_1^3 \quad = \sum_{j=1}^{4} f_j^3(x_{1j}) \cdot \eta_j = 1.$$

So

$$\sigma_1 = \left(\sigma_1^1, \sigma_1^2, \sigma_1^3\right) = (0.12, 0.18, 1).$$

When $i = 2$,

$$\sigma_2^1 = \sum_{j=1}^{4} f_j^1(x_{2j}) \cdot \eta_j = 0.53, \sigma_2^2 = \sum_{j=1}^{4} f_j^2(x_{2j}) \cdot \eta_j = 0.71,$$

$$\sigma_2^3 = \sum_{j=1}^{4} f_j^3(x_{2j}) \cdot \eta_j = 0.25.$$

So,

$$\sigma_2 = \left(\sigma_2^1, \sigma_2^2, \sigma_2^3\right) = (0.53, 0.71, 0.25).$$

When $i = 3$,

$$\sigma_3^1 = \sum_{j=1}^{4} f_j^1(x_{3j}) \cdot \eta_j = 0.54, \sigma_3^2 = \sum_{j=1}^{4} f_j^2(x_{3j}) \cdot \eta_j = 0.75,$$

$$\sigma_3^3 = \sum_{j=1}^{4} f_j^3(x_{3j}) \cdot \eta_j = 0.31.$$

So

$$\sigma_3 = \left(\sigma_3^1, \sigma_3^2, \sigma_3^3\right) = (0.54, 0.75, 0.31).$$

When $i = 4$,

$$\sigma_4^1 = \sum_{j=1}^{4} f_j^1(x_{4j}) \cdot \eta_j = 0.70, \sigma_4^2 = \sum_{j=1}^{4} f_j^2(x_{4j}) \cdot \eta_j = 0.28,$$

$$\sigma_4^3 = \sum_{j=1}^{4} f_j^3(x_{4j}) \cdot \eta_j = 0.37.$$

So,

$$\sigma_4 = \left(\sigma_4^1, \sigma_4^2, \sigma_4^3\right) = (0.70, 0.28, 0.37)\,.$$

Similarly, we can calculate and obtain

$$\sigma_5 = \left(\sigma_5^1, \sigma_5^2, \sigma_5^3\right) = (0.81, 0.20, 0.25)\,,$$

$$\sigma_6 = \left(\sigma_6^1, \sigma_6^2, \sigma_6^3\right) = (0.76, 0.66, 0.00)\,,$$

$$\sigma_7 = \left(\sigma_7^1, \sigma_7^2, \sigma_7^3\right) = (0.95, 0.23, 0.00)\,,$$

$$\sigma_8 = \left(\sigma_8^1, \sigma_8^2, \sigma_8^3\right) = (0.89, 0.23, 0.04)\,,$$

$$\sigma_9 = \left(\sigma_9^1, \sigma_9^2, \sigma_9^3\right) = (0.86, 0.46, 0.00)\,,$$

$$\sigma_{10} = \left(\sigma_{10}^1, \sigma_{10}^2, \sigma_{10}^3\right) = (0.75, 0.68, 0.00)\,,$$

$$\sigma_{11} = \left(\sigma_{11}^1, \sigma_{11}^2, \sigma_{11}^3\right) = (0.99, 0.35, 0.00)\,,$$

$$\sigma_{12} = \left(\sigma_{12}^1, \sigma_{12}^2, \sigma_{12}^3\right) = (0.91, 0.48, 0.00)\,,$$

$$\sigma_{13} = \left(\sigma_{13}^1, \sigma_{13}^2, \sigma_{13}^3\right) = (0.82, 0.48, 0.04)\,,$$

$$\sigma_{14} = \left(\sigma_{14}^1, \sigma_{14}^2, \sigma_{14}^3\right) = (0.79, 0.55, 0.00)\,,$$

$$\sigma_{15} = \left(\sigma_{15}^1, \sigma_{15}^2, \sigma_{15}^3\right) = (0.59, 0.63, 0.39)\,,$$

$$\sigma_{16} = \left(\sigma_{16}^1, \sigma_{16}^2, \sigma_{16}^3\right) = (0.77, 0.39, 0.20)\,,$$

and

$$\sigma_{17} = \left(\sigma_{17}^1, \sigma_{17}^2, \sigma_{17}^3\right) = (0.08, 0.13, 1.00)\,.$$

Step 4: From the following facts

$$\max_{1 \le k \le 3} \left\{\sigma_1^k\right\} = 1.00 = \sigma_1^3,\ \max_{1 \le k \le 3} \left\{\sigma_2^k\right\} = 0.71 = \sigma_2^2,$$

$$\max_{1 \le k \le 3} \left\{\sigma_3^k\right\} = 0.75 = \sigma_3^2,\ \max_{1 \le k \le 3} \left\{\sigma_4^k\right\} = 0.70 = \sigma_4^1,$$

$$\max_{1 \le k \le 3} \left\{\sigma_5^k\right\} = 0.81 = \sigma_5^1,\ \max_{1 \le k \le 3} \left\{\sigma_6^k\right\} = 0.76 = \sigma_6^1,$$

$$\max_{1 \le k \le 3} \left\{\sigma_7^k\right\} = 0.95 = \sigma_7^1,\ \max_{1 \le k \le 3} \left\{\sigma_8^k\right\} = 0.89 = \sigma_8^1,$$

$$\max_{1 \le k \le 3} \left\{\sigma_9^k\right\} = 0.86 = \sigma_9^1,\ \max_{1 \le k \le 3} \left\{\sigma_{10}^k\right\} = 0.75 = \sigma_{10}^1,$$

$$\max_{1 \le k \le 3} \left\{\sigma_{11}^k\right\} = 0.99 = \sigma_{11}^1,\ \max_{1 \le k \le 3} \left\{\sigma_{12}^k\right\} = 0.91 = \sigma_{12}^1,$$

$$\max_{1 \le k \le 3} \left\{\sigma_{13}^k\right\} = 0.82 = \sigma_{13}^1,\ \max_{1 \le k \le 3} \left\{\sigma_{14}^k\right\} = 0.79 = \sigma_{14}^1,$$

$$\max_{1 \le k \le 3} \left\{ \sigma_{15}^k \right\} = 0.63 = \sigma_{15}^2, \max_{1 \le k \le 3} \left\{ \sigma_{16}^k \right\} = 0.77 = \sigma_{16}^1,$$

and

$$\max_{1 \le k \le 3} \left\{ \sigma_{17}^k \right\} = 1.00 = \sigma_{17}^3,$$

it follows that the trees with the numberings 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, are strains with wide adaptability. They are diversiform-leaved poplars, sacsaouls, Chinese pines, oriental arborvitaes, white elms, dryland willows, Chinese white poplars, oaks, Huashan pines, masson pines, China firs, and camphor trees. These trees have an extremely strong ability to adapt themselves to natural ecological environments, and can grow well in most parts of China, and should be widely introduced. The trees named Korean pine, Northeast China Ash, and bamboo with numberings 2, 3, and 15, respectively, belong to the grey class of medium adaptability, and can be introduced to a relatively large area in China. And the trees with the names camphor pine and Southern Asian pine and numberings 1 and 17, respectively, belong to the grey class of narrow adaptability, where camphor pines are found near the northern border of China and southern Asian pines are mainly located near the Southern border of China.

6.5 Grey Evaluation Based on Triangular Whitenization Functions

Assume that n objects have been clustered into s different grey classes according to m evaluation criteria. Let x_{ij}, $i = 1, 2, ..., n$, $j = 1, 2, ..., m$, be the observational value of object i in terms of criterion j. We need to apply the values x_{ij}, $j = 1, 2, ..., m$, to evaluate and analyze object i, $i = 1, 2, ..., n$. To achieve this end, we only need to go through the following steps.

Step 1. Based on the predetermined number s of grey classes for the planned evaluation, divide the individual ranges of the criteria into s grey classes. For example, let $[a_1, a_{s+1}]$ be the range of the values of criterion j. Now, divide $[a_1, a_{s+1}]$ into s grey classes as follows,

$$[a_1, a_2], ..., [a_k, a_{k+1}], ..., [a_{s-1}, a_s], [a_s, a_{s+1}],$$

where a_k, $k = 1, 2, ..., s$, in general, can be determined based on specific requirements of a situation or relevant qualitative analysis.

Step 2. Let the whitenization weight function value for $(a_k + a_{k+1})/2$ to belong to the kth grey class be 1. When $\left(\frac{a_k+a_{k+1}}{2}, 1\right)$ is connected to the starting point a_{k-1} of the $(k-1)$th grey class and the ending point a_{k+2} of the $(k+1)$th grey class, one obtains a triangular whitenization weight function $f_j^k(\cdot)$ in terms of criterion j about the kth grey class, $j = 1, 2,$

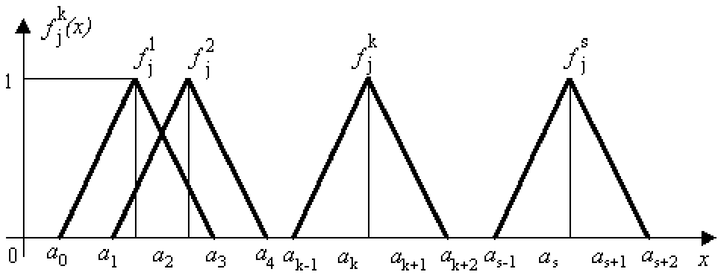

FIGURE 6.12. Construction of triangular whitenization weight functions

..., m, $k = 1, 2, ..., s$. For $f_j^1(\cdot)$ and $f_j^s(\cdot)$, the range of criterion j can be extended to the left and the right to a_0 and a_{s+2}, respectively (see Figure 6.12 for more details).

For any observational value x of criterion j, one can use the following

$$f_j^k(x) = \begin{cases} 0, & x \notin [a_{k-1}, a_{k+2}] \\ \dfrac{x - a_{k-1}}{\lambda_k - a_{k-1}}, & x \in [a_{k-1}, \lambda_k] \\ \dfrac{a_{k+2} - x}{a_{k+2} - \lambda_k}, & x \in [\lambda_k, a_{k+2}] \end{cases} \tag{6.2}$$

to compute the degree of membership $f_j^k(x)$ for x to belong to the kth grey class, $k = 1, 2, ..., s$, and $\lambda_k = \dfrac{a_k + a_{k+1}}{2}$.

Step 3. Compute the cluster coefficient σ_i^k for object i, $i = 1, 2, ..., n$, in terms of the kth grey class, $k = 1, 2, ..., s$:

$$\sigma_i^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \cdot \eta_j, \tag{6.3}$$

where $f_j^k(x_{ij})$ stands for the whitenization weight function for object i to belong to the kth grey class under criterion j, and η_j the weight of criterion j of the clustering.

Step 4. If $\max_{1 \le k \le s} \{\sigma_i^k\} = \sigma_i^{k^*}$, then object i belongs to the k^*th grey class. When more than one object belongs to the k^*th grey class, one can further determine the order of preference among these objects based on the magnitudes of their individual cluster coefficients.

6.6 Grey Statistics

In this section, let us look at how criteria used in a study can be clustered.

Definition 6.6.1. *Assume that there exist n statistical objects and m statistical criteria, and these criteria will be clustered into s grey classes. Classifying criterion j into the kth grey class, $k = 1, 2, ..., s$, based on the observational values x_{ij}, of the n objects with respect to the jth criterion, $i = 1, 2, ..., n$, and $j = 1, 2, ..., m$, is called grey statistics.*

Definition 6.6.2. *Assume that x_{ij} $(i = 1, 2, ..., n;\ j = 1, 2, ..., m)$ is the observational value of the ith statistical object with respect to the jth criterion, $f^k(\cdot)$ $(k = 1, 2, ..., s)$ the whitenization weight function of the kth grey class, and η_i $(i = 1, 2, ..., n)$ the weight of object i satisfying $\sum_{i=1}^{n}\eta_i = 1$.*

1. When $\eta_i = \frac{1}{n}$, $i = 1, 2, ..., n$, then

$$\sigma_j^k = \frac{\sum_{i=1}^{n} f^k(x_{ij})\cdot\eta_i}{\sum_{k=1}^{s}\sum_{i=1}^{n} f^k(x_{ij})\cdot\eta_i},$$

$j = 1, 2, ..., m$; $k = 1, 2, ..., s$, is called the statistical coefficient of equal weight objects.

2. When there are i_1 and $i_2 \in \{1, 2, ..., n\}$ so that $\eta_{i_1} \neq \eta_{i_2}$, the following

$$\sigma_j^k = \frac{\sum_{i=1}^{n} f^k(x_{ij})\cdot\eta_i}{\sum_{k=1}^{s}\sum_{i=1}^{n} f^k(x_{ij})\cdot\eta_i},$$

$j = 1, 2, ..., m$; $k = 1, 2, ..., s$, is called the statistical coefficient of unequal weight objects.

Definition 6.6.3. *Let σ_j^k be defined as in Definition 6.6.2. Then*

$$\sigma_j = \left(\sigma_j^1, \sigma_j^2, ..., \sigma_j^s\right),$$

$j = 1, 2, ..., m$, is called the vector of statistical coefficients of the criterion j.

Definition 6.6.4. *The matrix*

$$\sum = \left[\sigma_j^k\right]_{m\times s} = \begin{bmatrix} \sigma_1^1 & \sigma_1^2 & \cdots & \sigma_1^s \\ \sigma_2^1 & \sigma_2^2 & \cdots & \sigma_2^s \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_m^1 & \sigma_m^2 & \cdots & \sigma_m^s \end{bmatrix}$$

is called the matrix of statistical coefficients.

Definition 6.6.5. *If* $\sigma_j^{k^*} = \max_{1 \le k \le s} \left\{ \sigma_j^k \right\}$, *then we say that the* j*th statistical criterion, seen from the whole of the system under consideration, belongs to the grey class* k^*.

Grey statistics can be performed according to the following steps.

Step 1: Construct the observation matrix of the object i with respect to the criterion j, $i = 1, 2, ..., n$ $j = 1, 2, ..., m$:

$$A = [x_{ij}]_{n \times m} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}.$$

Step 2: Determine statistical grey classes and the whitenization weight functions $f^k(\cdot)$, $k = 1, 2, ..., s$.

Step 3: Determine the grey statistical weight η_i for object i, $i = 1, 2, ..., n$.

Step 4: Compute statistical coefficients.

$$\sigma_j^k = \frac{\sum_{i=1}^{n} f^k(x_{ij}) \cdot \eta_i}{\sum_{k=1}^{s} \sum_{i=1}^{n} f^k(x_{ij}) \cdot \eta_i}.$$

Step 5: Construct the statistical vectors

$$\sigma_j = \left(\sigma_j^1, \sigma_j^2, ..., \sigma_j^s \right)$$

and the matrix of statistical coefficients

$$\sum = \left[\sigma_j^k \right]_{m \times s}.$$

Step 6: Decide to which grey class the n statistical objects, seen from the whole of the system under consideration with respect to the criterion j, belong. If $\sigma_j^{k^*} = \max_{1 \le k \le s} \left\{ \sigma_j^k \right\}$, then the jth criterion belong to the grey class k^*.

When the statistical objects under consideration are different economic districts or different departments of a government entity related to economic development, and the economic criteria are different business types, grey statistics can be applied to analyze and to synthetically evaluate groups of economic bodies with respect to different business types to decide

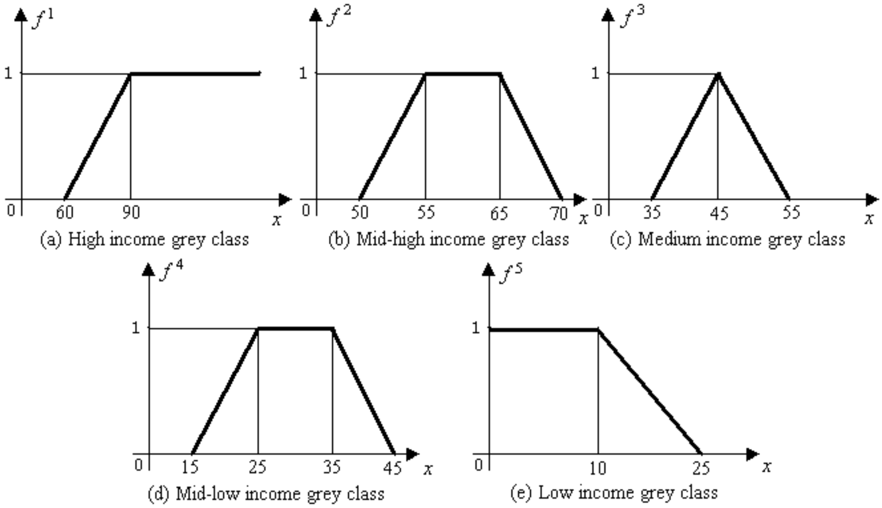

FIGURE 6.13. The whitenization weight functions of all grey classes

which business types warrant more attention. When the statistical objects are different sectors of a decision-making unit, and when the statistical criteria are different decision-making plans, the relevant grey statistics can synthesize all different ideas from different sectors, evaluate all decision-making plans, and select the optimal plan.

Example 6.6.1. For the three economic districts discussed in Example 6.3.1, the observational matrix for the three revenue criteria: farming, livestock husbandry, and industry, is given as follows,

$$A = [x_{ij}]_{3\times 3} = \begin{bmatrix} 80 & 20 & 100 \\ 40 & 30 & 30 \\ 10 & 90 & 60 \end{bmatrix}.$$

Perform a grey statistics for the three kinds of revenues of the economic districts based on the following five grey classes: high income class, mid-high income class, medium class, mid-low income class, and low income class.

Solution: Step 1: The observational matrix X has been given.

Step 2: Assume that the whitenization weight functions of all grey classes are given as in Figure 6.13.

Combining Figure (6.13) and Proposition 6.3.1, we can obtain that

$$f^1(x) = \begin{cases} 0, & x < 60 \\ \dfrac{x-60}{30}, & 60 \le x \le 90 \\ 1, & x > 90, \end{cases} \quad f^2(x) = \begin{cases} 0, & x \notin [50, 70] \\ \dfrac{x-50}{5}, & 50 \le x < 55 \\ 1, & 55 \le x \le 65 \\ \dfrac{70-x}{5}, & 65 < x \le 70, \end{cases}$$

$$f^3(x) = \begin{cases} 0, & x \notin [35, 55] \\ \dfrac{x-35}{10}, & 35 \le x \le 45 \\ \dfrac{55-x}{10}, & 45 < x \le 55, \end{cases} \quad f^4(x) = \begin{cases} 0, & x \notin [15, 45] \\ \dfrac{x-15}{10}, & 15 \le x < 25 \\ 1, & 25 \le x \le 35 \\ \dfrac{45-x}{10}, & 35 < x \le 45, \end{cases}$$

and

$$f^5(x) = \begin{cases} 0, & x \notin [0, 25] \\ 1, & 0 \le x \le 10 \\ \dfrac{25-x}{15}, & 10 < x \le 25. \end{cases}$$

Step 3: Now, we do grey statistical analysis with equal weight objects. That is, we take $\eta_i = \frac{1}{3}$, $i = 1, 2, 3$.

Step 4: Compute statistical coefficients. From that

$$\sigma_j^k = \frac{\sum_{i=1}^{3} f^k(x_{ij}) \cdot \eta_i}{\sum_{k=1}^{5}\sum_{i=1}^{3} f^k(x_{ij}) \cdot \eta_i},$$

it follows that when $j = 1$,

$$\sum_{i=1}^{3} f^1(x_{i1}) \cdot \eta_i = \frac{2}{9}, \ \sum_{i=1}^{3} f^2(x_{i1}) \cdot \eta_i = 0,$$

$$\sum_{i=1}^{3} f^3(x_{i1}) \cdot \eta_i = \frac{1}{6}, \ \sum_{i=1}^{3} f^4(x_{i1}) \cdot \eta_i = \frac{1}{6},$$

$$\sum_{i=1}^{3} f^5(x_{i1}) \cdot \eta_i = \frac{1}{3}.$$

Therefore,

$$\sum_{k=1}^{5}\sum_{i=1}^{3} f^k(x_{i1}) \cdot \eta_i = \frac{2}{9} + 0 + \frac{1}{6} + \frac{1}{6} + \frac{1}{3} = \frac{8}{9}.$$

So,

$$\sigma_1^1 = \frac{2}{9} \div \frac{8}{9} = \frac{1}{4}, \sigma_1^2 = 0, \sigma_1^3 = \frac{1}{6} \div \frac{8}{9} = \frac{3}{16},$$
$$\sigma_1^4 = \frac{1}{6} \div \frac{8}{9} = \frac{3}{16}, \text{ and } \sigma_1^5 = \frac{1}{3} \div \frac{8}{9} = \frac{3}{8}.$$

Then

$$\sigma_1 = \left(\sigma_1^1, \sigma_1^2, \sigma_1^3, \sigma_1^4, \sigma_1^5\right) = \left(\frac{1}{4}, 0, \frac{3}{16}, \frac{3}{16}, \frac{3}{8}\right).$$

When $j = 2$, the similar computations can be used to obtain that

$$\sum_{i=1}^{3} f^1(x_{i2}) \cdot \eta_i = \tfrac{1}{3}, \sum_{i=1}^{3} f^2(x_{i2}) \cdot \eta_i = 0, \sum_{i=1}^{3} f^3(x_{i2}) \cdot \eta_i = 0,$$

$$\sum_{i=1}^{3} f^4(x_{i2}) \cdot \eta_i = \tfrac{1}{2}, \sum_{i=1}^{3} f^5(x_{i2}) \cdot \eta_i = \tfrac{1}{9}.$$

So,

$$\sum_{k=1}^{5}\sum_{i=1}^{3} f^k(x_{i2}) \cdot \eta_i = \frac{17}{18}.$$

Therefore,

$$\sigma_2^1 = \frac{6}{17},\ \sigma_2^2 = 0,\ \sigma_2^3 = 0,\ \sigma_2^4 = \frac{9}{17},\ \sigma_2^5 = \frac{2}{17}$$

and

$$\sigma_2 = \left(\sigma_2^1, \sigma_2^2, \sigma_2^3, \sigma_2^4, \sigma_2^5\right) = \left(\frac{6}{17}, 0, 0, \frac{9}{17}, \frac{2}{17}\right).$$

When $j = 3$, we can obtain

$$\sum_{i=1}^{3} f^1(x_{i3}) \cdot \eta_i = \tfrac{1}{3}, \sum_{i=1}^{3} f^2(x_{i3}) \cdot \eta_i = \tfrac{1}{3}, \sum_{i=1}^{3} f^3(x_{i3}) \cdot \eta_i = 0,$$

$$\sum_{i=1}^{3} f^4(x_{i3}) \cdot \eta_i = \tfrac{1}{3}, \sum_{i=1}^{3} f^5(x_{i3}) \cdot \eta_i = 0.$$

So,

$$\sum_{k=1}^{5}\sum_{i=1}^{3} f^k(x_{i3}) \cdot \eta_i = 1.$$

Therefore,

$$\sigma_3^1 = \frac{1}{3},\ \sigma_3^2 = \frac{1}{3},\ \sigma_3^3 = 0,\ \sigma_3^4 = \frac{1}{3},\ \sigma_3^5 = 0,$$

and

$$\sigma_3 = \left(\sigma_3^1, \sigma_3^2, \sigma_3^3, \sigma_3^4, \sigma_3^5\right) = \left(\frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3}, 0\right).$$

Step 5: Write down the matrix of statistical coefficients.

$$\sum = [\sigma_j^k]_{3\times 5} = \begin{bmatrix} \frac{1}{4} & 0 & \frac{3}{16} & \frac{3}{16} & \frac{3}{8} \\ \frac{6}{17} & 0 & 0 & \frac{9}{17} & \frac{2}{17} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix}.$$

Step 6: Determine grey classes. From

$$\max_{1\le k\le 5} \{\sigma_1^k\} = \frac{3}{8} = \sigma_1^5,\ \max_{1\le k\le 5} \{\sigma_2^k\} = \frac{9}{17} = \sigma_2^4$$

and

$$\max_{1\le k\le 5} \{\sigma_3^k\} = \frac{1}{3} = \sigma_3^1 = \sigma_3^2 = \sigma_3^4,$$

it follows that among the three economic districts, seen from the whole of the system, the revenue from farming belongs to the low income class, the revenue from livestock husbandry belongs to the mid-low income class. As for the revenue from industry, the conclusion is not unique. Now, we further consider

$$\frac{1}{3}(1+2+4) = \frac{7}{3},$$

on the account of

$$\min\left\{\left|1-\frac{7}{3}\right|, \left|2-\frac{7}{3}\right|, \left|4-\frac{7}{3}\right|\right\} = \left|2-\frac{7}{3}\right|,$$

so, we could take $k^* = 2$. That is, from the angle of the whole system, it can be seen that the revenue from industry belongs to the grey class of mid-high income.

6.7 Entropy of Coefficient Vector of Grey Evaluations

For evaluations of clusters of grey incidences, the greyness of their results is expressed in terms of closeness among degrees ε_i, $i = 1, 2, ..., m$, or ε_{ij}, $i, j = 1, 2, ..., s$, of grey incidence. If the degrees ε_i, $i = 1, 2, ..., m$, or ε_{ij}, $i, j = 1, 2, ..., s$, are very close to each other, the degree of greyness of the outcome of the evaluation will be increased. For evaluations of grey clusters with either variable weights or fixed weights or those based on triangular whitenization weight functions, their degrees of greyness are expressed in terms of the closeness among the cluster coefficients σ_i^k, $k = 1, 2, ..., s$. For the sake of convenience, in the following, we use the vector

$$\sigma_i = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^s\right)$$

to represent the result of various chosen grey evaluations. Here, the degree of greyness of an evaluation result is expressed by the degree of balance among the components of σ_i. The more balanced the components of σ_i are, the more greyness the evaluation result. Again, for the sake of convenience and without loss of generality, let us assume $\sum_{k=1}^{s} \sigma_i^k = 1$.

Definition 6.7.1. *The following sum*

$$I(\sigma_i) = -\sum_{k=1}^{s} \sigma_i^k \ln \sigma_i^k \tag{6.4}$$

is called the entropy of the grey evaluation coefficient vector σ_i.

The entropy value $I(\sigma_i)$, as defined in equ. (6.4), can be used as a measure for the degree of balance among the components of the grey evaluation coefficient vector σ_i. The more balanced the σ_i^k values, $k = 1, 2, ..., s$, are, the larger the $I(\sigma_i)$ value.

As for a single cluster coefficient σ_i^k, the smaller its value is, the larger $\left(-\ln \sigma_i^k\right)$, and the larger influence $\left(-\ln \sigma_i^k\right)$ has on $I(\sigma_i)$. And, at the same time, the weight of $\left(-\ln \sigma_i^k\right)$ gets smaller. On the other hand, when the value of σ_i^k is relatively large, the value of $\left(-\ln \sigma_i^k\right)$ will be relatively small with a relatively large weight.

The entropy $I(\sigma_i)$ of the grey evaluation coefficient vector σ_i satisfies the following properties.

Property 6.7.1. *(Non-negativeness)*

$$I(\sigma_i) \geq 0. \tag{6.5}$$

Proof. 1. If there exists k' such that $\sigma_i^{k'} = 1$, from $\sigma_i^k \geq 0$ and $\sum_{k=1}^{s} \sigma_i^k = 1$, it follows that for any $k = 1, 2, ..., s$, when $k \neq k'$, $\sigma_i^k = 0$. Therefore,

$$I(\sigma_i) = -\sigma_i^{k'} \ln \sigma_i^{k'} = 0.$$

2. If for $k = 1, 2, ..., s$, $\sigma_i^k \neq 1$, then from $0 \leq \sigma_i^k < 1$, it follows that $\ln \sigma_i^k < 0$. So

$$I(\sigma_i) = -\sum_{k=1}^{s} \sigma_i^k \ln \sigma_i^k > 0. \square$$

In fact, $I(\sigma_i) = 0$ is the special situation where the result of the relevant grey evaluation is completely white. When some degree of greyness appears in a result of a grey evaluation, one must have $I(\sigma_i) > 0$.

Property 6.7.2. *(Symmetry) If*

$$\sigma_i = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^k, \sigma_i^{k+1}, ..., \sigma_i^s\right)$$

and

$$\sigma_i' = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^{k+1}, \sigma_i^k, ..., \sigma_i^s\right)$$

are two evaluation coefficient vectors with only the positions of the kth and $(k+1)$th components switched, then

$$I(\sigma_i) = I(\sigma_i').$$

Proof. It suffices to show that

$$-\sigma_i^k \ln \sigma_i^k - \sigma_i^{k+1} \ln \sigma_i^{k+1} = -\sigma_i^{k+1} \ln \sigma_i^{k+1} - \sigma_i^k \ln \sigma_i^k.$$

Based on the property of commutative property of real number addition, this end is obvious. □

This property states that the entropy of each grey evaluation coefficient vector σ_i is related to the values of the evaluation coefficients σ_i^1, σ_i^2, ..., σ_i^s and has nothing to do with the order in which these coefficients are arranged.

Property 6.7.3. *(Expendability) If*

$$\sigma_i = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^k, \sigma_i^{k+1}, ..., \sigma_i^s\right)$$

and

$$\sigma_i' = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^k, 0, \sigma_i^{k+1}, ..., \sigma_i^s\right)$$

are two grey evaluation coefficient vectors with a 0 component inserted at the kth location of σ_i, then

$$I(\sigma_i) = I(\sigma_i').$$

This property states that when a grey class with evaluation coefficient 0 is added, the entropy of the resultant grey evaluation coefficient vector will not be changed. This property can also be generalized to the following. For any $\sigma_i^k > 0$, if $\sigma_i' = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^k - \varepsilon, \varepsilon, \sigma_i^{k+1}, ..., \sigma_i^s\right)$, then

$$\lim_{\varepsilon \to 0^+} I(\sigma_i') = I(\sigma_i).$$

Property 6.7.4. *If $\sigma_i = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^s\right)$ and $\delta_i = \left(\delta_i^1, \delta_i^2, ..., \delta_i^s\right)$ are two grey evaluation coefficient vectors, then*

$$I(\sigma_i \cdot \delta_i) = I(\sigma_i) + I(\delta_i), \tag{6.6}$$

where

$$\begin{aligned}\sigma_i \cdot \delta_i \quad = \quad & \left(\sigma_i^1 \cdot \delta_i^1, \sigma_i^1 \cdot \delta_i^2, ..., \sigma_i^1 \cdot \delta_i^s,\right. \\ & \sigma_i^2 \cdot \delta_i^1, \sigma_i^2 \cdot \delta_i^2, ..., \sigma_i^2 \cdot \delta_i^s, \\ & \cdots\cdots \\ & \left.\sigma_i^s \cdot \delta_i^1, \sigma_i^s \cdot \delta_i^2, ..., \sigma_i^s \cdot \delta_i^s\right).\end{aligned}$$

Proof. From Definition 6.7.1, it follows that

$$\begin{aligned}I\left(\sigma_i \cdot \delta_i\right) &= -\sum_{j=1}^s \sum_{k=1}^s \sigma_i^j \cdot \delta_i^k \ln\left(\sigma_i^j \cdot \delta_i^k\right) \\ &= -\sum_{j=1}^s \sigma_i^j \left(\sum_{k=1}^s \delta_i^k \left(\ln \sigma_i^j + \ln \delta_i^k\right)\right) \\ &= -\sum_{j=1}^s \sigma_i^j \ln \sigma_i^j \sum_{k=1}^s \delta_i^k - \sum_{j=1}^s \sigma_i^j \sum_{k=1}^s \delta_i^k \ln \delta_i^k \\ &= -\sum_{j=1}^s \sigma_i^j \ln \sigma_i^j + \sum_{j=1}^s \sigma_i^j I\left(\delta_i\right) \\ &= I\left(\sigma_i\right) + I\left(\delta_i\right) \sum_{j=1}^s \sigma_i^j = I\left(\sigma_i\right) + I\left(\delta_i\right). \ \square\end{aligned}$$

Equation (6.6) indicates that the entropy of the product of two grey evaluation coefficient vectors equals the sum of the entropies of the individual vectors.

Property 6.7.5. *(Separability) Assume that*

$$\sigma_i = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^k + \sigma_i^{k+1}, ..., \sigma_i^s\right), \ \sigma_i' = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^k, \sigma_i^{k+1}, ..., \sigma_i^s\right)$$

are two grey evaluation coefficient vectors, where the kth entry $\left(\sigma_i^k + \sigma_i^{k+1}\right)$ of σ_i is split into two separate entries in σ_i' as the kth and the $(k+1)$th entries. Then, the following holds true,

$$I\left(\sigma_i'\right) = I\left(\sigma_i\right) - \left(\sigma_i^k + \sigma_i^{k+1}\right) I\left(\theta_i\right), \tag{6.7}$$

where

$$\theta_i = \left(\frac{\sigma_i^k}{\sigma_i^k + \sigma_i^{k+1}}, \frac{\sigma_i^{k+1}}{\sigma_i^k + \sigma_i^{k+1}}\right).$$

Proof. It suffices to show that

$$\begin{aligned}-\sigma_i^k \ln \sigma_i^k - \sigma_i^{k+1} \ln \sigma_i^{k+1} \quad &= -\left(\sigma_i^k + \sigma_i^{k+1}\right) \ln\left(\sigma_i^k + \sigma_i^{k+1}\right) \\ &\quad - \left(\sigma_i^k + \sigma_i^{k+1}\right) I\left(\theta_i\right)\end{aligned} \tag{6.8}$$

From

$$\left(\sigma_i^k + \sigma_i^{k+1}\right) I\left(\theta_i\right) = \left(\sigma_i^k + \sigma_i^{k+1}\right)\left(\frac{-\sigma_i^k}{\sigma_i^k + \sigma_i^{k+1}} \ln \frac{\sigma_i^k}{\sigma_i^k + \sigma_i^{k+1}} - \frac{\sigma_i^{k+1}}{\sigma_i^k + \sigma_i^{k+1}} \ln \frac{\sigma_i^{k+1}}{\sigma_i^k + \sigma_i^{k+1}}\right)$$

$$= -\sigma_i^k \ln \frac{\sigma_i^k}{\sigma_i^k + \sigma_i^{k+1}} - \sigma_i^{k+1} \ln \frac{\sigma_i^{k+1}}{\sigma_i^k + \sigma_i^{k+1}}$$

$$= -\sigma_i^k \ln \sigma_i^k - \sigma_i^{k+1} \ln \sigma_i^{k+1} + \sigma_i^k \ln\left(\sigma_i^k + \sigma_i^{k+1}\right) + \sigma_i^{k+1} \ln\left(\sigma_i^k + \sigma_i^{k+1}\right),$$

it follows that equ. (6.8) holds true. □

The separability property states that if grey clusters are further refined with the relevant grey evaluation coefficients separated, the entropy of the grey evaluation coefficient vector will increase.

Property 6.7.6. *(Maximum Valuation) When all the components of a grey evaluation coefficient vector* σ_i *are equal to the same value* $\frac{1}{s}$, *the entropy* $I(\sigma_i)$ *reaches its maximum value* $\ln s$. *That is,*

$$I\left(\sigma_i\right) \leq \ln s. \tag{6.9}$$

Proof. Because $I(\sigma_i) = -\sum_{k=1}^{s} \sigma_i^k \ln \sigma_i^k$, let $L = I(\sigma_i) + \lambda\left(1 - \sum_{k=1}^{s} \sigma_i^k\right)$. From $\frac{\partial L}{\partial \sigma_i^k} = -\ln \sigma_i^k - 1 - \lambda = 0$, $k = 1, 2, ..., s$, we can solve this system of equations and obtain $\sigma_i^1 = \sigma_i^2 = \cdots = \sigma_i^s =$ a constant $= \frac{1}{s}$. Therefore,

$$\max I\left(\sigma_i\right) = -\sum_{k=1}^{s} \frac{1}{s} \ln \frac{1}{s} = \frac{1}{s} \sum_{k=1}^{s} \ln s = \ln s. \ \square$$

Equation (6.9) indicates that when all components of σ_i equal the same value $\frac{1}{s}$, the entropy $I(\sigma_i)$ of σ_i reaches its maximum possible value $\ln s$.

From equs. (6.5) and (6.9) it follows that

$$0 \leq I\left(\sigma_i\right) \leq \ln s.$$

When $I(\sigma_i) = 0$, all hard results obtained from the grey evaluation have maximum reliability. In this case, the evaluation results contain the least uncertainty so that it is difficult for anyone to draw any soft result. When

$I(\sigma_i) \to 0$, the results of the grey evaluation contain a little uncertainty so that each hard result obtained from the evaluation can be seen with good reliability. At the same time, there is not sufficient evidence to support any strong soft result. When $I(\sigma_i) \to \ln s$, the results of the grey evaluation contain a relatively large amount of uncertainty so that the reliability of any hard result of the evaluation is not too good. At the same time, it is relatively easy for one to obtain soft results from the evaluation. When $I(\sigma_i) = \ln s$, the results of the grey evaluation contain the maximum amount of uncertainty. So, in this case, it is difficult to obtain any convincing hard result. On the other hand, the researcher can only draw soft results, if any.

6.8 Practical Examples

In this section, we look at three real-life projects in which we had the honor to be directly involved.

Example 6.8.1. In this example, we perform a grey evaluation analysis for the combined strength in the area of science and technology for Henan Province, The People's Republic of China.

With the start of reform in the infrastructure of science and technology in 1985, the operational mechanism and organizational structure of the system of science and technology in Henan Province have experienced a major change. The equipment of various scientific and technological resources has become more feasible. The overall provincial strength in science and technology has been constantly improving. The relationship between science and technology and economic and social development has been more mature and harmonic. In 1995, there were over 957,000 scholars and technicians in the province, representing an increase of more than 82.6% when compared to the year 1985. At the same time, the academic quality of the scientific team was obviously more advanced than before with much increased funding and a great many new scientific leaders appearing. In the year of 1995, the total funding allocated to the area of science and technology surpassed ¥2,900,000,000. During this ten-year period of time, over 12,000 major scientific and technological achievements and over 7500 patents had been recorded, which had helped and supported the healthy economic development of the province. According to relevant statistics, in the last five years, progress in science and technology has made over a 44% contribution to the overall economic growth of the province. Based on a combination of methods of Delphi correspondence and interviews, we have organized such an index system for evaluating regional strength in the area of science and technology such as shown in Table 6.3.

Table 6.3. Index system for evaluating regional strength in science and technology

Symbol	Weight	Weak Class	Medium Class	Strong Class
X_1	8	$2 \le x_1^1 < 20$	$20 \le x_1^2 < 70$	$70 \le x_1^3 < 110$
X_2	5	$80 \le x_2^1 < 150$	$150 \le x_2^2 < 250$	$250 \le x_2^3 < 500$
X_3	6	$5 \le x_3^1 < 12$	$12 \le x_3^2 < 20$	$20 \le x_3^3 < 30$
X_4	5	$3 \le x_4^1 < 6$	$6 \le x_4^2 < 10$	$10 \le x_4^3 < 15$
X_5	5	$2 \le x_5^1 < 5$	$5 \le x_5^2 < 8$	$8 \le x_5^3 < 12$
X_6	4	$1 \le x_6^1 < 2$	$2 \le x_6^2 < 5$	$5 \le x_6^3 < 12$
X_7	4	$5 \le x_7^1 < 10$	$10 \le x_7^2 < 20$	$20 \le x_7^3 < 40$
X_8	4	$0.5 \le x_8^1 < 2.5$	$2.5 \le x_8^2 < 5.5$	$5.5 \le x_8^3 < 9$
X_9	4	$8 \le x_9^1 < 12$	$12 \le x_9^2 < 20$	$20 \le x_9^3 < 50$
X_{10}	4	$30 \le x_{10}^1 < 100$	$100 \le x_{10}^2 < 200$	$200 \le x_{10}^3 < 350$
X_{11}	6	$2 \le x_{11}^1 < 4$	$4 \le x_{11}^2 < 7$	$7 \le x_{11}^3 < 15$
X_{12}	5	$8 \le x_{12}^1 < 15$	$15 \le x_{12}^1 < 25$	$25 \le x_{12}^1 < 50$
X_{13}	5	$4 \le x_{13}^1 < 6$	$6 \le x_{13}^2 < 9$	$9 \le x_{13}^3 < 11$
X_{14}	5	$0.5 \le x_{14}^1 < 1$	$1 \le x_{14}^2 < 2$	$2 \le x_{14}^3 < 4$
X_{15}	5	$3 \le x_{15}^1 < 5$	$5 \le x_{15}^2 < 8$	$8 \le x_{15}^3 < 15$
X_{16}	5	$2 \le x_{16}^1 < 5$	$5 \le x_{16}^2 < 10$	$10 \le x_{16}^3 < 20$
X_{17}	3	$5 \le x_{17}^1 < 20$	$20 \le x_{17}^3 < 60$	$60 \le x_{17}^3 < 120$
X_{18}	4	$5 \le x_{18}^1 < 8$	$8 \le x_{18}^2 < 12$	$12 \le x_{18}^3 < 20$
X_{19}	4	$2 \le x_{19}^1 < 5$	$5 \le x_{19}^2 < 8$	$8 \le x_{19}^3 < 11$
X_{20}	3	$3 \le x_{20}^1 < 6$	$6 \le x_{20}^2 < 9$	$9 \le x_{20}^3 < 15$
X_{21}	6	$25 \le x_{21}^1 < 35$	$35 \le x_{21}^2 < 45$	$45 \le x_{21}^3 < 55$

Here, all X_i, $i = 1, 2, \cdots, 21$, are defined as follows. For all input in terms of science and technology, we have

1. X_1 stands for the number of scientists and technicians with 10,000 as its unit,
2. X_2 the average number of scientists and technicians in each population of 10,000 people,
3. X_3 the concentration of engineers with % as its unit,
4. X_4 R&D spending/GDP with $^0/_{00}$ as its unit,
5. X_5 spending in science, technology, and applications with $^0/_{00}$ as its unit,
6. X_6 average funding available to the individual scientist and/or technician with ¥1,000 as its unit,
7. X_7 total value of equipment used in scientific research with ¥1 billion as its unit,

8. X_8 sale of scientific and technological books with 1 billion as its unit, and

9. X_9 the number of personal computers owned by each 10,000 residents.

For all input in terms of activities related to science and technology, we have

1. X_{10} number of existing agencies engaged in activities for scientific research and applications,

2. X_{11} number of ongoing scientific research projects with 1000 as its unit,

3. X_{12} number of people currently enrolled in a college per 10,000 residents, and

4. X_{13} average year of formal education of individual person in the work force.

In terms of products produced as a consequence of application of progress in science and technology, we have

1. X_{14} number of scientific achievements and patents with 1000 as its unit,

2. X_{15} number of research papers published in 1000 articles,

3. X_{16} amount in terms of money of commercial contracts assigned in the market of technology in ¥1 billion,

4. X_{17} amount of increase in terms of money in the area of manufacturing industry in ¥10 billion,

5. X_{18} concentration of high tech with % as its unit,

6. X_{19} amount of taxes collected compared to spending in industry,

7. X_{20} productivity with ¥1,000 as its unit, and

8. X_{21} percent of contribution made by progress in technology.

For the year of 1995, the materialized values of the criteria X_i, $i = 1, 2, ..., 21$, collected in Henan Province, are given in Table 6.4.

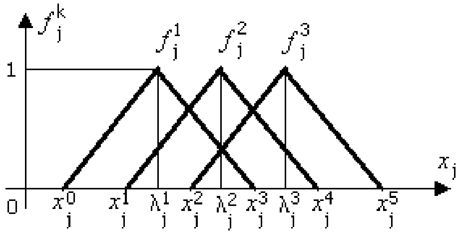

FIGURE 6.14. The general form of triangular whitenization weight functions

Table 6.4. Materialized values of the criteria X_i

Symbol	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Value	95.700	104.53	7.800	8	7	3,3500	18.6
Symbol	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}
Value	7.9	22	199	9.7030	13.4	8	2,2440
Symbol	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{21}
Value	10.1	12.88	722.97	9.4	10.450	6.7	43.5

Based on Table 6.3, the general form of triangular whitenization weight functions for the criteria X_i, $i = 1, 2, ..., 21$, is shown in Figure 6.14. Here x_j^0 and x_j^5 represent expanded values of the range of values of the criteria X_j, $j = 1, 2, ..., 21$. For an observational value x of criterion j, we can use equ. (6.10) to compute the whitenization weight function value $f_j^k(x)$ for the kth grey class, $k = 1, 2, 3$, as follows.

$$f_j^k(x) = \begin{cases} 0, & x \notin \left[x_j^{k-1}, x_j^{k+2}\right] \\ \dfrac{x - x_j^{k-1}}{\lambda_j^k - x_j^{k-1}}, & x \in \left[x_j^{k-1}, \lambda_j^k\right] \\ \dfrac{x_j^{k+2} - x}{x_j^{k+2} - \lambda_j^k}, & x \in \left[\lambda_j^k, x_j^{k+2}\right]. \end{cases} \tag{6.10}$$

For example, when $j = 1$, we expand the range of values for criterion X_1: the number of scientists and technicians, to $x_1^0 = 0.5$ and $x_1^5 = 160$. In this case, x_1^1, x_1^2, x_1^3, and x_1^4 are respectively the threshold values for the three grey classes: "weak", "medium", and "strong". That is, $x_1^1 = 2$, $x_1^2 = 20$, $x_1^3 = 70$, and $x_1^4 = 110$. Now, we let λ_1^k be the average value of x_1^k and

x_1^{k+1}. That is,

$$\lambda_1^1 = \frac{1}{2}\left(x_1^1 + x_1^2\right) = 11, \lambda_1^2 = \frac{1}{2}\left(x_1^2 + x_1^3\right) = 45, \lambda_1^3 = \frac{1}{2}\left(x_1^3 + x_1^4\right) = 90.$$

Now, substituting these specific values into equ. (6.10) leads to the following triangular whitenization weight functions for the case $j = 1$.

$$f_1^1(x) = \begin{cases} 0, & x \notin [0.5, 70] \\ \dfrac{x - 0.5}{11 - 0.5}, & x \in [0.5, 11] \\ \dfrac{70 - x}{70 - 11}, & x \in [11, 70], \end{cases} \tag{6.11}$$

$$f_1^2(x) = \begin{cases} 0, & x \notin [2, 110] \\ \dfrac{x - 2}{45 - 2}, & x \in [2, 45] \\ \dfrac{110 - x}{110 - 45}, & x \in [45, 110], \end{cases} \tag{6.12}$$

$$f_1^3(x) = \begin{cases} 0, & x \notin [20, 160] \\ \dfrac{x - 20}{90 - 20}, & x \in [20, 90] \\ \dfrac{160 - x}{160 - 90}, & x \in [90, 160]. \end{cases} \tag{6.13}$$

After substituting $x_1 = 95.70$, the value materialized in 1995 in Henan Province, into equs. (6.11) through (6.13), we obtain the whitenization weight values of the criterion X_1 in terms of the three grey classes "weak", "medium", and "strong" as follows:

$$f_1^1(95.70) = 0,\ f_1^2(95.70) = 0.22,\ f_1^3(95.70) = 0.919.$$

From these values, it can be seen that in terms of the number of scientists and technicians, Henan Province has entered the group of provinces with relatively strong scientific and technological strength. Table 6.5 lists all the expanded values for the criteria X_1 to X_{21} and Table 6.4 the relevant whitenization weight function values regarding the three grey classes: "weak", "medium", and "strong".

Table 6.5. Expanded values for the criteria X_1 to X_{21}

Symbol	X_1	X_2	X_3	X_4	X_5	X_6	X_7
x_j^0	0.5	40	2	1	1	300	2
x_j^5	160	800	40	20	18	20,000	60
Symbol	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}
x_j^0	0.20	3	10	11	3	3	300
x_j^5	12	80	500	20	80	12	6,000
Symbol	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{21}
x_j^0	1	0.5	20	3	1	1	15
x_j^5	20	30	1,800	30	15	20	65

Table 6.6. Whitenization weight function values regarding the three grey classes: "weak", "medium", and "strong"

Symbol	X_1	X_2	X_3	X_4	X_5	X_6	X_7
$f_j^1(x)$	0	0.860	0.892	0.636	0.222	0.471	0.112
$f_j^2(x)$	0.22	0.204	0.225	1	0.909	0.925	0.86
$f_j^3(x)$	0.919	0	0	0.308	0.4	0.208	0.43
Symbol	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}
$f_j^1(x)$	0	0	0.007	0	0.859	0.25	0
$f_j^2(x)$	0.22	0.824	0.755	0.558	0.45	0.857	0.702
$f_j^3(x)$	0.863	0.435	0.566	0.814	0	0.5	0.622
Symbol	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{21}
$f_j^1(x)$	0	0	0	0.473	0	0.511	0.233
$f_j^2(x)$	0.576	0.618	0.596	0.88	0.122	0.822	0.9
$f_j^3(x)$	0.785	0.728	0.747	0.175	0.827	0.117	0.433

From equ. (6.3), we can compute the cluster coefficients $\sigma_{HN}^k, k = 1, 2, 3,$ for the provincial strength of Henan in areas of science and technology in terms of the three grey classes: "weak", "medium", and "strong" as follows.

$$\sigma_{HN}^1 = \sum_{j=1}^{21} f_j^1(x_j) \cdot \eta_j = 26.67,$$

$$\sigma_{HN}^2 = \sum_{j=1}^{21} f_j^2(x_j) \cdot \eta_j = 61.261,$$

and

$$\sigma_{HN}^3 = \sum_{j=1}^{21} f_j^3(x_j) \cdot \eta_j = 48.157.$$

From the fact that $\max\left\{\sigma_{HN}^k : 1 \le k \le 3\right\} = 61.216 = \sigma_{HN}^2$, it can be seen that the comprehensive strength of Henan Province in areas of science and technology belongs to the grey class of "medium". However, the value of $\sigma_{HN}^3 = 48.157$ is relatively close to the value of σ_{HN}^2. We can see that the overall strength of Henan Province is getting near the grey class of "strong". This end discovery agrees well with the fact that in recent years, Henan Province has been recognized as being above the medium of the national level in terms of strength in areas of science and technology.

Table 6.6 also shows that the Henan Provincial strength has been mainly affected by such factors as $X_{2,}$ $X_{3,}$ X, $X_{12,}$ and $X_{20,}$ which are all about per capita averages, as well as such factors as $X_{2,}$ that is, the concentration of high tech. So, in order to strengthen its ranks when compared to other provinces and major cities in China, Henan Province needs to put its emphasis on improving its per capita averages. And, it needs to apply effective means to speed up the development of high tech in the province. Henan Province is the largest province in China in terms of population with a relatively weak foundation in such areas as high tech. In order to improve its per capita averages and develop high tech manufacturing industry, it will not be simply a matter of days. It will possibly take years to accomplish such a goal.

Example 6.8.2. In this example, we look at a topic such as evaluation criteria for regional signature industry. And, then we will look at the selection of signature industries for Wu Dou County.

Regional signature industry(ies) stands for a group of businesses that play a leading role in the development of the region's economy. These businesses determine the formation and evolution of the region's economic system. In terms of regional economic planning and healthy regional economic development, it is very critical to (1) correctly evaluate the levels of importance of various types of businesses located at the region of interest, (2) clearly determine which types of businesses are the key to the local economy, and (3) rightly depict the relationship between the key types of businesses and assistant and other types of businesses. On the other hand, the formation and evaluation of a region's key types of businesses are, to a certain degree, influenced by the region's economic structure, market characteristics, and resources. Therefore, encouraging the development of key types of businesses should be both the starting and the ending points of any activity in terms of policy making and regional economic planning. The ability to correctly define the key types of businesses will directly affect the economic growth of any region.

In general, a regional key type of business possesses the following main characteristics: (1) great potential for development with relatively high growth rate; (2) strong ability to take large market share; (3) representing a leader in the progress of technology; (4) having a wide range of influence with relatively strong forward and backward incidences. In our study, we employ these main characteristics as our criteria for selecting and determining regional key types of businesses. More specifically, we have the following quantitative criteria to determine a regional key type of business.

1. Income elasticity ε_i.

Assume that y stands for the per capita income, x_i the demand for the products or service of the ith type of business. Then, the income elasticity ε_i of the ith type of business is

$$\varepsilon_i = \frac{y}{x_i} \cdot \frac{\partial x_i}{\partial y}. \tag{6.14}$$

2. Growth rate r_i.

Assume that x_i^0 stands for the initial demand for the products of the ith type of business and r_i the average growth rate. Then, the demand for its products at time t is given by

$$x_i^t = x_i^0 \left(1 + r_i\right)^t. \tag{6.15}$$

3. Technological improvement η_i.

Assume that x_i stands for the total production value of the ith type of business, L_i its labor input, K_i the capital input, in the Cobb–Douglas production function

$$x_i = Ae^{bt} L_i^{\alpha} K_i^{\beta}, \tag{6.16}$$

b represents technological progress, and $\dfrac{\Delta x_i}{x_i}$ the production growth rate. Then,

$$\eta_i = \frac{b}{\dfrac{\Delta x_i}{x_i}} \tag{6.17}$$

describes the amount of contribution made by technological progress in the production growth of the ith type of business.

4. Degree of business incidence.

The degree of business incidence is divided into forward incidence and backward incidence. They all stand for the degree of direct or indirect impact caused by changes of the demand of one type of business on the input and output of other types of businesses. Such impact is also known as a wave effect. What is most used here are the concepts of coefficients of sensitivity and of impact. Assume that b_{ij} stands for the complete consumption coefficient in an input and output analysis and n the number of different types of businesses. Then, the coefficient of sensitivity of the ith type of business is given by

$$\mu_i = \frac{\sum_{j=1}^{n} b_{ij}}{\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}} \tag{6.18}$$

and the coefficient of impact of the jth type of business is given by

$$\nu_j = \frac{\sum_{i=1}^{n} b_{ij}}{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}}. \tag{6.19}$$

5. Coefficient of Comparative Advantage θ_i.

Each regional signature type of business should be such a group of developed enterprises that when compared to other regional businesses or when compared to its environment, it has its own business advantage. The relative advantage of the ith type of business can be expressed using the coefficient of comparative advantage θ_i. Here, this coefficient θ_i of comparative advantage can be written as the product of r_{i_1}, the coefficient of comparative concentration, r_{i_2}, the coefficient of comparative output, r_{i_3}, the coefficient of comparative productivity, and r_{i_4}, the coefficient of comparative tax rate. That is, we have

$$\theta_i = r_{i_1} \cdot r_{i_2} \cdot r_{i_3} \cdot r_{i_4}, \tag{6.20}$$

where

$$r_{i_1} = \frac{x_{k_i}/x_k}{x_i/x},\ r_{i_2} = \frac{o_{k_i}/o_k}{o_i/o},\ r_{i_1} = \frac{p_{k_i}/p_k}{p_i/p},\ r_{i_1} = \frac{T_{k_i}}{T_i}. \tag{6.21}$$

In equ. (6.21), the symbols $x_{k_i}, o_{k_i}, p_{k_i}$, and T_{k_i} represent, respectively, the production value, amount of production output, overall productivity, and the tax rate of the production of the ith type of business located in the kth region. The symbols x_k, o_k, and p_k respectively represent the total production value, production output, and average productivity of the kth region. And, x_i, o_i, p_i, and T_i respectively represent the production value, amount of production output, overall productivity, and the tax rate of the production of the ith type of business of a larger system within which the region of concern is located.

All the criteria discussed above are non-dimensional. In order to make our evaluation easier, we divide all regional business types into three categories based on their levels of importance in the economic development of their regions: general type, assistant type, and signature type. Then, through the use of the Delphi investigation method, we determine the domains of various categories and the weights of the criteria to be used in the selection of signature types of businesses. Our results are provided in Table 6.7.

Table 6.7. Criteria system for evaluating regional businesses

Symbol	Weight	General type	Assistant Type	Signature Type
ε_i	2	$\varepsilon_i < 1$	$1 \leq \varepsilon_i \leq 3$	$\varepsilon_i > 3$
r_i	2	$r_i < 0.9\%$	$9\% \leq r_i \leq 18\%$	$r_i > 18\%$
η_i	2	$\eta_i < 30\%$	$30\% \leq \eta_i \leq 50\%$	$\eta_i > 50\%$
μ_i	1.2	$\mu_i < 1$	$1 \leq \mu_i \leq 2$	$\mu_i > 2$
ν_i	1.2	$\nu_i < 1$	$1 \leq \nu_i \leq 1.2$	$\nu_i > 1.2$
θ_i	1.6	$\theta_i < 2$	$2 \leq \theta_i \leq 5$	$\theta_i > 5$

When practically determining regional signature types of businesses, other than the above standards and criteria, one often needs to develop a set of general principles for the selection of regional signature types of businesses based on the specific situations of the region of study. For example, one often has to consider such principles as sufficient employment, low energy cost, high addition values, advanced development of "bottle neck" businesses, and so on.

In the following, we focus on our selection of signature types of industrial businesses for Wu Dou County, the People's Republic of China.

Wu Dou County is located on the north shore of the Yellow River and is a typical agricultural county. In order for the county to compete in the coming age commercially, it needs to push strongly for development in modern industries. To this end, the first problem facing county officials is to determine the types of signature businesses. Based on the available data collected at Wu Dou County, we have produced Table 6.8.

Table 6.8. Criteria evaluation data for Wu Dou County

Business Type	ε_i	r_i	η_i	μ_i	ν_i	θ_i
BT 1	3.03	0.41	0.42	1.07	1.13	5.19
BT 2	1.63	0.22	0.53	2.85	1.16	1.82
BT 3	0.12	0.23	0.39	1.28	1.07	5.55
BT 4	3.99	0.37	0.53	6.89	1.20	5.13
BT 5	7.87	0.33	0.47	0.90	0.98	6.42
BT 6	3.86	0.24	0.52	0.78	0.81	4.82
BT 7	1.97	0.24	0.34	1.30	1.02	3.00
BT 8	1.85	0.21	0.32	1.06	1.10	2.43
BT 9	1.63	0.21	0.44	1.17	0.87	0.11
BT 10	1.05	0.16	0.37	0.88	0.90	2.26
BT 11	3.01	0.23	0.21	0.79	0.82	1.03
BT 12	1.69	0.21	0.42	1.27	1.05	1.19
BT 13	0.00	0.00	0.00	1.04	1.08	1.04
BT 14	0.00	0.00	0.00	0.66	1.15	0.55
BT 15	1.74	0.21	0.31	0.98	0.53	2.06

Business types are defined as follows.

- BT 1 stands for paper manufacturing and paper related industries,
- BT 2 for chemical industry,
- BT 3 for textile,
- BT 4 for leather,
- BT 5 for food and drinks,
- BT 6 for pharmaceutical,
- BT 7 for machinery,
- BT 8 for metallurgy,
- BT 9 for electricity,
- BT 10 for rubber,
- BT 11 for plastic,
- BT 12 for non-metal,
- BT 13 for electronics,
- BT 14 for instruments and meters, and
- BT 15 for water treatment.

Based on Tables 6.7 and 6.8 and the grey fixed weight clustering method, we obtain Table 6.9 listing all the evaluation results for the existing industries located in Wu Dou County.

Table 6.9. Evaluation results for existing industries in Wu Dou County

Business Type	General Type σ_i^1	Assistant Type σ_i^2	Signature Type σ_i^3
BT 1	1.032	2.608	6.360
BT 2	2.340	1.740	5.920
BT 3	3.088	3.312	3.600
BT 4	0.000	0.000	10.000
BT 5	2.400	0.600	7.000
BT 6	2.400	1.912	7.408
BT 7	3.233	4.767	2.000
BT 8	4.097	3.903	2.000
BT 9	4.332	2.868	2.800
BT 10	6.223	2.666	1.111
BT 11	6.000	0.000	4.000
BT 12	3.372	4.228	2.400
BT 13	8.944	1.056	0.000
BT 14	8.800	0.600	0.600
BT 15	6.256	1.744	2.000

From

$$\max_{1\le k\le 3}\left\{\sigma_1^k\right\}=6.36=\sigma_1^3,\ \max_{1\le k\le 3}\left\{\sigma_2^k\right\}=5.92=\sigma_2^3,\ \max_{1\le k\le 3}\left\{\sigma_3^k\right\}=3.60=\sigma_3^3,$$

$$\max_{1\le k\le 3}\left\{\sigma_4^k\right\}=10.0=\sigma_4^3,\ \max_{1\le k\le 3}\left\{\sigma_5^k\right\}=7.00=\sigma_5^3,\ \max_{1\le k\le 3}\left\{\sigma_6^k\right\}=7.41=\sigma_6^3,$$

$$\max_{1\le k\le 3}\left\{\sigma_7^k\right\}=4.77=\sigma_7^2,\ \max_{1\le k\le 3}\left\{\sigma_8^k\right\}=4.10=\sigma_8^1,\ \max_{1\le k\le 3}\left\{\sigma_9^k\right\}=4.33=\sigma_9^1,$$

$$\max_{1\le k\le 3}\left\{\sigma_{10}^k\right\}=6.22=\sigma_{10}^1,\ \max_{1\le k\le 3}\left\{\sigma_{11}^k\right\}=6=\sigma_{11}^1,$$

$$\max_{1\le k\le 3}\left\{\sigma_{12}^k\right\}=4.23=\sigma_{12}^2,\ \max_{1\le k\le 3}\left\{\sigma_{13}^k\right\}=8.94=\sigma_{13}^1,$$

$$\max_{1\le k\le 3}\left\{\sigma_{14}^k\right\}=8.8=\sigma_{14}^1,\ \max_{1\le k\le 3}\left\{\sigma_{15}^k\right\}=6.26=\sigma_{15}^1,$$

it follows that BT 1, BT 2, BT 3, BT 4, BT 5, and BT 6 are leading industries, and BT 7 and BT 12 are assistant industries with BT 8 to 11 and BT 13 to 15 being general types of businesses. Our research results here have provided the scientific basis for the economic planning of Wu Dou County at the time when it was already to welcome the arrival of the new millennium.

Example 6.8.3. Let us consider a grey evaluation analysis we did for Henan Province in 1993.

Since the start of the decade of the1990s, with an increased speed of economic reforms and rapid development of market economy, Henan Province entered a new period of time for growth. During the past several years, the relationship between regional economic systems and external environment, that between regional economic systems and internal subsystems, and that between various economic cells and the economic mechanism from the microscopic level to the macroscopic level gradually reached a relatively harmonic stability. In 1992, the GDP of the province reached a total of ¥1,207 billion, representing an increase of 13.6% over the previous year. Such a GDP level had materialized to a per capita income of ¥1,362, a 12.2% increase over the previous year.

In our study, we classified regional economic development into three categories: basic, comfortable, and wealthy. In fact, the development stages of the so-called "basic", "comfortable", and "wealthy" are all grey concepts. And, the relevant determination of categories is also grey. In order to reasonably determine the domain and weight of each criterion in terms of the categories to be employed in our evaluation, we organized three rounds of Delphi investigation. Based on experts' recommendations, we produced 16 regional economic evaluation criteria as listed in Table 6.10. And, Table 6.11 lists all the materialized criteria values for the year 1992 as collected in Henan Province.

Table 6.10. Criteria system for evaluating a regional economy

Symbol	Weight	Basic	Comfortable	Wealthy
E_1	8	[500, 1,500)	[1,500, 3,000)	[3,000, 7,000)
E_2	6	[900, 2,600)	[2,600, 5,400)	[5,400, 10,000)
E_3	6	[12, 19)	[19, 24)	[24, 30)
E_4	4	(5, 9]	(2.5, 5]	(1, 2.5]
E_5	6	[320, 370)	[370, 400)	[400, 450)
ST_1	7	[40, 60)	[60, 75)	[75, 90)
ST_2	7	[1, 1.8)	[1.8, 2.5)	[2.5, 3.6)
ST_3	7	[25, 40)	[40, 55)	[55, 70)
S_1	7	[35, 50)	[50, 75)	[75, 95)
S_2	7	[20, 30)	[30, 45)	[45, 60)
S_3	6	[25, 40)	[40, 50)	[50, 60)
S_4	5	[20, 35)	[35, 50)	[50, 60)
LS_1	6	(50, 60]	(40, 50]	(25, 40]
LS_2	6	[4, 8)	[8, 11)	[11, 14)
LS_3	5	[15, 20)	[20, 22)	[22, 25)
LS_4	7	[60, 65)	[65, 70)	[70, 75)

Five economic criteria, three criteria reflecting science and technology, four criteria about social structure, and four criteria on life quality are defined as follows.

1. E_1 stands for per capita GDP value with the yuan as its unit,
2. E_2 for productivity in yuan per person per year,
3. E_3 for tax rate per ¥10,000 industrial capital in the unit of yuan,
4. E_4 for energy cost per ¥10,000 industrial output in tons, and
5. E_5 for food occupation per person in kilograms;
6. ST_1 represents percent of adult literacy,
7. ST_2 percent of scientific workers and technicians in the labor force, and
8. ST_3 percent of contribution made by progress in science and technology;
9. S_1 indicates the percent of automation;
10. S_2 percent of informationalization,
11. S_3 percent of non-farming labor in the totality of the labor force, and

12. S_4 percent of concentration of city residents;

13. LS_1 is the Engal coefficient,

14. LS_2 average living space size of city residents,

15. LS_3 average living space size of country residents, and

16. LS_4 average life span of all residents.

In Table 6.10, the intervals for the categories "basic," "comfortable", and "wealthy" are used as follows; for example, when the materialized value $E_1(\otimes)$ of criterion E_1 falls in between [500, 1500), we talk about achieving a basic level of life. When the materialized value $E_1(\otimes)$ falls in the interval [1500, 3000), we talk about enjoying a comfortable life. When the materialized value $E_1(\otimes)$ falls in between 3000 and 7000, we have reached a level of wealthy lifestyle.

Table 6.11. Materialized economic values of Henan Province in 1992

Symbol	E_1	E_2	E_3	E_4	E_5	ST_1	ST_2	ST_3
Value	1,362	2,791	8.700	3.500	351.0	67.00	1.900	28.00
Symbol	S_1	S_2	S_3	S_4	LS_1	LS_2	LS_3	LS_4
Value	74.00	25.00	26.00	38.00	55.00	11.00	16.90	70.00

Based on the grey classes as defined in Table 6.10, the general form of triangular membership functions is given as in equ. (6.10) and shown in Figure 6.14. For example, for the case of $j = 1$, we expand the range of values of our grey number to $x_1^0 = 200$ and $x_1^5 = 10,000$. From $x_1^1 = 500$, $x_1^2 = 1500$, $x_1^3 = 3000$, and $x_1^4 = 7000$, it follows that

$$\begin{aligned}\lambda_1^1 &= \left(x_1^1 + x_1^2\right)/2 = 1000,\\ \lambda_1^2 &= \left(x_1^2 + x_1^3\right)/2 = 2250,\\ \lambda_1^3 &= \left(x_1^3 + x_1^4\right)/2 = 5000.\end{aligned}$$

For the per capita GDP $E_1 = 1362$ for the year 1992, we compute its membership in the class of "basic" as

$$f_1^1(1362) = \frac{(3000 - 1362)}{2000} = 0.819,$$

its membership in the class of "comfortable" as

$$f_1^2(1362) = 0.493$$

and its membership in the class of "comfortable" as

$$f_1^3(1362) = 0.$$

Therefore, our results indicate that in terms of per capita GDP, Henan Province still belongs to the class of "basic".

Table 6.12 lists all the extended values for each criterion and Table 6.13 provides the membership values for each criterion to belong to one grey class.

Table 6.12. Expanded fields of evaluation criteria

Symbol	E_1	E_2	E_3	E_4	E_5	ST_1	ST_2	ST_3
x_j^0	200	400	6	0.7	260	25	0.5	15
x_j^5	10,000	15,000	35	12	500	98	4.5	80
Symbol	S_1	S_2	S_3	S_4	LS_1	LS_2	LS_3	LS_4
x_j^0	20	10	15	10	20	2.5	10	15
x_j^5	97	70	70	80	70	16	30	80

Table 6.13. Memberships of main economic criteria

Symbol	E_1	E_2	E_3	E_4	E_5	ST_1	ST_2	ST_3
$f_j^1(\cdot)$	0.819	0.715	0.284	0.222	0.891	0.68	0.455	0.743
$f_j^2(\cdot)$	0.493	0.610	0	0.909	0.477	0.982	0.783	0.133
$f_j^3(\cdot)$	0	0.037	0	0.538	0	0.311	0.080	0
Symbol	S_1	S_2	S_3	S_4	LS_1	LS_2	LS_3	LS_4
$f_j^1(\cdot)$	0.031	1	0.629	0.467	1	0	0.92	0
$f_j^2(\cdot)$	0.646	0.286	0.050	0.8	0.333	0.667	0.317	0.667
$f_j^3(\cdot)$	0.686	0	0	0.133	0	0.667	0	0.667

Therefore, we have

$$\sigma_i^1 = \sum_{j=1}^{16} f_j^1(x_{ij}) \cdot \eta_j = 55.825,$$

$$\sigma_i^2 = \sum_{j=1}^{16} f_j^2(x_{ij}) \cdot \eta_j = 50.466$$

and

$$\sigma_i^3 = \sum_{j=1}^{16} f_j^3(x_{ij}) \cdot \eta_j = 19.249.$$

From $\max\left\{\sigma_i^k\colon 1 \le k \le 3\right\} = 55.825 = \sigma_1^1$, we conclude that Henan Province still belonged to the class of "basic" in 1992. However, because $\sigma_i^2 = 50.466$ is close to σ_i^1, we can say that Henan Province was on the edge of entering the class of "comfortable". So, with some effort, it would be possible for Henan Province to actually enter the class of "comfortable" by the end of the 20th century.

From Table 6.13, it can be seen that among the five economic criteria, the main problem was the low economic value of industrial output. Such a low value was still a large distance from the economy of the "basic" class, and any sustained increase in the industrial economy could only be maintained through high levels of input of both capital and labor. If this problem was not resolved in time, it would definitely cause serious effects on the economic health of the province. On the other hand, the situation with food production was not too good, either. Henan agricultural business had a relatively weak level of resistance against natural disasters. Therefore, mechanisms needed to be introduced in order to prevent and reduce the number and severity of natural disasters in order to help a stable and healthy development in agriculture.

Among the criteria of science and technology, science and technology did not contribute a high percentage to the economic growth. This problem of low contribution by science and technology was mainly caused by the blind emphasis on increased input of capital and labor without valuing the input and commercial effect of science and technology. The effect of science and technology being the first labor force was not effectively materialized. Among the criteria of social structure, two problems stood out. One was the low level of informationalization of the society and the other the low percentage of non-farming labor. With a large amount of labor restrained on the limited amount of farmland, its power and economic effect could not be released. This had been a chorionic problem hindering economic development in Henan Province. By sufficiently applying market and policy mechanisms, excess farm labor could be guided to other commercial endeavors so that these two problems could be resolved at the same time.

Food cost occupied as high as 55% of the total living expense. This was a relatively standard level for being in the "basic" class. It also explained the fact that most people still could not correctly allocate their spending funds among other areas of life, such as clothing, beautification of their living environment, entertainment, etc. So, it was necessary to help guide citizens to improve their consuming habits and tastes. However, in order to achieve a successful outcome along this line, a fundamental increase in the societal economic levels and standards were necessary.

7
Grey Systems Modeling

7.1 The Thought of Five-Step-Modeling

When studying an abstract system, establishing a mathematical model for the system is a quantitative study of the system with respect to its overall functions, synergic functions, incidence relationships among its factors, causal relationships, and dynamical relationships, etc. In this kind of study, a qualitative analysis must be first, and the relevant quantitative analysis must be closely related to the initial qualitative analysis. Therefore, in general, the establishment of a system's model needs to go through five steps: development of thoughts, factor analysis, quantification, dynamicalization, and optimization. This is the so-called *five-step modeling*. More specifically,

Step 1: Develop thoughts and form concepts. Through an initial qualitative analysis and research, one can clarify the research direction, goals, paths, and how to implement all the foreseeable details. Then, verbally and precisely describe the results. This is called a language modeling.

Step 2: Examine factors involved and any relationship between the factors contained in the language model to find causes and effects affecting the development of the system under consideration. Then, construct a diagram to depict the causal relation. See, for example, Figure 7.1.

Each pair of cause and effect (or a group of pairs of causes and effects) constitutes a link. Each system consists of several or many such links. Sometimes, a quantity is both a cause of a link and an effect of another link. When all such links are connected, one obtains a connected diagram

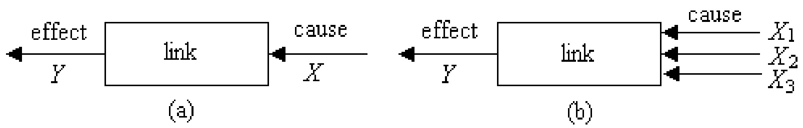

FIGURE 7.1. Flow diagram for a causal relation

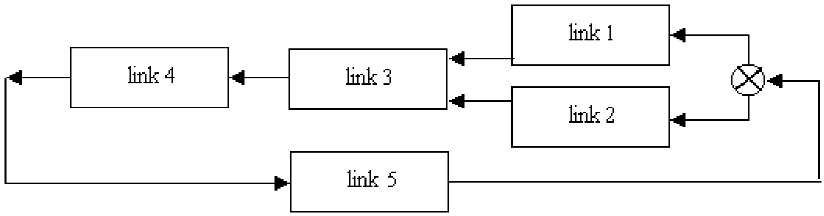

FIGURE 7.2. Connected diagram with several links

with several links, called network modeling. See Figure 7.2 for more details.

Step 3: Analyze quantitatively the causal relationship of each link and initially derive conceptual and quantified relations at some low level. This step is called a quantification modeling.

Step 4: Collect input and output data values of each link to establish a dynamic grey systems model (GM). This step is called dynamic modeling. Dynamic modeling is a high-level quantification modeling. It deeply reveals an existing quantitative relation or governing law of transformation from the input to the output. It is the foundation for systems analysis and optimization.

Step 5: Systemically study the dynamic model obtained in Step 4. Through adjustments on organizations, mechanisms, and parameters, reorganize the system in order to optimize allocations so that the goal of improving the system's dynamics can be reached. This way of constructing a model is called optimal modeling.

The entire process of the five-step-modeling is a process of establishing five models through five different steps:

$$\text{language modeling} \Rightarrow \text{network modeling} \Rightarrow \text{quantification modeling}$$
$$\Rightarrow \text{dynamic modeling} \Rightarrow \text{optimal modeling.}$$

During the process of establishing a model, one needs to constantly feed back the results obtained in later steps to earlier steps. After many iterations of such feedback, the entire modeling will become more mature and more adequate.

The concept of five-step-modeling establishes a bridge connecting the social and natural sciences, which mathematizes, computerizes, and natural-scientificalizes research in social science. At the same time, it brings the research of natural science to a higher level with more abstraction, precision, and philosophy.

Now, the fundamental idea of grey systems modeling can be summarized as follows.

1. A qualitative analysis is a prerequisite for modeling.
2. Quantitative modeling is a specialization of the initial qualitative analysis.
3. Qualitative and quantitative analyses are combined closely to complement each other.
4. Clarifying the system's factors, relationship between these factors, and relationship between the factors and the system is the center of the research of the system.
5. The factor analysis should not dwell on one state. Instead, it should involve the movement of time and change of states. That is, the research on system behaviors should be dynamicalized.
6. Any relationship between factors and between the factors and the system is relative, not absolute.
7. In order to generalize the effective methods and successful results in control theory to areas such as social science, economics, agriculture, ecology, etc., systems modeling needs to be controlled.
8. Through models to understand all properties of control of the system, such as controllability, observability, etc.
9. Through models to diagnose the system under consideration in order to clarify its current state and hidden problems.
10. From models, one should obtain as much information as possible, especially about development and change, such as information on whether the development of the system is permanent or limited; as for permanently developing systems, whether it is developing monotonically or vibrantly, whether the development is rapid or gradual; as for systems with limited development, what its limit is, whether it reaches its limit monotonically or wavelike, whether it reaches its limit quickly or gradually, and whether there exist shocks in the development.

11. The types of data often used in model building are given in the following,

    (1) data of scientific experiments; (2) empirical data;

    (3) production data; (4) decision-making data.

12. The fundamental data for grey modeling are sequence generations.
13. For sequences satisfying the conditions for being quasi-smooth, one can establish GM differential models. General non-negative sequences can become quasi-smooth if the accumulating generation is applied.
14. The accuracy of models can be improved through different methods of generation of grey numbers, choice of data values, reorganization or modification of sequences, and supplements with various levels of remnant GM models.
15. Grey systems theory employs three different methods to check and to determine model accuracies:
    (a) Remnant check. It does point-to-point checks on the accuracy between model values and values collected in practice;
    (b) Incidence degree check. This is done through a comparison between the model value curve and the sequence curve on which the model was built;
    (c) Remnant distribution check. It is done through the use of statistical characteristics of the remnant distributions.

7.2 Grey Differential Equations

Many scientific workers in systems research are interested in differential equations with the conviction that differential equations can deeply describe the essence of development of things. When faced with discrete data sequences, people often feel some difficulties, because only when the condition of differentiability is assumed, the concept of differential equations can be talked about. In grey systems theory, based on understanding of differential and integral calculus, the concept of grey derivatives is introduced so that we can establish models similar to differential equations for sequences of discrete data.

Definition 7.2.1. *Assume that we have the following differential equation,*

$$\frac{dx}{dt} + ax = b.$$

Then $\frac{dx}{dt}$ *is called the derivative of the unknown function* x, x *the background value of* $\frac{dx}{dt}$, *and* a *and* b *the parameters.*

That is, a first-order differential equation consists of three parts: derivative, background value, and some parameters.

Definition 7.2.2. *Assume that* $x(t)$ *is a function defined on a time set* T. *If when* $\Delta t \to 0$, *it is always true that*

$$x(t + \Delta t) - x(t) \neq 0,$$

then we say that the information density of $x(t)$ *on* T *is infinite.*

Proposition 7.2.1. *A function* $x(t)$ *satisfying the differential equation*

$$\frac{dx}{dt} + ax = b$$

satisfies the condition of infinite information density.

Proof. From the definition of derivatives

$$\frac{dx}{dt} = \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t}$$

the result follows. □

Definition 7.2.3. *Assume that* A *and* B *are sets, and* R *an operation between* A *and* B. *If for* a_1, $a_2 \in A$ *and* $b \in B$, *the following holds,*

$$a_1 R b = a_2 R b,$$

then b *is called a horizontal mapping of* a_1 *and* a_2.

Definition 7.2.4. *Assume that* R *is the operator of absolute difference. That is,*

$$aRb = |a - b|.$$

If $a_1 R b = a_2 R b$, *that is,* $|a_1 - b| = |a_2 - b|$, *for any* a_1, $a_2 \in A$, *and a fixed* $b \in B$, *then the operator* R *is called an arithmetic horizontal mapping or simple horizontal mapping.*

Definition 7.2.5. *Let*

$$\frac{dx}{dt} + ax = b,$$

$x(t + \Delta t)$ *and* $x(t)$ *be two elements in the background set, and*

$$X = \{x(t + \Delta t), x(t)\}.$$

1. *When*

$$\frac{dx}{dt} R x(t + \Delta t) = \frac{dx}{dt} R x(t),$$

the derivative and the background elements are said to satisfy the horizontal mapping relation.

2. If x is a value taken as a background value, satisfying

$$x(t) \neq x \neq x(t+\Delta t), x(t), x(t+\Delta t) \in X,$$

let $\delta(t+\Delta t)$ and $\delta(t)$ be components of

$$\frac{dx}{dt} \cong \frac{x(t+\Delta t)-x(t)}{\Delta t}.$$

When

$$\delta(t+\Delta t)Rx = \delta(t)Rx$$

we say that the background value and derivative components satisfy the horizontal mapping relation.

Proposition 7.2.2. *If $x(t)$ is a positive function, that is, for any t, $x(t) > 0$, then the derivative in the equation*

$$\frac{dx}{dt} + ax = b$$

and elements in the background set satisfy the simple horizontal mapping relation.

Theorem 7.2.1. *There are three fundamental conditions to form a differential equation:*

1. Information density is infinite;

2. Background values are a grey number;

3. The derivative and the background values satisfy the horizontal mapping relation.

Definition 7.2.6. *Assume that I is a set of units to measure time. If*

$$I = \{..., \text{year, month, day, hour, minute, second, } ... \}$$

then I is called the set of general time units.

Definition 7.2.7. *Assume that 1_i and 1_j are unit times of the ith level and the jth level in the set of general time units, respectively. If $1_i < 1_j$, then we say that the ith level time is denser than the jth level time.*

Definition 7.2.8. *Assume that*

$$X = (x(1_i), x(2_i), ..., x(n_i))$$

is a sequence of time units at the ith level. Then

$$d^{(i)} = x(k_i) - x(k_i - 1_i),$$

$k_i = 1_i, 2_i, ..., n_i$, *are called the information increments of the time units at the ith level.*

Definition 7.2.9. *Assume that X is a time sequence whose time unit can be infinitely divided, and 1_i a unit time at the ith level of time units. If when $1_i \to 0$,*

$$d^{(i)} = x(k_i) - x(k_i - 1_i) \neq 0.$$

Then X is said to be a sequence with the intension of differential equations, or a grey differential sequence, and

$$d^{(i)}(k_i) = \lim_{1_i \to 0} [x(k_i) - x(k_i - 1_i)],$$

is called the grey derivative of the sequence X. The grey derivative of a general sequence is written as $d(k)$.

Proposition 7.2.3. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence of raw data values and

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\right)$$

the sequence obtained through accumulating generation; that is,

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i),$$

k *= 1, 2, ..., n. Then the grey derivative of $X^{(1)}$ is*

$$d(k) \;=\; x^{(0)}(k).$$

Proof. Because the time unit is not divided into smaller units, we have

$$\begin{aligned} d(k) &= x^{(1)}(k) - x^{(1)}(k-1) \\ &= \sum_{i=1}^{k} x^{(0)}(i) - \sum_{i=1}^{k-1} x^{(0)}(i) = x^{(0)}(k). \; \square \end{aligned}$$

7.3 Model: GM(1,1)

In this section, we study the so-called GM(1,1) model. This model has been widely applied in various areas of practical applications of grey systems theory.

Definition 7.3.1. *The following*

$$d^{(i)}(k_i) + ax^{(1)}(k_i) = b$$

is called an equation of grey differential type.

Proposition 7.3.1. *For the following equation of grey differential type*

$$x^{(0)}(k) + ax^{(1)}(k) = b$$

the grey derivative $x^{(0)}(k)$ *and elements in the set of background values*

$$\left\{x^{(1)}(k), x^{(1)}(k-1)\right\}$$

do not satisfy the horizontal mapping relation.

Proof.

$$|x^{(0)}(k) - x^{(1)}(k)| = |-x^{(1)}(k-1)| = x^{(1)}(k-1),$$

but

$$|x^{(0)}(k) - x^{(1)}(k-1)| = |2x^{(0)}(k) - x^{(1)}(k)|,$$

so,

$$|x^{(0)}(k) - x^{(1)}(k)| \neq |x^{(0)}(k) - x^{(1)}(k-1)|. \square$$

Proposition 7.3.2. *If the background value is taken to be the mean of the entries in* $X^{(1)}$*, that is, let*

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1),$$

then the background value $z^{(1)}(k)$ *and the components* $x^{(1)}(k)$ *and* $x^{(1)}(k-1)$ *of the grey derivative satisfy the arithmetic horizontal mapping relation.*

Proof.

$$\begin{aligned} |z^{(1)}(k) - x^{(1)}(k)| &= |0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) - x^{(1)}(k)| \\ &= |0.5x^{(1)}(k-1) - 0.5x^{(1)}(k)| \end{aligned}$$

and

$$\begin{aligned} |z^{(1)}(k) - x^{(1)}(k-1)| &= |0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) - x^{(1)}(k-1)| \\ &= |0.5x^{(1)}(k) - 0.5x^{(1)}(k-1)|. \end{aligned}$$

Therefore,

$$z^{(1)}(k)Rx^{(1)}(k) = z^{(1)}(k)Rx^{(1)}(k-1). \square$$

Definition 7.3.2. *If an equation of grey differential type satisfies the following conditions,*

1. the information density is infinitely large;

2. the sequence possesses the intension of grey differentiation; and

3. the mapping from the set of background values to the components of the grey derivative satisfy the horizontal mapping relation,

then this equation of grey differential type is called a grey $differential$ $equation$.

Proposition 7.3.3. *The following equation*

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

where

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$$

is a grey differential equation.

Definition 7.3.3. *The equation*

$$x^{(0)}(k) + az^{(1)}(k) = b$$

is called a $GM(1,1)$ $model$.

The meaning of the symbol $GM(1,1)$ is given as follows:

G	M	$(1,$	$1)$
$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
Grey	Model	First Order	One Variable

Theorem 7.3.1. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a non-negative sequence, where $x^{(0)}(k) \geq 0, k = 1, 2, ..., n$, $X^{(1)}$ the 1-AGO sequence of $X^{(0)}$ with

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\right),$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i),$$

$k = 1, 2, ..., n$, and $Z^{(1)}$ is the mean generated sequence of consecutive neighbers of $X^{(1)}$ given by

$$Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)\right),$$

where

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1),$$

$k = 1, 2, ..., n$. *If* $\hat{a} = [a, b]^T$ *is a sequence of parameters, and*

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \cdot \\ \cdot \\ \cdot \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ -z^{(1)}(n) & 1 \end{bmatrix},$$

then the least squares estimate sequence of the grey differential equation

$$x^{(0)}(k) + az^{(1)}(k) = b$$

satisfies

$$\hat{a} = \left[B^T B\right]^{-1} B^T Y.$$

Proof. Substituting all data values into the grey differential equation

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

gives that

$$x^{(0)}(2) + az^{(1)}(2) = b,$$

$$x^{(0)}(3) + az^{(1)}(3) = b,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x^{(0)}(n) + az^{(1)}(n) = b.$$

That is,

$$Y = B\hat{a}.$$

For a pair of evaluated values of a and b, using -$az^{(1)}(k) + b$ to substitute $x^{(0)}(k)$, k = 2, 3, ..., n, gives the error sequence

$$\varepsilon = Y - B\hat{a}.$$

Let

$$\begin{aligned} s &= \varepsilon^T \varepsilon = \left[Y - B\hat{a}\right]^T \left[Y - B\hat{a}\right] \\ &= \sum_{k=2}^{n} [x^{(0)}(k) + az^{(1)}(k) - b]^2. \end{aligned}$$

The a and b values making s the minimum should satisfy

$$\begin{cases} \dfrac{\partial s}{\partial a} = 2\sum_{k=2}^{n}[x^{(0)}(k) + az^{(1)}(k) - b] \cdot z^{(1)}(k) = 0 \\ \dfrac{\partial s}{\partial b} = -2\sum_{k=2}^{n}[x^{(0)}(k) + az^{(1)}(k) - b] = 0. \end{cases}$$

That is,

$$\begin{cases} \sum_{k=2}^{n}\{x^{(0)}(k)z^{(1)}(k) + a[z^{(1)}(k)]^2 - b \cdot z^{(1)}(k)\} = 0 \\ \sum_{k=2}^{n}[x^{(0)}(k) + az^{(1)}(k) - b] = 0. \end{cases}$$

So, solving this system gives that

$$b = \frac{1}{n-1}\left[\sum_{k=2}^{n} x^{(0)}(k) + a\sum_{k=2}^{n} z^{(1)}(k)\right]$$

and

$$a = \frac{\frac{1}{n-1}\sum_{k=2}^{n} x^{(0)}(k) \cdot \sum_{k=2}^{n} z^{(1)}(k) - \sum_{k=2}^{n} x^{(0)}(k) \cdot z^{(1)}(k)}{\sum_{k=2}^{n}[z^{(1)}(k)]^2 - \frac{1}{n-1}\left[\sum_{k=2}^{n} z^{(1)}(k)\right]^2}.$$

From $Y = B\hat{a}$, it follows that

$$B^T B\hat{a} = B^T Y,\ \hat{a} = \left[B^T B\right]^{-1} B^T Y.$$

But

$$B^T B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ -z^{(1)}(n) & 1 \end{bmatrix}^T \cdot \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ -z^{(1)}(n) & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=2}^{n}[z^{(1)}(k)]^2 & -\sum_{k=2}^{n} z^{(1)}(k) \\ -\sum_{k=2}^{n} z^{(1)}(k) & n-1 \end{bmatrix},$$

$$\left[B^T B\right]^{-1} = \frac{\begin{bmatrix} n-1 & \sum_{k=2}^{n} z^{(1)}(k) \\ \sum_{k=2}^{n} z^{(1)}(k) & \sum_{k=2}^{n} [z^{(1)}(k)]^2 \end{bmatrix}}{(n-1)\sum_{k=2}^{n} [z^{(1)}(k)]^2 - \left[\sum_{k=2}^{n} z^{(1)}(k)\right]^2},$$

and

$$B^T Y = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ -z^{(1)}(n) & 1 \end{bmatrix}^T \cdot \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \cdot \\ \cdot \\ \cdot \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -\sum_{k=2}^{n} x^{(0)}(k) \cdot z^{(1)}(k) \\ \sum_{k=2}^{n} x^{(0)}(k) \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \hat{a} &= [B^T B]^{-1} B^T Y \\ &= \frac{\begin{bmatrix} -(n-1)\sum_{k=2}^{n} x^{(0)}(k) \cdot z^{(1)}(k) + \sum_{k=2}^{n} x^{(0)}(k) \cdot \sum_{k=2}^{n} z^{(1)}(k) \\ -\sum_{k=2}^{n} z^{(1)}(k) \cdot \sum_{k=2}^{n} x^{(0)}(k) \cdot z^{(1)}(k) + \sum_{k=2}^{n} x^{(0)}(k) \cdot \sum_{k=2}^{n} [z^{(1)}(k)]^2 \end{bmatrix}}{(n-1)\sum_{k=2}^{n} [z^{(1)}(k)]^2 - \left[\sum_{k=2}^{n} z^{(1)}(k)\right]^2} \\ &= \begin{bmatrix} \dfrac{\frac{1}{n-1}\sum_{k=2}^{n} x^{(0)}(k) \cdot \sum_{k=2}^{n} z^{(1)}(k) - \sum_{k=2}^{n} x^{(0)}(k) \cdot z^{(1)}(k)}{\sum_{k=2}^{n} [z^{(1)}(k)]^2 - \frac{1}{n-1}\left[\sum_{k=2}^{n} z^{(1)}(k)\right]^2} \\ \dfrac{1}{n-1}\left[\sum_{k=2}^{n} x^{(0)}(k) + a\sum_{k=2}^{n} z^{(1)}(k)\right] \end{bmatrix} \\ &= \begin{bmatrix} a & b \end{bmatrix}^T . \square \end{aligned}$$

Definition 7.3.4. *Assume that $X^{(0)}$ is a non-negative sequence, $X^{(1)}$ the sequence of 1-AGO generated from $X^{(0)}$, and $Z^{(1)}$ the sequence mean generated with consecutive neighbors of $X^{(1)}$. If*

$$[a, b]^T = \left[B^T B\right]^{-1} B^T Y,$$

then

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b$$

is called a whitenization (or image) equation of the grey differential equation

$$x^{(0)}(k) + az^{(1)}(k) = b.$$

Theorem 7.3.2. *Assume that B, Y, and $\hat{a}$ are the same as in Theorem 7.3.1. If*

$$[a, b]^T = \left[B^T B\right]^{-1} B^T Y.$$

then the following hold true.

1. The solution (or time response function) of the whitenization function

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b$$

is given by

$$x^{(1)}(t) = \left[x^{(1)}(0) - \frac{b}{a}\right] e^{-at} + \frac{b}{a}.$$

2. The time response sequence of the $GM(1,1)$ grey differential equation

$$x^{(0)}(k) + az^{(1)}(k) = b$$

is given by

$$\widehat{x}^{(1)}(k+1) = \left[x^{(1)}(0) - \frac{b}{a}\right] e^{-ak} + \frac{b}{a}, k = \ 1,\ 2,\ ...,\ n.$$

3. Let $x^{(1)}(0) = x^{(0)}(1)$, then

$$\widehat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-ak} + \frac{b}{a}, k = \ 1,\ 2,\ ...,\ n.$$

4. The restored values of $x^{(0)}(k)$s can be given by

$$\widehat{x}^{(0)}(k+1) = \alpha^{(1)}\widehat{x}^{(1)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k),$$

k = 1, 2, ..., n.

Definition 7.3.5. *The parameters $(-a)$ and b in the $GM(1,1)$ model are called the development coefficient and grey action quantity, respectively.*

The parameter $(-a)$ reflects development states of $\widehat{X}^{(1)}$ and $\widehat{X}^{(0)}$. In general, variables that act upon the system of interest should be external or predefined. However, GM(1, 1) is a single sequence modeling, which makes use of only the system's behavioral sequence (called output sequence, or background values) without considering any external acting sequences (called input sequences, or driving quantities). The grey action quantity in GM(1, 1) is a value derived from the background values. It reflects changes contained in the data and its exact intension is grey. The grey action quantity realizes the extension of the relevant intension. The existence of this grey action quantity distinguishes grey systems modeling from the general input-output modeling (or black box modeling), and is a teststone to separate the thoughts of grey systems and that of grey boxes.

Theorem 7.3.3. *The* $GM(1,1)$ *model*

$$x^{(0)}(k) + az^{(1)}(k) = b$$

can be transformed into

$$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1),$$

where

$$\beta = \frac{b}{1+0.5a}, \alpha = \frac{a}{1+0.5a}.$$

Proof. By substituting

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$$

into

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

we obtain that

$$\begin{aligned} x^{(0)}(k) & +0.5a\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \\ & = x^{(0)}(k) + 0.5a\left[x^{(1)}(k-1) + x^{(0)}(k) + x^{(1)}(k-1)\right] \\ & = (1+0.5a)x^{(0)}(k) + ax^{(1)}(k-1) = b. \end{aligned}$$

So,

$$x^{(0)}(k) = \frac{b}{1+0.5a} - \frac{a}{1+0.5a}x^{(1)}(k-1);$$

that is,

$$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1). \square$$

Theorem 7.3.4. *Assume that*

$$\beta = \frac{b}{1+0.5a}, \alpha = \frac{a}{1+0.5a}$$

and

$$\widehat{X}^{(1)} = \left(\widehat{x}^{(1)}(1), \widehat{x}^{(1)}(2), ..., \widehat{x}^{(1)}(n)\right)$$

is the time response sequence of the $GM(1,1)$ *model, where*

$$\widehat{x}^{(1)}(k) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-a(k-1)} + \frac{b}{a},$$

$k = 1, 2, ..., n$. Then

$$x^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right] e^{-a(k-2)}.$$

Proof. From Theorem 7.3.3, it followed that

$$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1).$$

Substituting the response value $\widehat{x}^{(1)}(k-1)$ of $x^{(1)}(k-1)$ into the equation above gives that

$$x^{(0)}(k) = \beta - \alpha \left[\left(x^{(0)}(1) - \frac{b}{a}\right) e^{-a(k-2)} + \frac{b}{a}\right];$$

that is,

$$x^{(0)}(k) = \beta - \alpha\frac{b}{a} + \left[\alpha\frac{b}{a} - \alpha x^{(0)}(1)\right] e^{-a(k-2)}.$$

Now from

$$\alpha\frac{b}{a} = \frac{a}{1+0.5a} \cdot \frac{b}{a} = \frac{b}{1+0.5a} = \beta,$$

we have

$$x^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right] e^{-a(k-2)}. \square$$

Example 7.3.1. Assume that

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{5} = (2.874, 3.278, 3.337, 3.390, 3.679)$$

is a given sequence of raw data. We use the following three GM models to simulate $X^{(0)}$ and compare their simulation accuracies.

1. $x^{(0)}(k) + az^{(1)}(k) = b$;
2. $x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1)$; and
3. $x^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right] e^{-a(k-2)}$.

Solution: 1. Step 1: Applying 1-AGO on $X^{(0)}$ gives us

$$X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{5} = (2.874, 6.152, 9.489, 12.879, 16.558).$$

Step 2: Perform a quasi-smoothness check on $X^{(0)}$. From

$$\rho(k) = \frac{x^{(0)}(k)}{x^{(1)}(k-1)},$$

it follows that

$$\rho(3) = \frac{x^{(0)}(3)}{x^{(1)}(2)} = \frac{3.337}{6.152} \approx 0.54,$$

$$\rho(4) = \frac{x^{(0)}(4)}{x^{(1)}(3)} = \frac{3.390}{9.489} \approx 0.36,$$

and

$$\rho(5) = \frac{x^{(0)}(5)}{x^{(1)}(4)} = \frac{3.679}{12.879} \approx 0.29.$$

So, $\rho(4) < 0.5$ and $\rho(5) < 0.5$. That is, for the case of $k > 3$, the condition of being quasi-smooth is satisfied.

Step 3: Check to see whether $X^{(1)}$ satisfies the law of quasi-exponentiality. From

$$\sigma^{(1)}(k) = \frac{x^{(1)}(k)}{x^{(1)}(k-1)},$$

it follows that

$$\sigma^{(1)}(3) = \frac{x^{(1)}(3)}{x^{(1)}(2)} = \frac{9.489}{6.152} \approx 1.54,$$

$$\sigma^{(1)}(4) = \frac{x^{(1)}(4)}{x^{(1)}(3)} = \frac{12.879}{9.489} \approx 1.36,$$

and

$$\sigma^{(1)}(5) = \frac{x^{(1)}(5)}{x^{(1)}(4)} = \frac{16.558}{12.879} \approx 1.29.$$

So, when $k > 3$, $\sigma^{(1)}(k) \in [1, 1.5]$, $\delta = 0.5$. That is, the law of quasi-exponentiality is satisfied. So, we can establish a GM(1, 1) model for $X^{(1)}$.

Step 4: Apply a consecutive neighbor generation to $X^{(1)}$. Let

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1).$$

Then, it follows that

$$\begin{aligned} Z^{(1)} &= \left(z^{(1)}(1), z^{(1)}(2), z^{(1)}(3), z^{(1)}(4), z^{(1)}(5)\right) \\ &= (2.874, 4.513, 7.820, 11.184, 14.718). \end{aligned}$$

Therefore,

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ -z^{(1)}(5) & 1 \end{bmatrix} = \begin{bmatrix} -4.513 & 1 \\ -7.820 & 1 \\ -11.184 & 1 \\ -14.718 & 1 \end{bmatrix},$$

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \end{bmatrix} = \begin{bmatrix} 3.278 \\ 3.337 \\ 3.390 \\ 3.679 \end{bmatrix};$$

and

$$\begin{aligned} B^T B &= \begin{bmatrix} -4.513 & 1 \\ -7.820 & 1 \\ -11.184 & 1 \\ -14.718 & 1 \end{bmatrix}^T \cdot \begin{bmatrix} -4.513 & 1 \\ -7.820 & 1 \\ -11.184 & 1 \\ -14.718 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 423.244 & -38.236 \\ -38.236 & 4 \end{bmatrix}. \end{aligned}$$

So,

$$\left[B^T B\right]^{-1} = \begin{bmatrix} 423.244 & -38.236 \\ -38.236 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 0.017317 & 0.165537 \\ 0.165537 & 1.83236 \end{bmatrix}.$$

Step 5: Perform a least squares estimate for the parametric sequence $\widehat{a} = [a, b]^T$. We can obtain that

$$\begin{aligned} \widehat{a} &= \left[B^T B\right]^{-1} B^T Y \\ &= \begin{bmatrix} 0.017317 & 0.165537 \\ 0.165537 & 1.83236 \end{bmatrix} \cdot \begin{bmatrix} -4.513 & 1 \\ -7.820 & 1 \\ -11.184 & 1 \\ -14.718 & 1 \end{bmatrix}^T \cdot \begin{bmatrix} 3.278 \\ 3.337 \\ 3.390 \\ 3.679 \end{bmatrix}; \end{aligned}$$

so,

$$\widehat{a} = \begin{bmatrix} 0.087385 & 1.085292 \\ 0.030118 & 0.537861 \\ -0.028136 & -0.019006 \\ -0.089335 & -0.604014 \end{bmatrix}^T \cdot \begin{bmatrix} 3.278 \\ 3.337 \\ 3.390 \\ 3.679 \end{bmatrix} = \begin{bmatrix} -0.0372 \\ 3.06536 \end{bmatrix}.$$

Step 6: Determine the model. We have

$$\frac{dx^{(1)}}{dt} - 0.0372x^{(1)} = 3.06536$$

and the time response sequence

$$\begin{aligned}\widehat{x}^{(1)}(k+1) &= \left[x^{(0)}(1) - \tfrac{b}{a}\right] e^{-ak} + \tfrac{b}{a} \\ &= 85.276151 \cdot e^{0.0372k} - 82.402151.\end{aligned}$$

Step 7: Solve the model obtained in Step 6 for the simulation value of $X^{(1)}$.

$$\widehat{X}^{(1)} = \left(\widehat{x}^{(1)}(i)\right)_{i=1}^{5} == (2.8740, 6.1060, 9.4605, 12.9422, 16.5558)\,.$$

Step 8: Restore the $\widehat{X}^{(1)}$-value to find the simulation value of $X^{(0)}$. From

$$\widehat{x}^{(0)}(k) = \alpha^{(1)}\widehat{x}^{(1)}(k) = \widehat{x}^{(1)}(k) - \widehat{x}^{(1)}(k-1),$$

it follows that

$$\widehat{X}^{(0)} = \left(\widehat{x}^{(0)}(i)\right)_{i=1}^{5} = (2.8740, 3.2320, 3.3545, 3.4817, 3.6136)\,.$$

Step 9: Evaluate the error. The following Table 7.1 gives the relevant error values.

Table7.1. Relevant error values

	Real Data	Simulated Data	Errors	Relative Errors (%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \frac{\lvert\varepsilon(k)\rvert}{x^{(0)}(k)}$
2	3.278	3.2320	0.0460	1.40
3	3.337	3.3545	-0.0175	0.52
4	3.390	3.4817	-0.0917	2.71
5	3.679	3.6136	0.0654	1.78

From Table 7.1, we can compute the square sum of the errors:

$$s = \varepsilon^T \varepsilon = \begin{bmatrix} \varepsilon(2) \\ \varepsilon(3) \\ \varepsilon(4) \\ \varepsilon(5) \end{bmatrix}^T \cdot \begin{bmatrix} \varepsilon(2) \\ \varepsilon(3) \\ \varepsilon(4) \\ \varepsilon(5) \end{bmatrix}$$

$$= \begin{bmatrix} 0.0460 \\ -0.0175 \\ -0.0917 \\ 0.0654 \end{bmatrix}^T \cdot \begin{bmatrix} 0.0460 \\ -0.0175 \\ -0.0917 \\ 0.0654 \end{bmatrix} = 0.01511$$

and obtain the average relative error:

$$\Delta = \frac{1}{4}\sum_{k=2}^{5} \Delta_k = 1.6025\%.$$

2. From 1, we know that

$$a = -0.03720 \text{ and } b = 3.06536.$$

So,

$$\alpha = \frac{a}{1+0.5a} = \frac{-0.03720}{1+0.5 \cdot (-0.0372)} = -0.0379$$

and

$$\beta = \frac{b}{1+0.5a} = \frac{3.06536}{1+0.5 \cdot (-0.0372)} = 3.1235,$$

and it follows that

$$\begin{aligned} \widehat{x}^{(0)}(k) &= \beta - \alpha x^{(1)}(k-1) \\ &= 3.1235 + 0.0379 x^{(1)}(k-1). \end{aligned}$$

Therefore,

$$\widehat{X}^{(0)} = \left(\widehat{x}^{(0)}(i)\right)_{i=1}^{5} = (3.2324, 3.2324, 3.3567, 3.4820, 3.6105).$$

Now, we evaluate errors.

Table 7.2. Computed error values

	Real Data	Simulated Data	Errors	Relative Errors (%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \frac{\lvert\varepsilon(k)\rvert}{x^{(0)}(k)}$
2	3.278	3.2324	0.0456	1.39
3	3.337	3.3567	-0.0197	0.59
4	3.390	3.4820	-0.0920	2.71
5	3.679	3.6105	0.0685	1.86

From Table 7.2, we can compute the sum of squares of errors as follows,

$$s = \varepsilon^T \varepsilon = 0.0156$$

and the average relative error as follows,

$$\Delta = \frac{1}{4}\sum_{k=2}^{5} \Delta_k = 1.6375\%.$$

3. From 2, it follows that

$$\alpha = -0.0379, \beta = 3.1235.$$

So, we have

$$\begin{aligned}\widehat{x}^{(0)}(k) &= \left[\beta - \alpha x^{(0)}(1)\right] e^{-a(k-2)} \\ &= [3.1235 + 0.0379 \cdot 2.874] \cdot e^{0.0372(k-2)} \\ &= 3.2324246 \cdot e^{0.0372(k-2)}.\end{aligned}$$

Therefore,

$$\widehat{X}^{(0)} = \left(\widehat{x}^{(0)}(i)\right)_{i=1}^{5} = (3.1144, 3.2324, 3.3549, 3.4821, 3.6141)\,.$$

Now, we do an error evaluation (see Table 7.3).

Table 7.3. Error evaluations

	Real Data	Simulated Data	Errors	Relative Errors (%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \frac{\lvert\varepsilon(k)\rvert}{x^{(0)}(k)}$
2	3.278	3.2324	0.0456	1.39
3	3.337	3.3549	-0.0179	0.54
4	3.390	3.4821	-0.0921	2.72
5	3.679	3.6141	0.0649	1.76

From Table 7.3, we can compute the sum of squares of errors as follows,

$$s = \varepsilon^T \varepsilon = 0.01509$$

and the average relative error as follows,

$$\Delta = \frac{1}{4}\sum_{k=2}^{5} \Delta_k = 1.6025\%.$$

4. From the sums of squares of errors and average relative errors of the three models used, it can be seen that the exponential models

$$\begin{cases} \widehat{x}^{(1)}(k) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-a(k-1)} + \frac{b}{a} \\ \widehat{x}^{(0)}(k) = \widehat{x}^{(1)}(k) - \widehat{x}^{(1)}(k-1) \end{cases}$$

and

$$\widehat{x}^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right] \cdot e^{-a(k-2)}$$

have relatively high accuracy, whereas the accuracy of the difference model

$$\widehat{x}^{(0)}(k) = \beta - \alpha x^{(1)}(k-1)$$

is relatively low.

Theorem 7.3.5. *If $X^{(0)}$ is a quasi-smooth sequence, then the development coefficient a of the $GM(1,1)$ model of its 1-AGO sequence $X^{(1)}$ can be written as*

$$a = \frac{\dfrac{b}{x^{(1)}(k-1)} - \rho(k)}{1 + 0.5\rho(k)},$$

where

$$\rho(k) = \frac{x^{(0)}(k)}{x^{(1)}(k-1)}.$$

Proof. From

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

it follows that

$$x^{(0)}(k) + a[0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)] = b,$$

$$x^{(0)}(k) + a[0.5x^{(0)}(k) + 0.5x^{(1)}(k-1) + 0.5x^{(1)}(k-1)] = b,$$

$$x^{(0)}(k) + a[0.5x^{(0)}(k) + x^{(1)}(k-1)] = b.$$

Therefore,

$$a = \frac{b - x^{(0)}(k)}{x^{(1)}(k-1) + 0.5x^{(0)}(k)},$$

and

$$a = \frac{\dfrac{b}{x^{(1)}(k-1)} - \dfrac{x^{(0)}(k)}{x^{(1)}(k-1)}}{1 + 0.5\dfrac{x^{(0)}(k)}{x^{(1)}(k-1)}} = \frac{\dfrac{b}{x^{(1)}(k-1)} - \rho(k)}{1 + 0.5\rho(k)}. \square$$

From Theorem 7.3.5, it follows that when b is limited and $x^{(1)}(k-1)$ is sufficiently large, the development coefficient $(-a)$ of the GM(1, 1) model is mainly determined by the smooth ratio $\rho(k)$.

In each GM(1,1) model, Theorem 7.3.1 is the basis for the construction of such a model. Because the result in Theorem 7.3.1 is established using vertical distances of data, it is reasonable to expect better simulation results if true distances of data are applied. To this end, we devote the rest of this section to establish a new result, similar to Theorem 7.3.1, but based on the estimate of least sum of squared true distances between data points and a special exponential curve. To achieve this end, let us first look at how the concept of differentiation, as studied in calculus, has been generalized to the study of discrete time series data. For a fixed whole number r, the derivative of the sequence $D^r(X^{(0)}) = X^{(r)}$, where $X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)}\right)$ is a sequence of raw data and $X^{(r)} = \left(x_1^{(r)}, x_2^{(r)}, ..., x_n^{(r)}\right)$ the sequence obtained by applying the accumulating generator D on $X^{(0)}$ r times, is defined as follows.

$$\begin{aligned} \frac{d}{dt}X^{(r)}_{(k)} &= \lim_{\text{time unit} \to \text{minimum}} \frac{x_k^{(r)} - x_{k-1}^{(r)}}{\text{time unit}} \\ &= \frac{\sum_{i=1}^{k} x_i^{(r-1)} - \sum_{i=1}^{k-1} x_i^{(r-1)}}{1} \\ &= x_k^{(r-1)}. \end{aligned} \tag{7.1}$$

That is, the derivative of $X^{(r)}$ with respect to time is $X^{(r-1)}$.

Assume that one of the exponential (solution) curves of the following differential equation

$$\frac{dy}{dx} + \alpha y = \beta, \tag{7.2}$$

where α and β are some fixed constants, has been used to fit the given set of data:

$$(1, x_1^{(1)}), (2, x_2^{(1)}), (3, x_3^{(1)}), ..., (n, x_n^{(1)}).$$

Let (x_k, y_k) be the point on that special curve such that the distance between (x_k, y_k) and $(k, x_k^{(1)})$ is the minimum of all the distances between points on the special curve of equ. (7.2) and the data point $(k, x_k^{(1)})$, for k = 1, 2, ..., n.

We now consider the sum of squared distances between (x_k, y_k) and $(k, x_k^{(1)})$, $k = 1, 2, ..., n$, as follows.

$$\begin{aligned} s^2 &= \sum_{k=1}^{n} \left[(x_k - k)^2 + \left(y_k - x_k^{(1)} \right)^2 \right] \\ &= \sum_{k=1}^{n} \left(y_k - x_k^{(1)} \right)^2 \left[\left(\frac{x_k - k}{y_k - x_k^{(1)}} \right)^2 + 1 \right]. \end{aligned} \tag{7.3}$$

Because the distance between (x_k, y_k) and $(k, x_k^{(1)})$ equals that between the point $(k, x_k^{(1)})$ and the special curve of equ. (7.2), we have

$$\frac{y_k - x_k^{(1)}}{x_k - k} = -\frac{1}{\frac{dy_k}{dx}} \tag{7.4}$$

because the slope $\frac{y_k - x_k^{(1)}}{x_k - k}$ is perpendicular to the tangent of the special curve of equ. (7.2) at the point (x_k, y_k).

If the special curve of equ. (7.2) fits the data points $\{\left(k, x_k^{(1)}\right)$: $k = 1, 2, ..., n\}$ well, we should have according to equ. (7.1) the following,

$$\frac{dy_k}{dx} \approx \frac{dx_k^{(1)}}{dt} = x_k^{(0)}, \; k = 1, 2, ..., n. \tag{7.5}$$

Substituting equ. (7.5) into equ. (7.3) provides

$$s^2 = \sum_{k=1}^{n} \left(y_k - x_k^{(1)} \right)^2 \left[\left(x_k^{(0)} \right)^2 + 1 \right]. \tag{7.6}$$

Because $dy_k/dx + \alpha y_k = \beta$, equ. (7.5) implies

$$y_k = \frac{\beta - dy_k/dx}{\alpha} = \frac{\beta - x_k^{(0)}}{\alpha}. \tag{7.7}$$

Substituting equ. (7.7) into equ. (7.6) offers

$$s^2 = \sum_{k=1}^{n} \left(\frac{\beta - x_k^{(0)}}{\alpha} - x_k^{(1)} \right)^2 \left[\left(x_k^{(0)} \right)^2 + 1 \right]. \tag{7.8}$$

Now, we find the estimated values a and b for α and β such that s^2 in equ. (7.8) will be the minimum possible. By differentiating s^2 with respect to α and β, respectively, we have

$$\frac{\partial s^2}{\partial \alpha} = \frac{2}{\alpha^3} \sum_{k=1}^{n} \left(x_k^{(0)} - \beta \right) \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\beta - x_k^{(0)} - \alpha x_k^{(1)} \right]$$

and

$$\frac{\partial s^2}{\partial \beta} = \frac{2}{\alpha^2} \sum_{k=1}^{n} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\beta - x_k^{(0)} - \alpha x_k^{(1)} \right].$$

Assume that when $\alpha = a$ and $\beta = b$, s^2-value, as shown in equ.(7.8), reaches its minimum. Then, $\alpha = a$ and $\beta = b$ satisfy the following system of equations.

$$\begin{cases} \sum_{k=1}^{n} \left(x_k^{(0)} - \beta \right) \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\beta - x_k^{(0)} - \alpha x_k^{(1)} \right] = 0 \\ \sum_{k=1}^{n} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\beta - x_k^{(0)} - \alpha x_k^{(1)} \right] = 0. \end{cases} \tag{7.9}$$

By applying the first equation to the second equation in equ. (7.9), we obtain the following equivalent system.

$$\begin{cases} \sum_{k=1}^{n} x_k^{(0)} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\beta - x_k^{(0)} - \alpha x_k^{(1)} \right] = 0 \\ \sum_{k=1}^{n} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\beta - x_k^{(0)} - \alpha x_k^{(1)} \right] = 0. \end{cases} \tag{7.10}$$

Solving this system for α and β provides the following.

$$\alpha = \frac{\sum_{j,k=1}^{n} x_j^{(0)} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\left(x_j^{(0)} \right)^2 + 1 \right] \left[x_k^{(0)} - x_j^{(0)} \right]}{\sum_{j,k=1}^{n} x_k^{(0)} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\left(x_j^{(0)} \right)^2 + 1 \right] \left[x_k^{(1)} - x_j^{(1)} \right]} \tag{7.11}$$

and

$$\beta = \frac{\sum_{j,k=1}^{n} x_k^{(0)} x_j^{(1)} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\left(x_j^{(0)} \right)^2 + 1 \right] \left[x_j^{(0)} - x_k^{(0)} \right]}{\sum_{j,k=1}^{n} x_k^{(0)} \left[\left(x_k^{(0)} \right)^2 + 1 \right] \left[\left(x_j^{(0)} \right)^2 + 1 \right] \left[x_k^{(1)} - x_j^{(1)} \right]}. \tag{7.12}$$

For $i = 1, 2, \cdots, n$, let us define the following matrices:

$${}_i^{(0)}A = \begin{bmatrix} 1 & x_i^{(0)} \end{bmatrix}_{1\times 2}, {}^{(r,j)}A = \begin{bmatrix} j & j & \cdots & j \\ x_1^{(r)} & x_2^{(r)} & \cdots & x_n^{(r)} \end{bmatrix}_{2\times n},$$

$${}^{(0)}X = diag\left[x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)} \right] = \begin{bmatrix} x_1^{(0)} & 0 & \cdots & 0 \\ 0 & x_2^{(0)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & x_n^{(0)} \end{bmatrix}_{n\times n},$$

$$
\begin{aligned}
{}^{(0)}A &= diag\left[{}^{(0)}_{1}A,{}^{(0)}_{2}A,\cdots,{}^{(0)}_{n}A\right] \\
&= \begin{bmatrix} 1 & x_1^{(0)} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & x_2^{(0)} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & x_n^{(0)} \end{bmatrix}_{n\times 2n},
\end{aligned}
$$

$$
B = \left[{}^{(0)}_{1}A,{}^{(0)}_{2}A,\cdots,{}^{(0)}_{n}A\right]^T_{1\times 2n} = \begin{bmatrix} 1 & x_1^{(0)} & 1 & x_2^{(0)} & \cdots & 1 & x_n^{(0)} \end{bmatrix}^T,
$$

$$
C = \begin{bmatrix} {}^{(0)}_{1}A/x_1^{(0)} & {}^{(0)}_{2}A/x_2^{(0)} & \cdots & \cdots & {}^{(0)}_{n}A/x_n^{(0)} \\ & & B & & \end{bmatrix}^T_{2\times 2n}.
$$

Theorem 7.3.6. Assume that the time series $X^{(0)}$ as above is nonnegative. If $\widehat{a} = \begin{bmatrix} a & b \end{bmatrix}^T$ is the least sum of squared distances estimate sequence of parameters $\begin{bmatrix} \alpha & \beta \end{bmatrix}^T$ such that a curve satisfying equ. (7.2) provides the best true distance fit of the time series $X^{(0)}$, then $\widehat{a}$ satisfies

$$
\widehat{a} = \frac{{}^{(1,1)}A^{(0)}X^{(0)}AC \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} {}^{(0,1)}A^{(0)}X^{(0)}AB}{\det({}^{(1,1)}A^{(0)}X^{(0)}AC)}. \tag{7.13}
$$

Proof. First, let us look at the denominator of equ. (7.13).

$$
\begin{aligned}
\det({}^{(1,1)}A \quad {}^{(0)}X^{(0)}AC) &= \det\left(\begin{bmatrix} x_1^{(0)} & x_2^{(0)} & \cdots & x_n^{(0)} \\ x_1^{(0)}x_1^{(1)} & x_2^{(0)}x_2^{(1)} & \cdots & x_n^{(0)}x_n^{(1)} \end{bmatrix}\right. \\
&\quad \left.\times \begin{bmatrix} x_1^{(0)} + 1/x_1^{(0)} & \left(x_1^{(0)}\right)^2 + 1 \\ x_2^{(0)} + 1/x_2^{(0)} & \left(x_2^{(0)}\right)^2 + 1 \\ \vdots & \vdots \\ x_n^{(0)} + 1/x_n^{(0)} & \left(x_2^{(0)}\right)^2 + 1 \end{bmatrix}\right) \\
&= \det \begin{bmatrix} \sum_{k=1}^n \left[\left(x_k^{(0)}\right)^2 + 1\right] & \sum_{k=1}^n x_k^{(0)}\left[\left(x_k^{(0)}\right)^2 + 1\right] \\ \sum_{k=1}^n x_k^{(0)}\left[\left(x_k^{(0)}\right)^2 + 1\right] & \sum_{k=1}^n x_k^{(0)}x_k^{(1)}\left[\left(x_k^{(0)}\right)^2 + 1\right] \end{bmatrix} \\
&= \sum_{j,k=1}^n x_k^{(0)}\left[\left(x_k^{(0)}\right)^2 + 1\right]\left[\left(x_j^{(0)}\right)^2 + 1\right]\left[x_k^{(1)} - x_j^{(1)}\right].
\end{aligned} \tag{7.14}
$$

Next, let us look at the numerator of equ. (7.13) in two steps. First, we compute

$$^{(1,1)}A^{(0)}X^{(0)}AC\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

as follows:

$$^{(1,1)}A \quad ^{(0)}X^{(0)}AC\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} x_1^{(0)} & x_2^{(0)} & \cdots & x_n^{(0)} \\ x_1^{(0)}x_1^{(1)} & x_2^{(0)}x_2^{(1)} & \cdots & x_n^{(0)}x_n^{(1)} \end{bmatrix}$$

$$\times \begin{bmatrix} \left(x_1^{(0)}\right)^2+1 & -\left[x_1^{(0)}+1/x_1^{(0)}\right] \\ \left(x_2^{(0)}\right)^2+1 & -\left[x_2^{(0)}+1/x_2^{(0)}\right] \\ \vdots & \vdots \\ \left(x_n^{(0)}\right)^2+1 & -\left[x_n^{(0)}+1/x_n^{(0)}\right] \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^n x_k^{(0)}\left[\left(x_k^{(0)}\right)^2+1\right] & -\sum_{k=1}^n \left[\left(x_k^{(0)}\right)^2+1\right] \\ \sum_{k=1}^n x_k^{(0)}x_k^{(1)}\left[\left(x_k^{(0)}\right)^2+1\right] & -\sum_{k=1}^n x_k^{(1)}\left[\left(x_k^{(0)}\right)^2+1\right] \end{bmatrix}. \tag{7.15}$$

Secondly, we compute $^{(0,1)}A^{(0)}X^{(0)}AB$ as follows.

$$^{(0,1)}A^{(0)}X^{(0)}AB = \begin{bmatrix} x_1^{(0)} & x_2^{(0)} & \cdots & x_n^{(0)} \\ \left(x_1^{(0)}\right)^2 & \left(x_2^{(0)}\right)^2 & \cdots & \left(x_n^{(0)}\right)^2 \end{bmatrix} \begin{bmatrix} \left(x_1^{(0)}\right)^2+1 \\ \left(x_2^{(0)}\right)^2+1 \\ \vdots \\ \left(x_n^{(0)}\right)^2+1 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^n x_k^{(0)}\left[\left(x_k^{(0)}\right)^2+1\right] \\ \sum_{k=1}^n \left(x_k^{(0)}\right)^2\left[\left(x_k^{(0)}\right)^2+1\right] \end{bmatrix}. \tag{7.16}$$

When we put equs. (7.14) through (7.16) together, and compare our result with those in equs. (7.11) and (7.12), we finish our proof. □

7.4 Model: Remnant GM(1,1)

When the accuracy of a GM(1, 1) model is not meeting a predetermined requirement, one can establish a GM(1, 1) model using the error sequence to remedy the original model in order to improve the accuracy.

Definition 7.4.1. *Assume that* $X^{(0)}$ *is a sequence of raw data,* $X^{(1)}$ *the 1-AGO sequence of* $X^{(0)}$*, and*

$$\widehat{x}^{(1)}(k+1) = [x^{(0)}(1) - \frac{b}{a}] \cdot e^{-ak} + \frac{b}{a}$$

the time response expression of GM(1, 1). Then,

$$d\widehat{x}^{(1)}(k+1) = (-a) \cdot [x^{(0)}(1) - \frac{b}{a}] \cdot e^{-ak}$$

is called the restored value through derivatives.

Proposition 7.4.1. *Assume that*

$$d\widehat{x}^{(1)}(k+1) = (-a) \cdot [x^{(0)}(1) - \frac{b}{a}] \cdot e^{-ak}$$

k = 0, 1, 2, ..., n− 1, are restored values through derivatives, and

$$\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k)$$

k = 0, 1, ..., n− 1, the restored values through inverse accumulating. Then

$$d\widehat{x}^{(1)}(k+1) \neq \widehat{x}^{(0)}(k+1).$$

Proof.

$$\begin{aligned}
\widehat{x}^{(0)}(k+1) &= \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k) \\
&= \left[x^{(0)}(1) - \tfrac{b}{a}\right] \cdot e^{-ak} + \tfrac{b}{a} \\
&\quad - \left[x^{(0)}(1) - \tfrac{b}{a}\right] \cdot e^{-a(k-1)} - \tfrac{b}{a} \\
&= (1 - e^{a}) \cdot \left[x^{(0)}(1) - \tfrac{b}{a}\right] \cdot e^{-ak}.
\end{aligned}$$

Because

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \cdots + \frac{a^m}{m!} + \cdots,$$

we have

$$1 - e^a = -a - \frac{a^2}{2!} - \frac{a^3}{3!} - \cdots - \frac{a^m}{m!} - \cdots \neq -a.$$

Therefore,

$$d\widehat{x}^{(1)}(k+1) \neq \widehat{x}^{(0)}(k+1). \ \square$$

From Proposition 7.4.1, it can be seen that GM(1, 1) is neither a differential equation nor a difference equation. However, when $|a|$ is sufficiently small, $1 - e^a \approx -a$. So,

$$d\widehat{x}^{(1)}(k+1) \approx \widehat{x}^{(0)}(k+1).$$

This implies that the results of differentiation and difference are very close. Therefore, GM(1, 1) can be seen as both a grey differential equation and a grey difference equation.

Due to the fact that the restored values through derivatives and through inverse accumulating are not the same, in order to reduce possible errors caused in reciprocating operations, we often use the errors of $X^{(1)}$ to improve the simulation values of $X^{(1)}$, where

$$\widehat{x}^{(1)}(k+1) = [x^{(0)}(1) - \frac{b}{a}] \cdot e^{-ak} + \frac{b}{a}.$$

Definition 7.4.2. *Assume that*

$$\varepsilon^{(0)} = \left(\varepsilon^{(0)}(1), \varepsilon^{(0)}(2), ..., \varepsilon^{(0)}(n)\right),$$

where

$$\varepsilon^{(0)}(k) = x^{(1)}(k) - \widehat{x}^{(1)}(k),$$

is the error sequence of $X^{(1)}$. If there exists k_0 satisfying:

1. *For any $k \geq k_0, \varepsilon^{(0)}(k)$ has the same sign, and*
2. *$n - k_0 \geq 4$, then*

$$\left(|\varepsilon^{(0)}(k_0)|, |\varepsilon^{(0)}(k_0+1)|, ..., |\varepsilon^{(0)}(n)|\right)$$

is called the error sequence of modelability, which is still denoted as

$$\varepsilon^{(0)} = \left(\varepsilon^{(0)}(k_0), \varepsilon^{(0)}(k_0+1), ..., \varepsilon^{(0)}(n)\right).$$

Proposition 7.4.2. *Assume that $\varepsilon^{(0)}$ is an error sequence of modelability with*

$$\varepsilon^{(1)} = \left(\varepsilon^{(1)}(k_0), \varepsilon^{(1)}(k_0+1), ..., \varepsilon^{(1)}(n)\right)$$

being its 1-AGO sequence, whose GM(1, 1) time response sequence is given by

$$\widehat{\varepsilon}^{(1)}(k+1) = \left[\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon}\right] \cdot e^{-a_\varepsilon(k-k_0)} + \frac{b_\varepsilon}{a_\varepsilon}, k \geq k_0.$$

Then the simulation sequence of the error sequence $\varepsilon^{(0)}$ is given by

$$\widehat{\varepsilon}^{(0)} = \left(\widehat{\varepsilon}^{(0)}(k_0), \widehat{\varepsilon}^{(0)}(k_0+1), ..., \widehat{\varepsilon}^{(0)}(n)\right),$$

where

$$\hat{\varepsilon}^{(0)}(k+1) = (-a_\varepsilon) \cdot \left[\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon}\right] \cdot e^{-a_\varepsilon(k-k_0)}, k \geq k_0.$$

Definition 7.4.3. *If $\varepsilon^{(0)}$ is used to modify $\hat{X}^{(1)}$, the time response sequence after modification*

$$\hat{x}^{(1)}(k+1) = \begin{cases} \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak} + \frac{b}{a}, & k < k_0 \\ \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak} + \frac{b}{a} & \\ \pm a_\varepsilon \cdot \left[\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon}\right] \cdot e^{-a_\varepsilon(k-k_0)}, & k \geq k_0 \end{cases}$$

is called the GM$(1,1)$ model with error modification, or remnant GM$(1,1)$ for short.

Here, the sign of

$$\hat{\varepsilon}^{(0)}(k+1) = a_\varepsilon \times \left[\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon}\right] \times e^{-a_\varepsilon(k-k_0)},$$

the error modification value, needs to be the same as the error $\varepsilon^{(0)}$.

If a modeling of the error sequence of $X^{(0)}$ and $\hat{X}^{(0)}$

$$\varepsilon^{(0)} = \left(\varepsilon^{(0)}(k_0), \varepsilon^{(0)}(k_0+1), ..., \varepsilon^{(0)}(n)\right)$$

is applied to modify the simulation value $\hat{X}^{(0)}$ of $X^{(0)}$, then different methods of restoration from $\hat{X}^{(1)}$ to $\hat{X}^{(0)}$ can produce different time response sequences of error modification.

Definition 7.4.4. *If*

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = (1-e^a) \cdot \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-a(k-1)},$$

then the corresponding time response sequence of error modification

$$\hat{x}^{(0)}(k+1) = \begin{cases} (1-e^a) \cdot \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak}, & k < k_0 \\ (1-e^a) \cdot \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak} & \\ \pm a_\varepsilon \cdot \left[\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon}\right] \cdot e^{-a_\varepsilon(k-k_0)}, & k \geq k_0 \end{cases}$$

is called the error modification model of inverse accumulating restoration.

Definition 7.4.5. *If*

$$\hat{x}^{(0)}(k+1) = (-a) \cdot \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak},$$

then the corresponding time response sequence of error modification

$$\widehat{x}^{(0)}(k+1) = \begin{cases} (-a)\cdot\left[x^{(0)}(1)-\frac{b}{a}\right]\cdot e^{-ak}, & k < k_0 \\ (-a)\cdot\left[x^{(0)}(1)-\frac{b}{a}\right]\cdot e^{-ak} & \\ \pm a_\varepsilon\cdot\left[\varepsilon^{(0)}(k_0)-\frac{b_\varepsilon}{a_\varepsilon}\right]\cdot e^{-a_\varepsilon(k-k_0)}, & k \geq k_0 \end{cases}$$

is called the error modification model of derivative restoration.

In the discussions above, all the error simulation terms in remnant GM(1, 1) have been taken to be the derivative restoration. Of course, they can be taken to be inverse accumulating restoration. That is, we can take

$$\widehat{\varepsilon}^{(0)}(k+1) = (1-e^{a_\varepsilon})\cdot\left[\varepsilon^{(0)}(k_0)-\frac{b_\varepsilon}{a_\varepsilon}\right]\cdot e^{-a_\varepsilon(k-k_0)}, k \geq k_0.$$

As long as $|a_\varepsilon|$ is sufficiently small, the effects of different error restoration methods on the modified $\widehat{x}^{(0)}(k+1)$ are almost the same.

Example 7.4.1. The following sequence of data represents the morbidity in rape at Yunmeng County of Hubei Province, The People's Republic of China,

$$X^{(0)} = (x^{(0)}(i))_{i=1}^{13} = (6, 20, 40, 25, 40, 45, 35, 21, 14, 18, 15.5, 17, 15)\,.$$

Establishing a GM(1, 1) model produces the following time response sequence

$$\widehat{x}^{(1)}(k+1) = -567.999\cdot e^{-0.06486k} + 573.999.$$

Applying inverse accumulating restoration gives

$$\begin{aligned} \widehat{X}^{(0)} &= \{\widehat{x}^{(0)}(k)\}_2^{13} \\ &= (35.6704, 33.4303, 31.3308, 29.3682, 27.5192, 25.7900, \\ &\quad 24.1719, 22.6534, 21.2307, 19.8974, 18.6478, 17.4768). \end{aligned}$$

To evaluate the accuracy, we obtain Table 7.4 of accuracies.

Table 7.4. Error evaluations for accuracies

	Real Data	Simulated Values	Errors	Relative Errors(%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \left\lvert \frac{\varepsilon(k)}{x^{(0)}(k)} \right\rvert$
2	20	35.6704	-15.6704	78.3540
3	40	33.4303	6.5697	16.4242
4	25	31.3308	-6.3308	25.3232
5	40	29.3682	10.6318	26.5795
6	45	27.5192	17.4808	38.8642
7	35	25.7901	9.2099	26.3140
8	21	24.1719	-3.1719	15.1043
9	14	22.6534	-8.6534	61.8100
10	18	21.2307	-3.2307	17.9483
11	15.5	19.8974	-4.3974	28.3703
12	17	18.6478	-1.6478	9.6929
13	15	17.4768	-2.4768	16.5120

From this table, it can be seen that the simulation error is relatively large. Now, we further calculate the sum of squares of errors

$$s = \varepsilon^T \varepsilon = 957.18,$$

and the average relative error

$$\Delta = \frac{1}{12}\sum_{k=2}^{13} \Delta_k = 30.11\%.$$

The sum of squares of errors is very large with the relevant accuracy less than 70%. So, it is necessary to apply a remnant model to do some remedies. Taking $k_0 = 9$ gives error sequence

$$\varepsilon^{(0)} = \left(\varepsilon^{(0)}(i)\right)_{i=9}^{13} = (-8.6534, -3.2307, -4.3974, -1.6478, -2.4768)\,,$$

which is an error sequence of modelability. Taking absolute value gives that

$$\varepsilon^{(0)} = (8.6534, 3.2307, 4.3974, 1.6478, 2.4768)\,.$$

Establishing a GM(1, 1) model for $\varepsilon^{(0)}$ produces the time response sequence for $\varepsilon^{(1)}$, the 1-AGO sequence of $\varepsilon^{(0)}$, as follows,

$$\widehat{\varepsilon}^{(1)}(k+1) = -24 \cdot e^{-0.16855(k-9)} + 32.7,$$

whose restored value of derivatives is

$$\begin{aligned} \widehat{\varepsilon}^{(0)}(k+1) &= (-0.16855) \cdot (-24) \cdot e^{-0.16855(k-9)} \\ &= 4.0452 \cdot e^{-0.16855(k-9)}. \end{aligned}$$

From

$$\begin{aligned}\hat{x}^{(0)}(k+1) &= \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \\ &= (1-e^{a}) \cdot \left[x^{(0)}(1) - \tfrac{b}{a}\right] \cdot e^{-ak} \\ &= 38.0614 \cdot e^{-0.06486k},\end{aligned}$$

we can obtain the remnant model of inverse accumulating restoration

$$\hat{x}^{(0)}(k+1) = \begin{cases} 38.0614 \cdot e^{-0.06486k}, & k < 9 \\ 38.0614 \cdot e^{-0.06486k} - 4.0452 \cdot e^{-0.16855(k-9)}, & k \geq 9, \end{cases}$$

where the sign of $\hat{\varepsilon}^{(0)}(k+1)$ is the same as the original error sequence.

Based on this model, we can modify the four simulation values with $k = 10, 11, 12, 13$, with improved accuracy listed in Table 7.5.

Table 7.5. Improved error checks

	Real Data	Modified simulated values	Errors	Relative Errors(%)
No.	$x^{(0)}(k)$	$\hat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$	$\Delta_k = \left\lvert \frac{\varepsilon(k)}{x^{(0)}(k)} \right\rvert$
10	18	17.1858	0.8142	4.52
11	15.5	16.4799	−0.9799	6.32
12	17	15.7604	1.2396	7.29
13	15	15.0372	−0.0372	0.25

From this table, we can compute the sum of squares of errors as follows,

$$s = \varepsilon^T \varepsilon = 3.1611,$$

and the average relative error

$$\Delta = \frac{1}{4} \sum_{k=10}^{13} \Delta_k = 4.595\%.$$

Here, the simulation accuracy of the remnant GM(1, 1) has been obviously increased. However, the current error sequence no longer satisfies the modeling requirement. So, if the improved accuracy is still unsatisfactory, we will have to consider other models or some appropriate choice of data to the original sequence.

7.5 Model Group of GM(1,1) Type

In practical modeling, some of the data values in the original sequence may not be applied to do the modeling. Each sub-sequence of the original data values can be used to establish a model. In general, different models are built based on different sub-sequences of the original sequence. Even though we might establish GM(1, 1) models of the same kind, different choices of the sub-sequences will in general give different values for the parameters a and b. These variations are exactly the reflection in modeling of the fact that different circumstances and different conditions in the environment affect the characteristics of the system under consideration. For example, for the food production in China, if we use the data values collected since 1949 to establish a GM(1, 1) model, the development coefficient $(-a)$ will be on the small side, but if only the values collected after 1978 are used, the corresponding development coefficient $(-a)$ will be obviously increased.

Definition 7.5.1. *Assume that the following sequence*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

takes $x^{(0)}(n)$ *as the origin of the time axis. Then the time with* $t < n$ *is called the past,* $t = n$ *the present, and* $t > n$ *the future.*

Definition 7.5.2. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence and

$$\widehat{x}^{(0)}(k+1) = (1 - e^{a}) \cdot \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak}$$

the restored value of the $GM(1,1)$ *time response sequence of* $X^{(0)}$. *Then*

1. *When* $t < n$, $\widehat{x}^{(0)}(t)$ *is called a simulation value of the model, or model simulation (value);*

2. *When* $t = n$, $\widehat{x}^{(0)}(t)$ *is called a filter value of the model;*

3. *When* $t > n$, $\widehat{x}^{(0)}(t)$ *is called a prediction (value) of the model.*

The main purpose of modeling is to make predictions. So, in order to improve the accuracy of prediction, we first need to guarantee a sufficiently high accuracy of simulation at the time moment $t = n$. That is to say, in general, the data for modeling purposes should be taken as an equal-time-interval sequence including $x^{(0)}(n)$.

Definition 7.5.3. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence of raw data.

1. *The* $GM(1,1)$ *model, built on the entire sequence* $X^{(0)}$, *is called all-data-*$GM(1,1)$.

2. For $k_0 > 1$, the GM$(1,1)$ model, built on the following sub-sequence

$$X^{(0)} = \left(x^{(0)}(k_0), x^{(0)}(k_0+1), ..., x^{(0)}(n)\right)$$

is called a partial-data GM$(1,1)$.

3. Let $x^{(0)}(n+1)$ be the newest piece of information. When $x^{(0)}(n+1)$ is inserted into the sequence $X^{(0)}$, the GM$(1,1)$ model, built on

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n), x^{(0)}(n+1)\right)$$

is called a new information GM$(1,1)$.

4. The GM$(1,1)$ model, built on the following new sequence obtained by inserting $x^{(0)}(n+1)$ and deleting $x^{(0)}(1)$,

$$X^{(0)} = \left(x^{(0)}(2), ..., x^{(0)}(n), x^{(0)}(n+1)\right)$$

is called a metabolic GM$(1,1)$.

Example 7.5.1. Let us use the last four data values of the sequence of raw data in Example 7.4.1 to build a GM(1, 1) model. We now have

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{4} = (18, 15.5, 17, 15),$$

whose 1-AGO sequence is

$$X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{4} = (18, 33.5, 50.5, 65.5),$$

and the sequence of the mean generation of consecutive neighbors of $X^{(1)}$ is given by

$$Z^{(1)} = \left(z^{(1)}(i)\right)_{i=1}^{4} = (18, 25.75, 42, 58).$$

So, it follows that

$$B = \begin{bmatrix} -25.75 & 1 \\ -42 & 1 \\ -58 & 1 \end{bmatrix}, Y = \begin{bmatrix} 15.5 \\ 17 \\ 15 \end{bmatrix}.$$

So

$$\left[B^T B\right]^{-1} = \begin{bmatrix} 5791.0625 & -125.75 \\ -125.75 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 0.001923 & 0.080603 \\ 0.080603 & 3.711922 \end{bmatrix}$$

and

$$B^T Y = \begin{bmatrix} -1983.125 \\ 47.5 \end{bmatrix}.$$

Therefore,

$$\widehat{a} = \left[B^T B\right]^{-1} \cdot B^T Y = \begin{bmatrix} 0.015093 \\ \\ 16.47047 \end{bmatrix}.$$

The GM(1, 1) time response sequence of $X^{(1)}$ is

$$\begin{aligned} \widehat{x}^{(1)}(k+1) &= \left[x^{(0)}(1) - \frac{b}{a}\right] \cdot e^{-ak} + \frac{b}{a} \\ &= -1073.265487 \cdot e^{-0.015093k} + 1091.265487, \end{aligned}$$

with the following simulated sequence

$$\widehat{X}^{(1)} = (18, 34.0772, 49.9135, 65.5126).$$

Let

$$\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k);$$

we obtain a simulated sequence of $X^{(0)}$ as follows,

$$\widehat{X}^{(0)} = (18, 16.0772, 15.8363, 15.5991).$$

Here, it can be seen that the accuracy of these simulated values is much higher than that obtained in Example 7.4.1, where an all-data GM(1, 1) was used. Especially, for the relative error Δ_{13} of the filter value $\widehat{x}^{(0)}(13)$, the relevant error value of the all-data GM(1, 1) is four times greater than that of the partial-data GM(1, 1). This fact implies that filter accuracy can be improved by appropriately choosing the data to be used in the process of modeling.

Example 7.5.2. For the sequence of raw data given in Example 7.3.1

$$X^{(0)} = (2.874, 3.278, 3.337, 3.39, 3.679),$$

let us add a piece of new information $x^{(0)}(6) = 3.85$. We now establish a new information and a metabolic GM(1, 1) model, and compare the results.

Solution: 1. The new information model.

The sequence with the new piece of information added is given by

$$X^{(0)} = (2.874, 3.278, 3.337, 3.39, 3.679, 3.85).$$

Now,

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ -\frac{1}{2}[x^{(1)}(3) + x^{(1)}(4)] & 1 \\ -\frac{1}{2}[x^{(1)}(4) + x^{(1)}(5)] & 1 \\ -\frac{1}{2}[x^{(1)}(5) + x^{(1)}(6)] & 1 \end{bmatrix} = \begin{bmatrix} -4.513 & 1 \\ -7.82 & 1 \\ -11.184 & 1 \\ -14.7185 & 1 \\ -18.488 & 1 \end{bmatrix}, \ Y = \begin{bmatrix} 3.278 \\ 3.337 \\ 3.390 \\ 3.679 \\ 3.850 \end{bmatrix},$$

and

$$\hat{a} = \left[B^T B\right]^{-1} B^T Y = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -0.0429 \\ 3.02 \end{bmatrix}.$$

So, we obtain the GM(1, 1) time response sequence as follows,

$$\begin{cases} \hat{x}^{(1)}(k+1) = 73.263569 \cdot e^{0.0429k} - 70.3895692 \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \end{cases}$$

with the filter value at $k = 6$:

$$\hat{x}^{(0)}(6) = 3.8126,$$

error

$$\varepsilon(6) = x^{(0)}(6) - \hat{x}^{(0)}(6) = 0.0374,$$

and relative error

$$\Delta_6 = \frac{0.0374}{3.85} = 0.97\%.$$

2. The metabolic model.

Inserting a piece of new information and deleting an old piece of information give us the modeling sequence

$$X^{(0)} = (3.278, 3.337, 3.39, 3.679, 3.85).$$

Now,

$$B = \begin{bmatrix} -4.9465 & 1 \\ -8.31 & 1 \\ -11.8446 & 1 \\ -15.609 & 1 \end{bmatrix}, Y = \begin{bmatrix} 3.337 \\ 3.390 \\ 3.679 \\ 3.850 \end{bmatrix},$$

and

$$\hat{a} = \left[B^T B\right]^{-1} B^T Y = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -0.0523 \\ 3.0392 \end{bmatrix}.$$

So, the GM(1, 1) time response sequence is

$$\begin{cases} \hat{x}^{(1)}(k+1) = 61.388899 \cdot e^{0.0523k} - 58.110899 \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), \end{cases}$$

with the filter value at $k = 6$

$$\hat{x}^{(0)}(6) = 3.856,$$

error

$$\varepsilon(6) = x^{(0)}(6) - \widehat{x}^{(0)}(6) = 0.006,$$

and relative error

$$\Delta_6 = \frac{0.006}{3.85} = 0.16\%.$$

3. Comparison of accuracies.

Table 7.6. Comparison of errors

	Parameter		Filter Values	
Model Type	a	b	$\widehat{x}^{(0)}(5)$	$\widehat{x}^{(0)}(6)$
Old Model	-0.0372	3.0653	3.6136	
New-Info. Model	-0.0429	3.02	3.653	3.8126
Metabolic Model	-0.0523	3.0392	3.6595	3.856
	Errors		Relative Errors(%)	
Model Type	$\varepsilon(5)$	$\varepsilon(6)$	Δ_5	Δ_6
Old Model	0.0654		1.78	
New-Info. Model	0.026	0.0374	0.71	0.97
Metabolic Model	0.0195	0.006	0.53	0.16

From Table 7.6, for the accuracy of the filter value $x^{(0)}(5)$, both the new information model and the metabolic model are better than the old model. This end implies that the new information GM(1, 1) and the metabolic GM(1, 1) have better prediction abilities than the old model. As a matter of fact, in the development process of a grey system, there always exist some stochastic interferences or some driving forces entering the system as time goes on so that the consequent development of the system is accordingly affected. Therefore, when using the GM(1, 1) model to do predictions, high accuracy can be achieved only for the first or the second data values after the origin value $x^{(0)}(n)$. In general, the farther away into the future, and the farther away from the origin data value, the weaker the prediction ability of GM(1, 1) becomes. In practical applications, one needs to constantly consider those interferences and driving factors entering the system as time goes on and promptly add new pieces of information to the original sequence $X^{(0)}$ and establish consequent new information GM(1, 1) models.

From the accuracy of the filter value $x^{(0)}(6)$, it can be seen that the metabolic model is better than the new information model. From the angle of prediction, it can be seen that the metabolic model is the most ideal prediction model. As the system develops further, the significance of the older data reduces so that at the same time when new data are added, the older data are deleted promptly, the constantly renewing modeling sequence can better reflect the current characteristics of the system. Especially, as the accumulation of quantitative changes increases, a jump or sudden change in the system will occur. At this very moment, compared with the older system, the current system is completely different. Hence, deleting those older data values is obviously very reasonable. Besides, constantly accepting and giving can avoid some of the difficulty of the huge amount of computation involved in modeling due to the fact that increased information can increase computer storage space tremendously.

7.6 GM(1,N) and GM(0,N)

In this section, we look at how to generalize GM(1,1) modeling to other practical situations.

Definition 7.6.1. *Assume that*

$$X_1^{(0)} = \left(x_1^{(0)}(1), x_1^{(0)}(2), ..., x_1^{(0)}(n)\right)$$

is a sequence of data of a system's characteristics,

$$X_i^{(0)} = \left(x_i^{(0)}(1), x_i^{(0)}(2), ..., x_i^{(0)}(n)\right), i = 2, 3, ..., N,$$

sequences of relevant factors, $X_i^{(1)}$ *the 1-AGO sequence of* $X_i^{(0)}$, $i = 1, 2, ...,$ N, *and* $Z_1^{(1)}$ *the sequence mean generated based on consecutive neighbors of* $X_1^{(1)}$. *Then*

$$x_1^{(0)}(k) + az_1^{(1)}(k) = \sum_{i=2}^{N} b_i x_i^{(1)}(k)$$

is called a $GM(1,N)$ *grey differential equation.*

Definition 7.6.2. *In a* $GM(1,N)$ *grey differential equation,* $(-a)$ *is called the development coefficient of the system,* $b_i x_i^{(1)}(k)$ *the driving term,* b_i *the driving coefficient, and*

$$\widehat{a} = [a, b_2, b_3, ..., b_N]^T$$

the sequence of parameters.

Theorem 7.6.1. *Assume that* $X_1^{(0)}$, $X_i^{(0)}, i = 2, 3, ..., N$, $X_i^{(1)}$, *and* $Z_1^{(1)}$ *are the same as defined in Definition 7.6.1. Let*

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix}, Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \cdot \\ \cdot \\ \cdot \\ x_1^{(0)}(n) \end{bmatrix}.$$

Then the least squares estimate of the sequence of parameters

$$\widehat{a} = [a, b_2, b_3, ..., b_N]^T$$

satisfies

$$\widehat{a} = [B^T B]^{-1} B^T Y.$$

Definition 7.6.3. *Assume*

$$\widehat{a} = [a, b_2, b_3, ..., b_N]^T.$$

Then

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b_2 x_2^{(1)} + b_3 x_3^{(1)} + \cdots + b_N x_N^{(1)}$$

is called a whitenization equation (or shadow equation) of the following $GM(1, N)$ grey differential equation

$$x_1^{(0)}(k) + az_1^{(1)}(k) = b_2 x_2^{(1)}(k) + b_3 x_3^{(1)}(k) + \cdots + b_N x_N^{(1)}(k).$$

Theorem 7.6.2. *Assume that $X_i^{(0)}, X_i^{(1)}, i = 1, 2, ..., N$, $Z_1^{(1)}$, B, Y are all defined as in Theorem 7.6.1, and*

$$\widehat{a} = [a, b_2, b_3, ..., b_N]^T = [B^T B]^{-1} B^T Y.$$

Then, we have

1. The solution of the whitenization equation

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = \sum_{i=2}^{N} b_i x_i^{(1)}$$

is given by

$$\begin{aligned} x^{(1)}(t) &= e^{-at}\left[\sum_{i=2}^{N} \int b_i x_i^{(1)}(t) \cdot e^{at} dt + x^{(1)}(0) - \sum_{i=2}^{N} \int b_i x_i^{(1)}(0) dt\right] \\ &= e^{-at}\left[x^{(1)}(0) - t\sum_{i=2}^{N} b_i x_i^{(1)}(0) + \sum_{i=2}^{N} \int b_i x_i^{(1)}(t) \cdot e^{at} dt\right]. \end{aligned}$$

2. When all of $X_i^{(1)}, i = 2, 3, ..., N$, vary slightly,

$$\sum_{i=2}^{N} b_i x_i^{(1)}(k)$$

is seen as a grey constant. Then the approximate time response sequence of the $GM(1, N)$ grey differential equation

$$x_1^{(0)}(k) + az_1^{(1)}(k) = \sum_{i=2}^{N} b_i x_i^{(1)}(k)$$

is given by

$$\widehat{x}_1^{(1)}(k+1) = \left[x_1^{(1)}(0) - \frac{1}{a}\sum_{i=2}^{N} b_i x_i^{(1)}(k+1)\right] e^{-ak} + \frac{1}{a}\sum_{i=2}^{N} b_i x_i^{(1)}(k+1),$$

where $x_1^{(1)}(0)$ *is taken to be* $x_1^{(0)}(1)$.

3. We have the restoration of inverse accumulation:

$$\widehat{x}_1^{(0)}(k+1) = \alpha^{(1)}\widehat{x}_1^{(1)}(k+1) = \widehat{x}_1^{(1)}(k+1) - \widehat{x}_1^{(1)}(k).$$

4. We also have the $GM(1,N)$ *difference simulation expression:*

$$x_1^{(0)}(k) = -az_1^{(1)}(k) + \sum_{i=2}^{N} b_i x_i^{(0)}(k).$$

Definition 7.6.4. *Assume that* $X_1^{(0)}$ *is a data sequence of a system's characteristics,* $X_i^{(0)}, i = 2, 3, ..., N$, *are sequences of relevant factors, and* $X_i^{(1)}$ *the 1-AGO sequences of* $X_i^{(0)}$, i = *1, 2, ..., n. Then*

$$X_1^{(1)} = b_2 X_2^{(1)} + b_3 X_3^{(1)} + \cdots + b_N X_N^{(1)} + a$$

is called a $GM(0,N)$ *model.*

Because GM(0, N) does not contain derivatives, it is a static model. It, in form, is similar to that of linear regression models, but has some essential differences from linear regression models. In general, linear regression models are built based on the original data sets, and the foundation for GM(0, N) models is the 1-AGO sequences of the original data.

Theorem 7.6.3. *Assume that* $X_i^{(0)}$ *and* $X_i^{(1)}$ *are defined the same way as in Definition 7.6.4, and*

$$B = \begin{bmatrix} x_2^{(1)}(2) & x_3^{(1)}(2) & \cdots & x_N^{(1)}(2) & 1 \\ x_2^{(1)}(3) & x_3^{(1)}(3) & \cdots & x_N^{(1)}(3) & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_2^{(1)}(n) & x_3^{(1)}(n) & \cdots & x_N^{(1)}(n) & 1 \end{bmatrix}, Y = \begin{bmatrix} x_1^{(1)}(2) \\ x_1^{(1)}(3) \\ \cdots \\ x_1^{(1)}(n) \end{bmatrix}.$$

Then the least squares estimate of the parameter sequence

$$\widehat{b} = [b_2, b_3, \cdots, b_N, a]^T$$

is

$$\widehat{b} = \left[B^T B\right]^{-1} B^T Y.$$

Example 7.6.1. Assume that

$$X_1^{(0)} = \left(x_1^{(0)}(i)\right)_{i=1}^5 = (2.874, 3.278, 3.307, 3.390, 3.679)$$

is a sequence of a system's characteristics, and

$$X_2^{(0)} = \left(x_2^{(0)}(i)\right)_{i=1}^5 = (7.04, 7.645, 8.075, 8.530, 8.774)$$

a data sequence of a relevant factor. We establish a GM(1, 2) and a GM(0, 2) model, respectively.

Solution: 1. Assume that the GM(1, 2) whitenization equation is

$$\frac{dx_1^{(1)}}{dt} + ax_1^{(1)} = bx_2^{(1)}.$$

Performing an 1-AGO to $X_1^{(0)}$ and $X_2^{(0)}$ individually gives

$$X_1^{(1)} = \left(x_1^{(1)}(i)\right)_{i=1}^5 = (2.874, 6.152, 9.459, 12.849, 16.528),$$

and

$$X_2^{(1)} = \left(x_2^{(1)}(i)\right)_{i=1}^5 = (7.04, 14.685, 22.760, 31.290, 40.064).$$

The sequence mean generated based on consecutive neighbors of $X_1^{(1)}$ is given as follows.

$$Z_1^{(1)} = \left(z_1^{(1)}(i)\right)_{i=1}^5 = (2.874, 4.513, 7.8055, 11.154, 14.6885).$$

So, it follows that

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) \\ -z_1^{(1)}(4) & x_2^{(1)}(4) \\ -z_1^{(1)}(5) & x_2^{(1)}(5) \end{bmatrix} = \begin{bmatrix} -4.513 & 14.685 \\ -7.8055 & 22.760 \\ -11.154 & 31.290 \\ -14.6885 & 40.064 \end{bmatrix},$$

$$Y = \left[x_1^{(0)}(2), x_1^{(0)}(3), x_1^{(0)}(4), x_1^{(0)}(5)\right]^T = [3.278, 3.307, 3.390, 3.679]^T,$$

and

$$B^TB = \begin{bmatrix} 421.456748 & -1181.415309 \\ -1181.415309 & 3317.855021 \end{bmatrix},$$

$$\left[B^T B\right]^{-1} = \begin{bmatrix} 1.280899 & 0.456100 \\ 0.456100 & 0.162709 \end{bmatrix},$$

$$B^T Y = \begin{bmatrix} -4.513 & 14.685 \\ -7.8055 & 22.760 \\ -11.154 & 31.290 \\ -14.6885 & 40.064 \end{bmatrix}^T \cdot \begin{bmatrix} 3.278 \\ 3.307 \\ 3.390 \\ 3.679 \end{bmatrix} = \begin{bmatrix} -132.457454 \\ 376.873306 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \left[\begin{array}{cc} a & b \end{array}\right]^T &= \left[B^T B\right]^{-1} B^T Y \\ &= \begin{bmatrix} 1.280899 & 0.456100 \\ 0.456100 & 0.162709 \end{bmatrix} \cdot \begin{bmatrix} -132.457454 \\ 376.873306 \end{bmatrix} \\ &= \begin{bmatrix} 2.2273 \\ 0.9068 \end{bmatrix}. \end{aligned}$$

So, we obtain the estimation model

$$\frac{dx_1^{(1)}}{dt} + 2.2273 x_1^{(1)} = 0.9068 x_2^{(1)},$$

and approximate time response sequence

$$\begin{aligned} \widehat{x}_1^{(1)}(k+1) &= \left[x_1^{(0)}(1) - \tfrac{b}{a} x_2^{(1)}(k+1)\right] \cdot e^{-ak} + \tfrac{b}{a} x_2^{(1)}(k+1) \\ &= \left[2.874 - 0.4071 x_2^{(1)}(k+1)\right] e^{-2.2273k} + 0.4071 x_2^{(1)}(k+1). \end{aligned}$$

From this end, it follows that

$$\widehat{x}_1^{(1)}(2) = 5.6436, \widehat{x}_1^{(1)}(3) = 9.1913,$$

$$\widehat{x}_1^{(1)}(4) = 12.7258, \widehat{x}_1^{(1)}(5) = 16.3082.$$

From the IAGO restoration:

$$\widehat{x}_1^{(0)}(k) = \widehat{x}_1^{(1)}(k) - \widehat{x}_1^{(1)}(k-1),$$

we have

$$\widehat{X}_1^{(0)} = \left(\widehat{x}_1^{(0)}(i)\right)_{i=1}^{5} = (2.874, 2.770, 3.548, 3.535, 3.582)$$

and the table of errors (see Table 7.7 for more details).

Table 7.7. Computed errors

	Origianl Data	Simulated Data	Errors	Relative Errors(%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \left\lvert \frac{\varepsilon(k)}{x^{(0)}(k)} \right\rvert$
2	3.278	2.770	0.508	15.5
3	3.307	3.518	-0.241	7.3
4	3.390	3.535	-0.145	4.3
5	3.679	3.582	0.097	2.6

2. Assume that

$$X_1^{(1)} = bX_2^{(1)} + a$$

is the GM(0, 2) model. From

$$B = \begin{bmatrix} x_2^{(1)}(2) & 1 \\ x_2^{(1)}(3) & 1 \\ x_2^{(1)}(4) & 1 \\ x_2^{(1)}(5) & 1 \end{bmatrix} = \begin{bmatrix} 14.685 & 1 \\ 22.760 & 1 \\ 31.290 & 1 \\ 40.064 & 1 \end{bmatrix}$$

and

$$Y = \begin{bmatrix} x_1^{(1)}(2) \\ x_1^{(1)}(3) \\ x_1^{(1)}(4) \\ x_1^{(1)}(5) \end{bmatrix} = \begin{bmatrix} 6.152 \\ 9.459 \\ 12.849 \\ 16.528 \end{bmatrix},$$

we can obtain a least squares estimate of $\widehat{b} = \begin{bmatrix} b & a \end{bmatrix}^T$ as follows,

$$\widehat{b} = \begin{bmatrix} b & a \end{bmatrix}^T = \left[B^T B\right]^{-1} B^T Y.$$

That is

$$\begin{aligned}\widehat{b} &= \begin{bmatrix} b \\ a \end{bmatrix} \\ &= \begin{bmatrix} 3570.455021 & 113.799 \\ 113.799 & 4 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1417.146962 \\ 44.988 \end{bmatrix} \\ &= \begin{bmatrix} 0.003004 & -0.085460 \\ -0.085460 & 2.681312 \end{bmatrix} \cdot \begin{bmatrix} 1417.146962 \\ 44.988 \end{bmatrix} \\ &= \begin{bmatrix} 0.412435 \\ -0.482515 \end{bmatrix}.\end{aligned}$$

So, we have the GM(0, 2) approximation

$$\widehat{x}_1^{(1)}(k) = 0.412435 x_2^{(1)}(k) - 0.482515, k = 1, 2, ..., 5.$$

And, therefore, it follows that

$$\widehat{x}_1^{(1)}(1) = 2.421, \widehat{x}_1^{(1)}(2) = 5.574, \widehat{x}_1^{(1)}(3) = 8.905,$$

$$\widehat{x}_1^{(1)}(4) = 12.423, \widehat{x}_1^{(1)}(5) = 16.042.$$

From the IAGO restoration:

$$\widehat{x}_1^{(0)}(k) = \widehat{x}_1^{(1)}(k) - \widehat{x}_1^{(1)}(k-1),$$

we have that

$$\widehat{X}_1^{(0)} = \left(\widehat{x}_1^{(0)}(i)\right)_{i=1}^{5} = (2.421, 3.153, 3.331, 3.518, 3.619)$$

and the table of errors in Table 7.8.

Table 7.8. Computed errors

	Original Data	Simulated Data	Errors	Relative Errors(%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \left\lvert \frac{\varepsilon(k)}{x^{(0)}(k)} \right\rvert$
2	3.278	3.153	0.125	3.8
3	3.307	3.331	-0.024	0.7
4	3.390	3.518	-0.128	3.8
5	3.679	3.619	0.060	1.6

7.7 GM(2,1) and Verhulst Model

GM(1, 1) is useful for sequences almost satisfying the law of exponentiality, and can only be applied to describe monotonic processes of change. As for non-monotonic wavelike development sequences, or saturated sigmoid sequences, we can consider building GM(2, 1) and Verhulst models.

Definition 7.7.1. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right),$$

is a sequence of raw data, its 1-AGO sequence $X^{(1)}$ *is*

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\right),$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i),$$

$k = 1, 2, ..., n$, *and the 1-IAGO sequence* $\alpha^{(1)}X^{(0)}$ *of* $X^{(0)}$ *is*

$$\alpha^{(1)}X^{(0)} = \left(\alpha^{(1)}x^{(0)}(1), \alpha^{(1)}x^{(0)}(2), ..., \alpha^{(1)}x^{(0)}(n)\right),$$

where

$$\alpha^{(1)}x^{(0)}(k) = x^{(0)}(k) - x^{(0)}(k-1),$$

$k = 1, 2, ..., n$, *and the sequence mean generated of consecutive neighbors of* $X^{(1)}$ *is*

$$Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), ..., z^{(1)}(n)\right),$$

where

$$z^{(1)}(k) = \frac{1}{2}[x^{(1)}(k) + x^{(1)}(k-1)],$$

$k = 1, 2, ..., n$. *Then*

$$\alpha^{(1)}X^{(0)} + a_1 X^{(0)} + a_2 Z^{(1)} = b$$

is called a GM(2, 1) grey differential equation.

Definition 7.7.2. *The equation*

$$\frac{d^2 x^{(1)}}{dt} + a_1 \frac{dx^{(1)}}{dt} + a_2 x^{(1)} = b$$

is called a whitenization equation of a GM(2, 1) grey differential equation.

Theorem 7.7.1. *Assume that $X^{(0)}, X^{(1)}, Z^{(1)}$, and $\alpha^{(1)}X^{(0)}$ are defined the same way as in Definition 7.7.1, and*

$$B = \begin{bmatrix} -x^{(0)}(2) & -z^{(1)}(2) & 1 \\ -x^{(0)}(3) & -z^{(1)}(3) & 1 \\ \cdots & \cdots & \cdots \\ -x^{(0)}(n) & -z^{(1)}(n) & 1 \end{bmatrix},$$

$$Y = \begin{bmatrix} \alpha^{(1)}x^{(0)}(2) \\ \alpha^{(1)}x^{(0)}(3) \\ \cdots \\ \alpha^{(1)}x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} x^{(0)}(2) - x^{(0)}(1) \\ x^{(0)}(3) - x^{(0)}(2) \\ \cdots \\ x^{(0)}(n) - x^{(0)}(n-1) \end{bmatrix}.$$

Then the least squares estimate of the $GM(2,1)$ parameter sequence

$$\widehat{a} = \begin{bmatrix} a_1 & a_2 & b \end{bmatrix}^T$$

is given by

$$\widehat{a} = \left[B^T B\right]^{-1} B^T Y.$$

Theorem 7.7.2. *As for the solution of the $GM(2,1)$ whitenization equation, the following hold true.*

1. If $X^{(1)*}$ is a special solution of

$$\frac{d^2x^{(1)}}{dt} + a_1\frac{dx^{(1)}}{dt} + a_2x^{(1)} = b$$

and $\overline{X}^{(1)}$ a general solution of the homogeneous equation

$$\frac{d^2x^{(1)}}{dt} + a_1\frac{dx^{(1)}}{dt} + a_2x^{(1)} = 0,$$

then $X^{(1)*} + \overline{X}^{(1)}$ is the general solution of the $GM(2,1)$ whitenization equation;

2. There are the following three cases for the general solution of the homogeneous equation above.

(a) When the characteristic equation

$$r^2 + a_1r + a_2 = 0$$

has two distinct real solutions r_1 *and* r_2*,*

$$\overline{X}^{(1)} = C_1 e^{r_1 t} + C_2 e^{r_2 t}.$$

(b) When the characteristic equation

$$r^2 + a_1 r + a_2 = 0$$

has a real solution r *of multiplicity 2,*

$$\overline{X}^{(1)} = e^{rt}(C_1 + C_2 t).$$

(c) When the characteristic equation

$$r^2 + a_1 r + a_2 = 0$$

has two complex conjugate solutions $r_1 = \alpha + i\beta, r_2 = \alpha - i\beta$*,*

$$\overline{X}^{(1)} = e^{\alpha t}(C_1 \cos \beta t + C_2 \sin \beta t).$$

3. There exist three possibilities for a special solution of a whitenization equation:

(a) When zero is not a solution of the characteristic equation,

$$X^{(1)*} = C;$$

(b) When zero is a solution of multiplicity 1 of the characteristic equation,

$$X^{(1)*} = Cx;$$

(c) When zero is a multiplicate solution of the characteristic equation,

$$X^{(1)*} = Cx^2.$$

Example 7.7.1. Assume that

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{5} = (2.874, 3.278, 3.337, 3.390, 3.679)$$

is a sequence of raw data. We now establish a GM(2, 1) grey differential equation.

Solution: The 1-AGO and 1-IAGO sequences of $X^{(0)}$ are respectively given as

$$X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{5} = (2.874, 6.152, 9.489, 12.879, 16.558)$$

and

$$\alpha^{(1)}X^{(0)} = \left(\alpha^{(1)}x^{(0)}(i)\right)_{i=1}^{5} = (0, 0.404, 0.059, 0.053, 0.289)\,,$$

and the sequence mean generated of consecutive neighbors of $X^{(1)}$

$$Z^{(1)} = \left(z^{(1)}(i)\right)_{i=1}^{5} = (2.874, 4.513, 7.820, 11.184, 14.7185)\,.$$

Now,

$$B = \begin{bmatrix} -x^{(0)}(2) & -z^{(1)}(2) & 1 \\ -x^{(0)}(3) & -z^{(1)}(3) & 1 \\ -x^{(0)}(4) & -z^{(1)}(4) & 1 \\ -x^{(0)}(5) & -z^{(1)}(5) & 1 \end{bmatrix} = \begin{bmatrix} -3.287 & -4.513 & 1 \\ -3.337 & -7.820 & 1 \\ -3.390 & -11.184 & 1 \\ -3.679 & -14.7185 & 1 \end{bmatrix},$$

and

$$Y = \left[\alpha^{(1)}x^{(0)}(i)\right]^T = [0.404, 0.059, 0.053, 0.289]^T\,.$$

So,

$$\widehat{a} = \left[\begin{array}{ccc} a_1 & a_2 & b \end{array}\right]^T = \left[B^T B\right]^{-1} B^T Y = [30.48, -1.04, 92.90]^T\,.$$

So, the GM(2, 1) whitenization equation

$$\frac{d^2x^{(1)}}{dt} + 30.48\frac{dx^{(1)}}{dt} - 1.04x^{(1)} = 92.90$$

has its characteristic equation

$$\lambda^2 + 30.48\lambda - 1.04 = 0,$$

which has two distinct real solutions $\lambda_1 = 0.0341$, and $\lambda_2 = -30.514$. So, the general solution of the homogeneous equation

$$\frac{d^2x^{(1)}}{dt} + 30.48\frac{dx^{(1)}}{dt} - 1.04x^{(1)} = 0$$

of the whitenization equation is

$$\overline{X}^{(1)}(t) = C_1 e^{0.0341t} + C_2 e^{-30.514t}.$$

Because zero is not a solution of the previous characteristic equation, we can easily obtain a special solution of the GM(2, 1) whitenization equation as follows.

$$X^{(1)*}(t) = -\frac{92.9}{1.04} = -89.3269.$$

Therefore, we have

$$\begin{aligned}\widehat{X}^{(1)}(t) &= \overline{X}^{(1)}(t) + X^{(1)*}(t)\\ &= C_1 e^{0.0341t} + C_2 e^{-30.514t} - 89.3269.\end{aligned}$$

Assume that $x^{(0)}(0) = 2.643$; then

$$\frac{dx^{(1)}}{dt}|_{t=0} = x^{(0)}(0) = 2.643.$$

Now, substituting

$$x^{(1)}(t)|_{t=0} = x^{(1)}(0) = x^{(0)}(1) = 2.874$$

and

$$\frac{dx^{(1)}}{dt}|_{t=0} = x^{(0)}(t)|_{t=0} = x^{(0)}(0) = 2.643$$

into

$$\widehat{X}^{(1)}(t) = C_1 e^{0.0341t} + C_2 e^{-30.514t} - 89.3269,$$

we obtain that

$$\begin{cases} 2.874 = C_1 + C_2 - 89.3269 \\ 2.643 = 0.0341C_1 - 30.514C_2. \end{cases}$$

So, it follows that

$$C_1 = 92.107983, C_2 = 2.931917.$$

So,

$$\widehat{X}^{(1)}(t) = 92.107983e^{0.0341t} + 2.931917e^{-30.514t} - 89.3269$$

Therefore, we obtain the GM(2, 1) time response sequence

$$\widehat{x}^{(1)}(k+1) = 92.107983e^{0.0341k} + 2.931917e^{-30.514k} - 89.3269$$

and

$$\widehat{X}^{(1)} = \left(\widehat{x}^{(1)}(i)\right)_{i=1}^{5} = (2.874, 5.9761, 9.2820, 12.7026, 16.2418)\,.$$

From the IAGO restoration:

$$\widehat{x}^{(0)}(k) = \widehat{x}^{(1)}(k) - \widehat{x}^{(1)}(k-1),$$

it follows that

$$\widehat{X}^{(0)} = \left(\widehat{x}^{(0)}(i)\right)_{i=1}^{5} = (2.874, 3.1021, 3.3059, 3.4206, 3.5392)\,.$$

Finally, we compute errors and relative errors, as shown in Table 7.9.

Table 7.9. Computed errors

	Original Data	Simulated Data	Errors	Relative Errors(%)
No.	$x^{(0)}(k)$	$\widehat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \widehat{x}^{(0)}(k)$	$\Delta_k = \left\lvert \frac{\varepsilon(k)}{x^{(0)}(k)} \right\rvert$
2	3.278	3.1021	0.1759	5.4
3	3.307	3.3059	0.0311	0.09
4	3.390	3.4206	-0.0306	0.09
5	3.679	3.5392	0.1399	3.8

Definition 7.7.3. *Assume that $X^{(0)}$ is a sequence of original data, $X^{(1)}$ the 1-AGO sequence of $X^{(0)}$, and $Z^{(1)}$ the sequence mean generated of consecutive neighbor of $X^{(1)}$. Then*

$$X^{(0)} + aZ^{(1)} = b[Z^{(1)}]^r$$

is called the $GM(1,1)$ power model.

Definition 7.7.4. *The equation*

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b[x^{(1)}]^r$$

is called the whitenization equation of the $GM(1,1)$ power model.

Theorem 7.7.3. *The solution of the whitenization equation of the $GM(1,1)$ power model is given by*

$$x^{(1)}(t) = \left\{ e^{-(1-r)at} \left[(1-r) \int be^{-(1-r)at} dt + C \right] \right\}^{\frac{1}{1-r}}.$$

Theorem 7.7.4. *Assume that $X^{(0)}, X^{(1)}$, and $Z^{(1)}$ are defined the same way as in Definition 7.7.3, and*

$$B = \begin{bmatrix} -z^{(1)}(2) & \left[z^{(1)}(2)\right]^r \\ -z^{(1)}(3) & \left[z^{(1)}(3)\right]^r \\ \cdots & \cdots \\ -z^{(1)}(n) & \left[z^{(1)}(n)\right]^r \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \cdots \\ x^{(0)}(n) \end{bmatrix}.$$

Then the least squares estimate of the parameter sequence $\widehat{a} = \begin{bmatrix} a & b \end{bmatrix}^T$ of the $GM(1,1)$ power model is given by

$$\widehat{a} = \left[B^T B\right]^{-1} B^T Y.$$

Definition 7.7.5. *When* $r = 2$,

$$X^{(0)} + aZ^{(1)} = b[Z^{(1)}]^2$$

is called the grey Verhulst model.

Definition 7.7.6. *The equation*

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b[x^{(1)}]^2$$

is called the whitenization equation of the grey Verhulst model.

Theorem 7.7.5. *1. The solution of the Verhulst whitenization equation is given by*

$$x^{(1)}(t) = \frac{ax^{(1)}(0)}{bx^{(1)}(0) + [a - bx^{(1)}(0)]e^{at}}.$$

2. The time response sequence of the grey Verhulst model is given by

$$\widehat{x}^{(1)}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0) + [a - bx^{(1)}(0)]e^{ak}}.$$

The Verhulst model is mainly used to describe and to study processes with saturated states (or say sigmoid processes). For example, this model is often used in the prediction of human populations, biological growth, reproduction, economic life span of consumable products, etc. From the solution of the Verhulst equation, it can be seen that when $t \to \infty$, if $a > 0$, then $x^{(1)}(t) \to 0$; if $a < 0$, then $x^{(1)}(t) \to \frac{a}{b}$. That is, when t is sufficiently large, for any $k > t$, $x^{(1)}(k+1)$ and $x^{(1)}(k)$ will be sufficiently close. At this time,

$$x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) > 0.$$

So, the system approaches extinction.

When resolving practical problems, we often face processes with sigmoid sequences of raw data. In this case, we can take the sequences of the original data as $X^{(1)}$ and the 1-IAGO sequence as $X^{(0)}$ to establish a Verhulst model to simulate $X^{(1)}$ directly.

Example 7.7.2. Study the number of large and medium-sized tractors used in agriculture in Henan Province from 1978 to 1982. The original data are given in Table 7.10.

Table 7.10. The original data

Year	1978	1979	1980	1981	1982
Number (in ten thousand)	4.1299	5.2382	5.9666	6.4590	6.3160

From Figure 7.3, it can be seen that the curve of the original data is similar to a sigmoid curve.

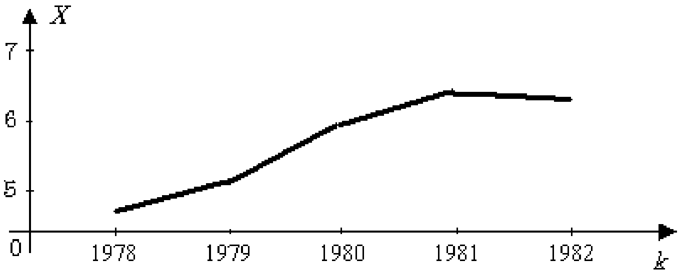

FIGURE 7.3. The curve of the original data

Let

$$X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{5} = (4.1299, 5.2382, 5.9666, 6.4590, 6.3160).$$

Then the 1-IAGO sequence $X^{(0)}$ of $X^{(1)}$ is

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{5} = (4.1299, 1.1083, 0.7284, 0.4924, -0.1430),$$

and the sequence $Z^{(1)}$ mean generated of consecutive neighbors of $X^{(1)}$ is

$$Z^{(1)} = \left(z^{(1)}(i)\right)_{i=1}^{5} = (4.1299, 4.68405, 5.6024, 6.2128, 6.3875).$$

Now,

$$B = \begin{bmatrix} -z^{(1)}(2) & \left[z^{(1)}(2)\right]^2 \\ -z^{(1)}(3) & \left[z^{(1)}(3)\right]^2 \\ -z^{(1)}(4) & \left[z^{(1)}(4)\right]^2 \\ -z^{(1)}(5) & \left[z^{(1)}(5)\right]^2 \end{bmatrix} = \begin{bmatrix} -4.68405 & 21.9403 \\ -5.6024 & 31.3869 \\ -6.2128 & 38.5989 \\ -6.3875 & 40.8002 \end{bmatrix},$$

$$\begin{aligned} Y &= \left[x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), x^{(0)}(5)\right]^T \\ &= [1.1083, 0.7284, 0.4924, -0.1430]^T, \end{aligned}$$

and

$$\begin{aligned} \left[B^T B\right]^{-1} &= \begin{bmatrix} 132.7263 & -779.0299 \\ -779.0299 & 4621.045657 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 0.716807 & 0.120842 \\ 0.120842 & 0.020588 \end{bmatrix}, \end{aligned}$$

$$B^T Y = \begin{bmatrix} -11.417292 \\ 60.350321 \end{bmatrix}.$$

So, it follows that

$$\widehat{a} = \begin{bmatrix} a \\ b \end{bmatrix} = \left[B^T B\right]^{-1} B^T Y = \begin{bmatrix} -0.891142 \\ -0.137196 \end{bmatrix}.$$

By taking $x^{(1)}(0) = x^{(1)}(1) = 4.1299$, we can obtain the Verhulst time response sequence as follows.

$$\widehat{x}^{(1)}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0) + [a - bx^{(1)}(0)]e^{ak}};$$

that is,

$$\widehat{x}^{(1)}(k+1) = \frac{3.680327}{-0.566606 - 0.324536e^{-0.891142k}}.$$

Based on this analysis, we can do simulation and prediction for the number of large and medium-sized tractors used in agriculture in Henan Province of China as follows.

$$\widehat{x}^{(1)}(1) = 4.1299, \widehat{x}^{(1)}(2) = 5.2597, \widehat{x}^{(1)}(3) = 5.9244,$$

$$\widehat{x}^{(1)}(4) = 6.2484, \widehat{x}^{(1)}(5) = 6.3918, \widehat{x}^{(1)}(6) = 6.4525,$$

where $\widehat{x}^{(1)}(6) = 6.4525$ is the predicted value for the number of large and medium sized tractors used in agriculture in Henan Province in 1983. The actual number for the year was $x^{(1)}(6) = 6.4389$. So, the error of prediction is

$$\varepsilon(6) = x^{(1)}(6) - \widehat{x}^{(1)}(6) = -0.0136,$$

and the relative error is

$$\Delta_6 = \left|\frac{\varepsilon(6)}{x^{(1)}(6)}\right| = \frac{0.0136}{6.4389} = 0.2\%.$$

Our accuracy of prediction has reached 99.8%. The simulation error is

$$\varepsilon^{(0)} = (\varepsilon(i))_{i=2}^{5} = (-0.0215, 0.0422, 0.02106, -0.0758).$$

So, we have achieved a very good simulation result.

8
Grey Combined Models

From examples and studies presented in the previous chapters, it can be seen that grey models possess such special abilities that they can weaken any randomness existing in the original data and uncover patterns of evolution of the system under investigation. When compared with other models widely used in practice, grey models seem to have very strong merging and penetrating capabilities. If one employs grey models in her entire process of modeling, she could supplement traditional models with the strength of grey models so that the accuracy of resultant predictions can be greatly improved. To this end, one can expect improvements in the following two aspects.

1. Model establishment is the core of systems analysis. When statistical models are considered, one needs to have large samples and large sets of observational values. However, in practice, due to various reasons, many available data sets just simply cannot satisfy these conditions. So, in these cases, the idea of establishing statistical models is out the window. When grey systems theory is applied, the ideal grey model is sought based on the principle that even though the original data need to be respected, one can employ scientific qualitative analysis to perform necessary modifications to the original experiment, observations, and statistical records. On the other hand, the centrally significant model in grey systems theory is the GM(1,1) model, where one only needs to have four data values to determine the parameters with a certain degree of accuracy. So, by using the principles and methodology of grey systems theory with the original observational data, one can expect to greatly improve the quality of modeling.

2. Each attempt at establishing a model is only a mapping reflecting one or several aspects of the subject of interest. Because the evolution of the system of interest is constantly influenced by many known and unknown, deterministic and random factors, it will be very limiting if one only applies one kind of modeling method in the endeavor to uncover the underlying development law of the system. Among all different methods of modeling, different models have their different characteristics. And, these characteristics can be employed to reveal different aspects of the subject under investigation. That is, when GM(1,1) models are applied together with other models, one can expect to deepen understanding about the system of concern and its evolution.

In this chapter, we study the grey econometric G-E model, grey Cobb–Douglas or production function G–C–D model, grey Markov G–M model, and grey time series model. In order to compare these grey combined models with traditional combined prediction models, in the last section of the chapter, we briefly look at the topic of combined predictions.

8.1 Econometric Models

8.1.1 Choice of Variables to Be Used in Modeling

Because it is very complicated to sort through various factors that affect the internal change of a system, the first question one needs to address before establish a meaningful model for the system is how to appropriately choose such variables to be applied in the system's model that they will eventually illustrate the changes of the system. The choice of variables depends on one's understanding of the system. On the other hand, one still needs to employ quantitative means to analyze the situation. To this end, the principle of grey incidences can be positively applied in order to resolve this problem.

Assume that y stands for an internal variable of the system of our study (for systems with more than one internal variable, these variables can be studied one by one), and x_1, x_2, ..., x_n the preimages of influencing factors that are either positively correlated or negatively correlated to y. These x-variables are seen as illustrating variables for behaviors of y. First, let us study the degree ε_i of incidence between y and x_i, $i = 1, 2, ..., n$. For a chosen lower threshold value ε_0, if $\varepsilon_i < \varepsilon_0$, let us delete x_i from the list of variables of our consideration. In this way, we can delete all those x-variables that have very weak correlation in terms of degree of incidence with the internal variable y. Assume that $x_{i_1}, x_{i_2}, ..., x_{i_m}$ are all the remaining variables. Now, we study the degrees $\varepsilon_{i_j i_k}$ of incidence between these variables, $i_j, i_k = i_1, i_2, ..., i_m$. For a chosen upper threshold value ε_0', if $\varepsilon_{i_j i_k} \geq \varepsilon_0'$, then the variables x_{i_j} and x_{i_k} are seen as the same kind of variables. In this way, all the remaining x- variables are divided into several subclasses. Now, we choose one variable from each subclass to enter into our final model building. The consequent result is that the final established econometric model can be greatly simplified without losing any expected power of description. At the same time, we can avoid the difficult collinearity problem.

8.1.2 Econometric Models

When one needs to estimate parameters of an econometric model, he often faces with difficulties of explaining some phenomena appearing out of his attempted approximation. For instance, such difficulties appear when the coefficients of some of the main variables are near zero, or when the signs of the estimated values of some parameters do not agree with the real life situation, or when slight vibrations in some observational values can cause several estimations of parameters to change largely, etc. Among the main reasons underlying these difficulties are (1) during the time frame within which observations are collected, the internal structure of the system of study has changed dramatically; (2) there exist multiple collinearity problems between the illustrating variables; and (3) there exist random fluctuations or noises in the observational values. In terms of the cases (1) and (2), one needs to re-consider the planned model structure or re-evaluate and re-organize the illustrating variables. When one is confronted with case (3), one can consider applying GM(1,1) simulated values of the original data to build the final model in order to reduce the effect of random fluctuation and noises existing in the original data. The consequent model, called a grey econometric combined model, in general can more adequately reflect the relationship between the internal variable y and its illustrating variables $x_i, i = 1, 2, ..., n$. At the same time, when predictions based on GM(1,1) made for the illustrating variables x_i are applied to produce predictions for the system's internal variable y of the grey econometric model, the results will be more scientifically solid. Also, by comparing predicted grey values made for the system's internal variable y with predictions obtained from an econometric model, the reliability of the study can be improved.

Grey econometric models can be applied to not only situations with known systems structures, but also to situations with the systems structure waiting to be studied and explored.

Example 8.1.1. Here, we use an example to illustrate how to apply the idea of grey econometric models to analyze and to make predictions on the food production of a specific region in the People's Republic of China. Here, due to our agreement with the officials of the region of concern, we do not reveal the identity of the region.

Based on the idea of grey combined econometric model building, in our study of that special region's food production system, we have abstracted the following 24 factors affecting the region's unit area food production.

The selection of these factors was done based on the results of three rounds of Delphi investigations conducted by 60 experts.

1. x_1: average amount of fertilizer applied per hectare of farmland in kilograms;
2. x_2: average amount of organic fertilizer applied per hectare of farmland in 100 kilograms;
3. x_3: percent of areas effectively irrigated;
4. x_4: percent of areas with guaranteed harvest in both a drought and a flood;
5. x_5: average number of irrigation wells per 10,000 hectares;
6. x_6: average annual capital input per hectare in \$10,000;
7. x_7: average automatic power applied per hectare in 10 watts;
8. x_8: percent of machine-cultivated farmland;
9. x_9: average electricity usage per hectare in kilowatt-hour;
10. x_{10}: percent of areas planted with scientifically hybridized seeds;
11. x_{11}: percent of areas planted with proven high-quality seeds;
12. x_{12}: average number of years of formal education received by farm workers;
13. x_{13}: percent of technicians among all farm workers;
14. x_{14}: number of scientists specialized in agriculture;
15. x_{15}: number of people engaged in research and application in agriculture;
16. x_{16}: percent of capital allocated to food production in the total government budget;
17. x_{17}: area affected by flood in 100,000 hectares;
18. x_{18}: area affect by drought in 100,000 hectares;
19. x_{19}: area affected by disaster and/or insects in 100,000 hectares;
20. x_{20}: area affected by wind disasters and/or hail in 100,000 hectares;
21. x_{21}: area affected by frost disasters in 100,000 hectares;
22. x_{22}: rate of idling farmland;

23. x_{23}: percent of areas planted with machines; and

24. x_{24}: percent of areas harvested with use of machines.

We then computed the degree ε_i of incidence of each of these variables with the variable of unit area food production. When taking a threshold value $\varepsilon_0 = 0.4$, such variables as $x_6, x_{12}, x_{16}, x_{19}, x_{20}, x_{21}, x_{22}, x_{23}$, and x_{24}, all have a degree ε_i of incidence with the variable of unit area food production smaller than 0.4. So, these variables are deleted from our consideration of illustrating variables affecting the unit area food production. Next, we computed the degree ε_{ij} of incidence between the remaining variables: x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{13}, x_{14}, x_{15}, x_{17}, and x_{18}. Now, we take the threshold value $\varepsilon_0' = 0.7$. When the degrees ε_{ij} of incidence were compared with this threshold value, we divided the 15 remaining illustrating variables into the following seven subclasses,

$$\{x_1\}, \{x_2\}, \{x_3, x_4, x_5\}, \{x_7, x_8, x_9\}, \{x_{10}, x_{11}, x_{13}, x_{14}, x_{15}\}, \{x_{17}\}, \{x_{18}\}.$$

By using x_3, x_7, and x_{14} as the representatives of the third, fourth, and the fifth subclasses, respectively, we obtained seven main illustrating variables x_1, x_2, x_3, x_7, x_{14}, x_{17}, and x_{18} affecting the unit area food production.

Before we established a systems model in the form of simple equations for the summer unit area food production y_1 and the autumn unit area food production y_2 in terms of these seven main illustrating variables, we first applied the data available for the time period from 1949 to 1997 to estimate our model parameters. The result was that we found several contradictions in our model development. To avoid these contradictions, we replaced the data we used earlier by those for the time period of 1957 to 1997. However, we still ran into some contradictions. Next, we tried to smooth the data using the exponential method. The result was still not satisfactory. At last, we applied GM(1,1) simulated values as our basis for the estimate of our parameters, and obtained the following system consisting of simply equations,

$$\begin{aligned} y_1 &= 126.4214 + 0.9686x_1 + 1.9669x_2 + 9.4071x_3 + \\ &\quad + 1.0212x_7 + 10.5503x_{14} - 0.6117x_{17} - 0.1853x_{18} + U_1 \end{aligned}$$

with the F-test value $F_1 = 679.2191$, R-test value $R_1 = 0.9799$, the standard error of our estimates $S_1 = 4.0174$, the Durbin–Watson test value $DW_1 = 1.3961$, and

$$\begin{aligned} y_2 &= 304.5194 + 0.7916x_1 + 1.7981x_2 + 12.8114x_3 + \\ &\quad + 5.3865x_7 + 9.1113x_{14} - 2.5417x_{17} - 3.6313x_{18} + U_2 \end{aligned}$$

with the F-test value $F_2 = 716.3874$, R-test value $R_2 = 0.9871$, the standard error of our estimates $S_2 = 3.9129$, and the Durbin–Watson test value

$DW_2 = 2.5346$. Here, we have obtained such a result that the illustrating variables have a clear influence on the unit area food production y_1 and y_2 where the explainability has achieved a level of 97.96% and 98.71%, respectively.

In order to further study the total food production, we still need to establish a model for the total planted area for the summer and autumn crops. The following are the main factors affecting the total planted area.

1. x_{25}: total area of farmland of the region in 100,000 hectares,
2. x_{26}: index of repeated planting,
3. x_{27}: percent of the area planted with grains, and
4. x_{28}: percent of the area planted with summer grains in the total area planted with grains.

Now, we have the formulas defining the areas y_3, indicating the total area planted for summer crops, and y_4, the total area planted for autumn crops as follows,

$$y_3 = x_{25} \cdot x_{26} \cdot x_{27} \cdot x_{28}$$

$$y_4 = x_{25} \cdot x_{26} \cdot x_{27} \cdot (1 - x_{28}) \, .$$

To estimate the internal variable, we need to know the values of the illustrating variables. In the study of development patterns of the illustrating variables, we have applied results of grey systems theory and established GM(1,1) models for the illustrating variables so that predictions about the systems internal variable could be made using predicted values of the illustrating variables.

The following lists the restored time response sequences for the illustrating variables x_1, x_2, x_3, x_7, x_{14}, x_{25}, x_{26}, x_{27}, and x_{28}. The predicted values of x_{17} and x_{18} are given by the disaster model (see next chapter for details).

$$
\begin{aligned}
\widehat{x}_1(1994+k) &= 960.42e^{0.0261k} + \eta_1, \\
\widehat{x}_2(1994+k) &= 553.41e^{0.0292k} + \eta_2, \\
\widehat{x}_3(1994+k) &= 40.06e^{0.0216k} + \eta_3, \\
\widehat{x}_7(1994+k) &= 23.16e^{0.0466k} + \eta_7, \\
\widehat{x}_{14}(1994+k) &= 20.84e^{0.043k} + \eta_{14}, \\
\widehat{x}_{25}(1994+k) &= 70.487e^{-0.003145k} + \eta_{25}, \\
\widehat{x}_{26}(1994+k) &= 169.00e^{0.004035k} + \eta_{26}, \\
\widehat{x}_{27}(1994+k) &= 78.71e^{-0.00455k} + \eta_{27}, \\
\widehat{x}_{28}(1994+k) &= 50.93e^{0.00561k} + \eta_{28}.
\end{aligned}
$$

Our predicted results for the three years of 2001, 2005, and 2010 are listed in Table 8.1. Here the predicted values for x_{17} and x_{18} are obtained by taking the averages of their individual areas affected by flood and drought for the relevant years.

Table 8.1. Predicted values for illustrating variables

Symbol	2001	2005	2010
$\widehat{x}_1$	1,152.94	1,279.81	1,458.22
$\widehat{x}_2$	678.92	763.03	882.98
$\widehat{x}_3$	46.60	50.80	56.60
$\widehat{x}_7$	32.09	38.67	48.81
$\widehat{x}_{14}$	28.16	33.44	41.47
$\widehat{x}_{17}$	5	1.7	11.7
$\widehat{x}_{18}$	5	11.7	32.38
$\widehat{x}_{25}$	68.95	68.09	67.03
$\widehat{x}_{26}$	173.84	176.67	180.27
$\widehat{x}_{27}$	76.24	74.87	73.18
$\widehat{x}_{28}$	52.97	54.17	55.71

Substituting the predicted values for the illustrating variables, as listed in Table 8.1, into the simplified equations for the summer and autumn unit area food productions y_1 and y_2 and the total planted areas for the summer crops and autumn crops y_3 and y_4 produces predicted values for these four variables. All the predicted values for these y-variables are listed in Table

8.2 as follows.

Table 8.2. Predicted values for y_1, y_2, y_3, and y_4

	2001	2005	2010
$\widehat{y}_1$ (kilograms/hectare)	3,342.77	3,733.80	4,282.22
$\widehat{y}_2$ (kilograms/hectare)	3,433.52	3,806.61	4,247.27
$\widehat{y}_3$ (in 100,000 hectares)	48.41	48.79	49.26
$\widehat{y}_4$ (in 100,000 hectares)	42.98	41.28	39.16

Therefore, we have obtained predicted values for the total summer food production y_5, total autumn food production y_6, and the total annual food production y for the region. These predicted values are given in Table 8.3 below.

Table 8.3. Predicted food productions in 10,000 tons

	2001	2005	2010
Total summer production y_5	1,618.23	1,821.72	2,109.42
Total autumn production y_6	1,475.73	1,571.37	1,663.23
Total annual production y	3,093.96	3,393.09	3,772.65

All the predicted values in Table 8.3 are obtained by considering the combined influence of many factors involved in the food production system of our special region. The importance level of each factor is determined by analyzing the system's history and the current situation. As the system develops and evolves, in the coming years, changes in the structure of the system might occur. In such a case, some of the main factors might go through relatively large fluctuations. And some of the minor factors might evolve into major factors influencing the final output of the system's production of food. In order to improve the reliability of our predictions, in our original study we further studied annual patterns existing in the production of each kind of crop and established GM(1,1) models for individual kinds of crops for their annual production. That is, we had produced predictions for the region's annual food production from a different angle in order to compare with results obtained from our earlier econometric prediction model. The following list the restored time sequences from different models and the defining formulas for the total crop productions.

- Total wheat production model:

$$\widehat{y}_5^1\,(1994+k) = 1,582.87\exp\left(0.0192k\right) + u_5^1;$$

- Total summer minor grain production model:

$$\widehat{y}_5^2\,(1994+k) = 28.18\exp\left(0.0152k\right) + u_5^2;$$

- Total rice production model:

$$\widehat{y}_6^1\,(1994+k) = 193.40\exp\left(0.02877k\right) + u_6^1;$$

- Total corn production model:

$$\widehat{y}_6^2(1994+k) = 519.66\exp(0.0468k) + u_6^2;$$

- Total production model for all kinds of potatoes:

$$\widehat{y}_6^3(1994+k) = 210.14\exp(0.0135k) + u_6^3;$$

- Total soybean production model:

$$\widehat{y}_6^4(1994+k) = 99.83\exp(0.0247k) + u_6^4;$$

- Total sorghum production model:

$$\widehat{y}_6^5(1994+k) = 17.39\exp(0.0502k) + u_6^5;$$

- Total millet production model:

$$\widehat{y}_6^6(1994+k) = 30.43\exp(0.0101k) + u_6^6;$$

- Total production model for all other autumn crops:

$$\widehat{y}_6^7(1994+k) = 21.12\exp(0.0236k) + u_6^7.$$

- The total summer food production is defined as:

$$\widehat{y}_5 = \widehat{y}_5^1 + \widehat{y}_5^2;$$

- The total autumn food production is defined as:

$$\widehat{y}_6 = \widehat{y}_6^1 + \widehat{y}_6^2 + \widehat{y}_6^3 + \widehat{y}_6^4 + \widehat{y}_6^5 + \widehat{y}_6^6 + \widehat{y}_6^7;$$

and the total food production of the region is defined as:

$$\widehat{y} = \widehat{y}_5 + \widehat{y}_6.$$

Here, minor grains include any grains other than wheat and rice. Based on these models and definitions, we obtained another set of predicted values for the region's main crop productions (see Table 8.4 for details).

Table 8.4. GM(1,1) predictions for total productions of major crops (in 10,000 tons)

	2001	2005	2010
Wheat production $\widehat{y}_5^1$	1,810.57	1,955.10	2,152.09
Summer minor grains production $\widehat{y}_5^2$	31.34	33.31	35.94
Rice production $\widehat{y}_6^1$	236.55	265.40	306.46
Corn production $\widehat{y}_6^2$	721.10	869.55	1,098.80
Potato production $\widehat{y}_6^3$	230.97	243.78	260.81
Soybean production $\widehat{y}_6^4$	118.67	131.00	148.22
Sorghum production $\widehat{y}_6^5$	12.23	10.01	7.79
Millet production $\widehat{y}_6^6$	32.66	34.01	35.77
Other autumn crop production $\widehat{y}_6^7$	24.91	27.38	30.81
Summer crop production $\widehat{y}_5$	1,841.91	1,988.41	2,188.03
Autumn crop production $\widehat{y}_6$	1,377.09	1,581.13	1,888.66
Annual crop production $\widehat{y}$	3,219.00	3,569.54	4,076.69

Comparing our two predictions, we see that they represent two different sets of values with differences in the amount of 4.04%, 5.2%, and 8.06%, respectively, for our region's total food production for the years of 2001, 2005, and 2010. These two different methods of prediction have produced such closely matched predictions that it explained that to a good degree, these prediction models have captured the essence of the objective law of development of the food production system of our concern. In this case, our prediction results had been applied as the basis for the government officials of the special region to adjust their policies in terms of food production and allocation. As the system evolves over time, the structure of our prediction model and its parameters need to be adjusted accordingly in order to adequately reflect the state of the system in different time moments in order to produce accurate and timely predictions in the years to come.

8.2 Cobb-Douglas Model

In this section, as the title suggests, we look at our second model, named either the Cobb–Douglas or production function model.

Definition 8.2.1. *Assume that K stands for the capital input, L the labor input, and Y the production output. Then,*

$$Y = A_0 e^{\gamma t} K^{\alpha} L^{\beta}$$

is called the C-D production function model, where α stands for the capital elasticity, β the labor elasticity, and γ the parameter for the progress of technology.

Definition 8.2.2. *The following*

$$\ln Y = \ln A_0 + \gamma t + \alpha \ln K + \beta \ln L$$

is called the log-linear form of the production function model.

For given time series data for the production output Y, capital input K, and labor input L,

$$Y = (y(1), y(2), ..., y(n)),$$

$$K = (k(1), k(2), ..., k(n)),$$

and

$$L = (\ell(1), \ell(2), ..., \ell(n)),$$

by using multivariate linear regression, one can obtain estimates for the parameters $\ln A_0$, γ, α, and β.

When Y, K, and L represent the time series of a special department, district or business, it is often the case that due to fluctuations existing in the data, the regressional estimates of the parameters contain serious errors leading to obviously unusable results. Some of the errors of estimates include near zero or negative value for γ, the parameter representing progress in technology, or the values of the elasticities α and β are out of their reasonable bounds.

In these circumstances, if one employs GM(1,1) simulated values of the time series Y, K, and L in place of the original data as the basis for multivariate regressional estimation, one could to an extent eliminate some of the random fluctuations existing in the original data. Consequently, the estimated values of the parameters would be more reasonable and the resultant model would more adequately reflect the relationship between the capital input, labor input, and progress in technology.

Definition 8.2.3. *Assume that*

$$\widehat{Y} = (\widehat{y}(1), \widehat{y}(2), ...\widehat{y}(n)),$$

$$\widehat{K} = \left(\widehat{k}(1), \widehat{k}(2), ..., \widehat{k}(n)\right),$$

and

$$\widehat{L} = \left(\widehat{\ell}(1), \widehat{\ell}(2), ..., \widehat{\ell}(n)\right)$$

are, respectively, GM(1,1) simulated values for Y, K, and L. Then,

$$\widehat{Y} = A_0 e^{\gamma t} \widehat{K}^{\alpha} \widehat{L}^{\beta}$$

is called a grey production function model.

In the grey production function model, no grey numbers appear explicitly. However, because it is a combination of a grey systems model and the C-D production function model, it contains a deep intension of “greyness”, and reflects the "principle of non-uniqueness of solutions" and the "principle of absoluteness of greyness". Therefore, one can expect to produce satisfactory results when such a combined model is applied in a real-life situation.

Example 8.2.1. Here, we look at how to make predictions about the rate of contribution made by progress in technology for different time frames in Henan Province, the People's Republic of China.

For the study of the rate of contribution made by progress in technology, in an economy, experts tend to apply a modified Cobb-Douglas production function model, using Solow's method of remnant values in the actual computation.

The formula of Solow's "remnant values" for calculating the speed of progress in technology is

$$\frac{\Delta A}{A} = \frac{\Delta Y}{Y} - \alpha\frac{\Delta K}{K} - \beta\frac{\Delta L}{L}. \tag{8.1}$$

If the influence of some non-technology aspects during the computed time frame of concern is very strong, equ. (8.1) generally does not produce a reasonable outcome. In this case, one can first apply the principles and results of grey systems theory to process the original data using buffer operators. Second, one can establish GM(1,1) models for the processed data. Thirdly, by using GM(1,1) simulated values one can establish a grey production function model. After substituting the output of the new production function model into equ. (8.1), one will be able to compute the rate of contribution to the production output made by progress in technology,

$$E_A = \left[\frac{\Delta\widehat{A}}{\widehat{A}} \div \frac{\Delta\widehat{Y}}{\widehat{Y}}\right] \times 100\%. \tag{8.2}$$

This quantity E_A is generally called the rate of contribution in economic growth made by progress in technology. Similarly, one can solve for the rates of contribution made by capital input and labor input as follows, respectively,

$$E_K = \left[\alpha\frac{\Delta\widehat{K}}{\widehat{K}} \div \frac{\Delta\widehat{Y}}{\widehat{Y}}\right] \times 100\% \tag{8.3}$$

and

$$E_L = \left[\beta\frac{\Delta\widehat{L}}{\widehat{L}} \div \frac{\Delta\widehat{Y}}{\widehat{Y}}\right] \times 100\%. \tag{8.4}$$

In order for us to describe accurately the historical evolutionary characteristics of the rate of contribution made by progress in technology for Henan Province in different time periods, we established grey production function models $\left(\widehat{Y}_1, \widehat{Y}_2, \widehat{Y}_3, \widehat{Y}_4\right)$ for the GDP of Henan Province for the following four time periods: 1952 to 1962, 1962 to 1970, 1970 to 1980, and 1980 to 1995, as follows.

$$\widehat{Y}_1 = 0.048 \exp(0.0581t)\, \widehat{K}_1^{0.2131} \widehat{L}_1^{0.7869},$$

$$\widehat{Y}_2 = 0.088 \exp(0.0072t)\, \widehat{K}_2^{0.5015} \widehat{L}_2^{0.4984},$$

$$\widehat{Y}_3 = 0.16 \exp(0.0098t)\, \widehat{K}_3^{0.5101} \widehat{L}_3^{0.4899},$$

and

$$\widehat{Y}_4 = 0.15 \exp(0.0161t)\, \widehat{K}_4^{0.3316} \widehat{L}_4^{0.6684},$$

where $\widehat{Y}_i$ stands for the GDP in 1 billion yuan, $\widehat{K}_i$ the fixed capital in 1 billion yuan, $\widehat{L}_i$ the number of people in the total work force, and t the time variable and $i = 1, 2, 3, 4$, the four time periods. Table 8.5 lists the computed speeds of technology progress and the rates of contribution made by progress in technology in economic growth for the four different time periods using equs. (8.1) and (8.2).

Table 8.5. Rates of contribution by technology for different times

Time Period	α	β	$\Delta\widehat{Y}/\widehat{Y}$	$\Delta\widehat{A}/\widehat{A}$	$E_A\,(\%)$
“First Five”	0.2131	0.7869	0.3296	0.1942	58.92
“Second Five”	0.2131	0.7869	-0.3833	–	–
1963 – 65	0.5015	0.4984	0.4969	0.2480	49.91
“Third Five”	0.5015	0.4984	0.4815	0.1553	32.25
“Fourth Five”	0.5101	0.4899	0.2573	0.0448	17.41
“Fifth Five”	0.5101	0.4899	0.5259	0.1564	29.75
“Sixth Five”	0.3316	0.6684	0.7397	0.2442	33.02
“Seventh Five”	0.3316	0.6684	0.4406	0.1841	41.78
“Eighth Five”	0.3316	0.6684	0.8389	0.3587	42.75

From Table 8.5, it can be seen that during the time period of the “first five”, the rate of contribution made by progress in technology was the highest for Henan Province, reaching 58.92%. The reason for such a high rate was because the establishment of the new China greatly liberated the productivity of the society, magnifying the “residual value” beyond the contributions made by capital input and labor input. However, during that time period, the level of technology was not high. During the “second five”, Henan Province suffered from severe natural disasters. Combined with human factors, the GDP in 1962 in Henan dropped 38.33% from the level of 1957 even though during the period the input of capital and labor had increased. Here, natural disasters and unrealistic blind leaps had consumed all the ”residual value” made from progress in technology. During 1963 to 1965, the rate of contribution made by progress in technology was relatively high. As a matter of fact, from looking at Henan’s GDP, the year of 1965 just recovered to the level of 1957. The relatively high “residual value” due to progress in technology contained influences of governmental economic policy adjustments. During the "fourth five" after the cultural revolution, both the speed of progress in technology and the rate of contribution made by progress in technology reached their lowest levels. After the "fifth five", both the speed of progress in technology and the rate of contribution made by progress in technology have been in an uptrend and reached their highest level during the "eighth five". Because annual statistical data fluctuate widely, it was difficult for us to obtain any meaningful results from an annual analysis in terms of progress in technology. However, when we employed longer time periods, such as the ones shown above, our computed results adequately reflected the history of Henan Province.

Based on equs. (8.3) and (8.4), we also computed the rates of contribution in the area of production output made by capital input and labor increase, respectively (See Table 8.6)). From Table 8.6 it can be seen that labor force increase contributed to the economic growth only slightly. To a degree, this fact indicates that the economic growth in Henan Province had mainly been a consequence of increased productivity. From the angles of progress in technology and amounts of capital input, other than the "first five", the three years of adjustment period, and the "seventh five", the rate of contribution by capital input had been greater than that by progress in technology. That implies that in Henan, the relatively high rate of economic growth had been sustained mainly by high levels of capital injection.

Table 8.6. Contributions due to capital & labor inputs

Time Period	$\alpha\Delta\widehat{K}/\widehat{K}$	$\beta\Delta\widehat{L}/\widehat{L}$	$E_K/\%$	$E_L/\%$
"First Five"	0.0671	0.0683	20.37	20.71
"Second Five"	0.3541	0.0826	–	–
1963 - 65	0.2117	0.0372	42.6	7.49
"Third Five"	0.2553	0.0709	53.02	14.72
"Fourth Five"	0.1287	0.0838	50.02	32.57
"Fifth Five"	0.3258	0.0437	61.95	8.31
"Sixth Five"	0.3606	0.1349	48.75	18.24
"Seventh Five"	0.1490	0.1075	33.82	24.39
"Eighth Five"	0.4110	0.0692	48.99	8.25

8.3 Markov Model

In this section, we study two classes of models named, respectively, grey moving probability and grey state Markov models.

8.3.1 Grey Moving Probability Markov Model

First, let us look at the class of grey moving probability Markov models.

Definition 8.3.1. *Assume that* $\{X_n : n \in T\}$ *is a stochastic process. If for any integer* $n \in T$ *and any states* $i_0, i_1, \cdots, i_{n+1} \in I$, *the following holds* *true,*

$$P\left(X_{n+1}=i_{n+1}|X_j=i_j, j=0,1,...,n\right)=P\left(X_{n+1}=i_{n+1}|X_n=i_n\right), \tag{8.5}$$

then $\{X_n : n \in T\}$ *is called a Markov chain.*

Equ. (8.5) is known as not having any post-effect. It represents that each future state (when $t = n+1$) of the system has something to do only with the present state without any influence from any other earlier state ($t \leq n-1$).

Definition 8.3.2. *For any* $n \in T$ *and states* i, $j \in I$, *the following*

$$P_{ij}\left(n\right)=P\left(X_{n+1}=j|X_n=i\right) \tag{8.6}$$

is called the moving probability of the Markov chain.

Definition 8.3.3. *If the moving probability* $P_{ij}\left(n\right)$ *in equ. (8.6) has nothing to do with* n, *then* $\{X_n : n \in T\}$ *is called a homogeneous Markov chain.*

For a homogeneous Markov chain, its moving probability $p_{ij}\left(n\right)$ is often written as p_{ij}. Because we focus on homogeneous Markov chains in this section, the word "homogeneous" is omitted from here on.

Definition 8.3.4. *Assume that* p_{ij} *stands for the moving probability as defined above. Then, the following*

$$P=[p_{ij}]=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} & \cdots \\ p_{21} & p_{22} & \cdots & p_{2n} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \tag{8.7}$$

is called the moving probability matrix for the system's state.

Proposition 8.3.1. *The elements of the moving probability matrix* P *satisfy the following properties:*

1. $p_{ij} \geq 0, i, j \in I$;

2. $\sum_{j\in I} p_{ij} = 1$, *for any* $i \in I$.

Property 2 here indicates that the sum of any chosen row of the moving probability matrix equals 1.

Definition 8.3.5. *The following*

$$p_{ij}^{(n)}=P\left(X_{m+n}=j|X_m=i\right), i, j \in I, n \geq 1,$$

is called the n*th step moving probability of the given Markov chain. And,* $P^{(n)}=\left[p_{ij}^{(n)}\right]$ *the* n*th step moving probability matrix.*

Proposition 8.3.2. *The* n*th step moving probability matrix* $P^{(n)}$ *satisfies the following properties:*

1. $p_{ij}^{(n)} \geq 0, i, j \in I$;

2. $\sum_{j\in I} p_{ij}^{(n)} = 1$, *for any* $i \in I$;
3. $P^{(n)} = P^n$.

Definition 8.3.6. *Any Markov chain with its moving probability matrix containing a grey entry(ies) is called a grey Markov chain.*

When solving a real-life problem, due to a lack of information, it is often difficult to determine the exact value of one or more entries in the moving probability matrix of a Markov chain. In such a case, the researcher will have to replace the uncertain entry by a grey interval $p_{ij}(\otimes)$ based on what is known. When the moving probability matrix is grey, in general the researcher will require the whitenization matrix $\widetilde{P}(\otimes) = \left[\widetilde{P}_{ij}(\otimes)\right]$ to satisfy:

1. $\widetilde{P}_{ij}(\otimes) \geq 0, i, j \in I$;
2. $\sum_{j\in I} \widetilde{P}_{ij} = 1$, for any $i \in I$.

Proposition 8.3.3. *Assume that the initial distribution of a finite state grey Markov chain is*

$$P^T(0) = (p_1, p_2, ..., p_n),$$

and the moving probability matrix is

$$P(\otimes) = [P_{ij}(\otimes)].$$

Then, the system distribution of the next state is given by

$$P^T(1) = P^T(0)\, P(\otimes), \tag{8.8}$$

the system distribution of the second state is

$$P^T(2) = P^T(0)\, P^2(\otimes), \tag{8.9}$$

$$\cdots\cdots\cdots\cdots\cdots,$$

and the system distribution of the s state is

$$P^T(s) = P^T(0)\, P^s(\otimes). \tag{8.10}$$

This proposition indicates that as long as one knows the initial distribution and the moving probability matrix of a system, one will be able to predict the distribution of the next state, the state after the next, and any state in the future.

8.3.2 Grey State Markov Model

Assume that $X = (x(1), x(2), ..., x(s))$ is a sequence of raw data and $\widehat{X} = (\widehat{x}(1), \widehat{x}(2), ..., \widehat{x}(s))$ is the simulated sequence using the GM(1,1) model. Then, the curve of $\widehat{X}$ reflects the trend existing in the original raw data sequence X. For a instable stochastic sequence X satisfying the conditions of a Markov chain, if we divide it into n states, then each state can be expressed as follows.

$$\otimes_i = \left[\widetilde{\otimes}_{1i}, \widetilde{\otimes}_{2i}\right], \widetilde{\otimes}_i \in \otimes_i$$

with

$$\widetilde{\otimes}_{1i} = \widehat{x}(k) + A_i, \widetilde{\otimes}_{2i} = \widehat{x}(k) + B_i,$$

for $i = 1, 2, ..., n$. Because $\widehat{X}$ is a function of time k, the grey elements $\widetilde{\otimes}_{1i}$ and $\widetilde{\otimes}_{2i}$ also change with time.

Definition 8.3.7. *If $M_{ij}(m)$ is the size of the sample of data representing the development from state $\otimes_i$ to state $\otimes_j$ through m steps, and M_i the size of the sample of data for staying at the state $\otimes_i$, then*

$$P_{ij}(m) = \frac{M_{ij}(m)}{M_i}, i = 1, 2, ..., n, \tag{8.11}$$

is called the state moving probability.

In practice, the researcher generally only needs to consider one state moving probability matrix P. Assume that the object to be predicted is located at $\otimes_k$ state. Then observe the kth row of P. If

$$\max_j p_{kj} = p_{k\ell}$$

then it will be reasonable to expect that the system will probably develop next from state $\otimes_k$ to state $\otimes_\ell$. If two or more probabilities in the kth row in the matrix P are equal or nearly so, then it will be difficult to determine to which state the system will develop next. In this case, one needs to observe the second step or the nth step moving probability matrices $P^{(2)}$ and $P^{(n)}$, where $n \geq 3$.

Example 8.3.1. Let us here look at how we did a prediction for the oil-tea production for Nan Hu Farm located in Zhejiang Province, the People's Republic of China.

The historical production data is given in Table 8.7 below.

Table 8.7. 1973–1987 Oil-tea production in 10,000 kilograms

Year	1973	1974	1975	1976	1977
Total production	25.60	14.05	32.75	4.85	6.05
Year	1978	1979	1980	1981	1982
Total production	19.10	16.50	23.80	48.00	28.50
Year	1983	1984	1985	1986	1987
Total production	28.70	12.80	35.20	60.65	3.60

1. We established a GM(1,1) model as follows.

$$\widehat{x}(k) = 480.33796 \exp(-0.0356923k) - 454.73796.$$

2. We then constructed the states. Based on the realistic situation of oil-tea production as reflected in the data, we established the following four states,

$$\begin{aligned}
&\otimes_1\colon \widetilde{\otimes}_{11} = \widehat{x}(k) - 20, \widetilde{\otimes}_{21} = \widehat{x}(k),\\
&\otimes_2\colon \widetilde{\otimes}_{12} = \widehat{x}(k) - 30, \widetilde{\otimes}_{22} = \widehat{x}(k) - 20,\\
&\otimes_3\colon \widetilde{\otimes}_{13} = \widehat{x}(k), \widetilde{\otimes}_{23} = \widehat{x}(k) + 20,\\
&\otimes_4\colon \widetilde{\otimes}_{14} = \widehat{x}(k) + 20, \widetilde{\otimes}_{24} = \widehat{x}(k) + 30,
\end{aligned}$$

$$\otimes_i = \left[\widetilde{\otimes}_{1i}, \widetilde{\otimes}_{2i}\right], \widetilde{\otimes}_i \in \otimes_i,\ i = 1, 2, 3, 4,$$

where $\widehat{x}(k)$ stands for the predicted oil-tea production at the time moment k based on our GM(1,1) model.

3. Next, we obtained the state moving probability matrix. Based on the state moving probability formula in equ. (8.11), we computed the next stage state moving probability matrix P as follows.

$$P = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

In 1987, the oil-tea production was located in state $\otimes_3$. After some development for one year, from

$$\max_j p_{3j} = p_{31}$$

we predicted that for 1988, the oil-tea production may most probably be in state $\otimes_1$.□

8.4 Combined Time Series Model

The GM(1,1) model possesses great strength when used to predict future development tendency of the system of concern, whereas time series models have special capability of describing random fluctuations of the system. So, it is natural to consider using grey time series combined models to study the governing laws of those systems whose changes are filled with both fluctuations and tendencies.

Definition 8.4.1. *For a long time series* $\{X(t)\}$,

$$X(t) = Ae^{-a(t-1)} + \sum_{i=1}^{k} \varphi_i f_i + at \tag{8.12}$$

is called a grey time series combined model. Here, the first term on the left-hand side is a part of the GM(1,1) model, which can be seen as the tendency term. The other two terms are time series model parts, which can be seen as fluctuation terms.

In this model, the GM(1,1) part is obtained using what has been learned in Chapter 7: "Grey Systems Modeling." The fluctuation parts can be obtained by modeling selected orthonormalized mean-value generation functions. The basic steps of such modeling can be described as follows.

1. Standardization of the original time series $\{X(t)\}$. Let

$$X'(t) = \frac{X(t) - \overline{X}}{\sigma}, \tag{8.13}$$

where $\overline{X}$ and σ are, respectively, the mean and standard deviation of the original series.

2. Compute the mean-value generation function f_ℓ, $\ell = 1, 2, \cdots, m$, and the period-expanding matrix F. Specifically, we have the following

$$F = [f_\ell(i)]_{n \times (m-1)}, \tag{8.14}$$

where

$$f_\ell(i) = \frac{1}{n_\ell} \sum_{j=0}^{n_\ell - 1} X'(i + j\ell), \ i = 1, 2, ..., \ell; 1 < \ell \le m \tag{8.15}$$

with n being the length of the original time series, $m = \left\lfloor \frac{n}{2} \right\rfloor$ the largest integer smaller than $\frac{n}{2}$, and $n_\ell = \left\lfloor \frac{n}{\ell} \right\rfloor$ the largest integer smaller than $\frac{n}{\ell}$.

3. Apply the Gram–Schmidt process to orthonormalize the mean-value generation functions obtained above.

First, taking $f_2(i)$ to be the initial vector for the orthonormalization of the other functions $f_3(i), f_4(i), ..., f_m(i)$ produces the sequence $\widetilde{f}_2, \widetilde{f}_3, ..., \widetilde{f}_m$. Then, take $\widetilde{f}_3$ to be the initial vector for the of the functions $\widetilde{f}_4, ..., \widetilde{f}_m$. After $m-1$ steps, the completely orthonormalized functions are denoted as follows,

$$\widetilde{F} = \left\{ \widetilde{f}_2, \widetilde{f}_3, ..., \widetilde{f}_m \right\}, m = \left\lfloor \frac{n}{2} \right\rfloor. \tag{8.16}$$

4. Using $f_\ell, \ell = 2, 3, ..., m$, as independent variables, a linear model can be established for $X'(t)$ as follows:

$$X'(t) = \sum_{i=2}^{m} \widetilde{\varphi}_i \widetilde{f}_i + e(t), \tag{8.17}$$

where $\widetilde{\varphi}_i$ are model coefficients. In the form of vectors and matrices, we have

$$X'_{n \times 1} = \widetilde{F}_{n \times (m-1)} \widetilde{\Phi}_{(m-1) \times 1}. \tag{8.18}$$

5. Evaluate the coefficients $\widetilde{\varphi}_i$. Applying the least squares estimate

$$\widetilde{\Phi} = \left(\widetilde{F}^T \widetilde{F}\right)^{-1} \widetilde{F}^T X' \tag{8.19}$$

produces the coefficient vector $\widetilde{\Phi}_{(m-1)\times 1}$.

6. Select mean-value generation functions. The magnitude of the absolute value of the coefficient $\widetilde{\varphi}_i$ in the linear model represents the level of significance of the relevant mean-value generation function. So, all the coefficients $\widetilde{\varphi}_i$ can be ordered based on the magnitudes of their absolute values, and the mean-value generation functions $\widetilde{f}_\ell$ will enter the equation according to the magnitudes of the absolute values of their coefficients $\widetilde{\varphi}_i$. The total number of mean-value generation functions entering the final equation is determined by using the following coupling score criterion,

$$CSC2_k = \frac{Q_k}{Q_x} + \frac{\lambda N}{N_k}, \tag{8.20}$$

where Q_k stands for the sum of squared residuals of the model when k mean-value functions enter into the equation, $Q_x = \frac{1}{n}\sum_{i=1}^{n}\left(X'(t) - \overline{X'}\right)^2$, k the number of independent parameters in the model, N_k the tendency score, and λ the weight of the tendency prediction. When $CSC2_k$ reaches its minimum value and N_k its maximum value, the corresponding model is optimal. By using this criterion, one determines k-value.

7. Establish the prediction model based on the selection of orthonormalized mean-value generation functions. The model coefficients selected above were determined on the basis of orthonormalized mean-value generation functions. Therefore, in order to make predictions, one needs to compute the coefficients in terms of the original mean-value generation functions $f_\ell(t)$. Now, by using the multivariate linear regression method, we establish and solve the following system of equations for the coefficients $\varphi_1, \varphi_2, \ldots$, and φ_k of our model.

$$\begin{cases} r_{11}\varphi_1 + r_{12}\varphi_2 + \cdots + r_{1k}\varphi_k & = & r_{1x} \\ r_{21}\varphi_1 + r_{22}\varphi_2 + \cdots + r_{2k}\varphi_k & = & r_{2x} \\ \cdots\cdots\cdots\cdots & & \\ r_{k1}\varphi_1 + r_{k2}\varphi_2 + \cdots + r_{kk}\varphi_k & = & r_{kx}, \end{cases} \tag{8.21}$$

where

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}, r_{ix} = \frac{s_{ix}}{\sqrt{s_{ii}s_{xx}}},$$

$$s_{ij} = \sum_{t=1}^{n}\left(f_i(t) - \widetilde{f}_i\right)\left(f_j(t) - \widetilde{f}_j\right),$$

$$s_{ix} = \sum_{t=1}^{n}\left(f_i(t) - \widetilde{f}_i\right)\left(X'(t) - \overline{X'}\right),$$

$$s_{xx} = \sum_{t=1}^{n} \left(X'(t) - \overline{X'} \right)^2 ,$$

for $i, j = 1, 2, ..., k$. Therefore, we obtain the following linear regression model in terms of the original mean-value generation functions,

$$\widehat{X'}(t) = \varphi_0 + \sum_{i=1}^{k} \varphi_i f_i(t) . \tag{8.22}$$

When equ. (8.22) is used to make predictions, as long as the original mean-value generation functions are extended a few steps into the future, the following prediction values can be obtained,

$$\widehat{X'}(n+q) = \varphi_0 + \sum_{i=1}^{k} \varphi_i f_i(n+q) , \tag{8.23}$$

where $q = 1, 2, 3, ...$ are the future time periods of predicted values. On the basis of GM(1,1) and time series models, combining the restored sequence $\widehat{X'}(t)$, as produced by the time series model, and the GM(1,1) model produces

$$\widehat{X} = \widehat{X}(t)\sigma + \widehat{y}(t) = Ae^{-a(t-1)} + \sum_{i=1}^{k} \varphi_i' f_i + a_t, \tag{8.24}$$

where

$$A = \left(1 - e^{-a}\right) \left(y^{(0)}(1) - \frac{u}{a} \right), \ a_i = \varphi_0 \sigma \text{ and } \varphi_i' = \varphi_i \sigma.$$

The optimal combined prediction model can be established by optimizing the model obtained above on the basis that the average error of the model simulation and the original data is minimum. That is, the following optimized combined model can be employed in prediction,

$$\widehat{X}(n+q) = Ae^{-a(n+q-1)} + \sum_{i=1}^{k} \varphi_i' f_i(n+q) + a_t. \tag{8.25}$$

8.5 Combined Predictions

In practice, there are generally many different methods that can be applied to make predictions for the future values of a variable. And sometimes many models can be employed to describe to a certain degree the patterns of change existing in the variable. When faced with these many "possible" models to use and various predicted values to choose from, how can the forecaster decide on which model to use and which prediction should be selected?

As a matter of fact, under this circumstance, each model reflects some aspects of the given information. Ignoring any of the potentially useful models means that some valuable information will be disregarded. At the same time, any single model selected can hardly describe the whole picture of evolution of the variable. To get around this problem, it is time for the forecaster to consider the method of combined predictions.

The idea of a combined prediction is about how to organically combine predictions produced by different models into a usable value with added accuracy. This method can very effectively employ many useful aspects of the available information so that the pattern of change of the system under consideration can be more adequately described. Each well-combined prediction method can not only avoid the pity that each single model will have to give up some aspects of the available information, but also reduce randomness and increase prediction accuracy. Currently, the idea of combined predictions has been widely employed in the practice of various predictions. However, the fundamental principle underlying the concept of combined predictions is different from that of grey combined model predictions.

8.5.1 The Combined Prediction Model

Suppose that there are n models potentially useful in the prediction of the value at time t of a variable. After applying statistical tests and non-statistical reasoning, m satisfactory models are selected from among the original n models. Now, a combination of these m models can be written as the following model.

$$\begin{array}{ll} Y = C\left\{f_1(t), f_2(t), ..., f_m(t)\right\} & \\ f_i(t) = S_t\left\{y_1, y_2, ..., y_n\right\}; & i = 1, 2, ..., m \\ s.t. \left\{\begin{array}{l} S(y_j) \in S \\ K(y_j) \in K \end{array}\right.; & j = 1, 2, ..., n, \end{array} \tag{8.26}$$

where C stands for a combination at time moment t, S_t a predicted value at time t, $S(y_j)$ the statistical test value of the jth model, S the set of threshold values selected by the researcher for statistical tests, $K(y_j)$ the knowledge and judgment about the feasibility of prediction model y_j, K the experts' collective knowledge about feasibility of predictions, and y_i the ith prediction model.

Throughout the process of making a prediction, because the internal structure of each model system participating in the prediction, may change over time, the combined prediction process can be classified into two possibilities: combined predictions with either changing structures or fixed structures.

8.5.2 Combined Predictions with Changing Structure

In this sub-section, we study combined predictions such that the internal structure of each model system, participating in the prediction, changes over time. And, structural changes can be categorized into the following three situations.

1. The number of models participating in the combined prediction changes over time;

2. The types of models participating in the combined prediction changes over time; and

3. The weights of individual models participating in the combined prediction changes over time.

For the first two situations, the forecaster needs to recombine the participating models with different weights that vary with the changing number or type of models. In terms of the third situation, the prediction values $f_i(t)$, $i = 1, 2, ..., m$, of the m participating models, their weights $\omega_i(t)$, $i = 1, 2, ..., m$, and the combined prediction value $f(t)$ are all functions of time t. In this case, the combined prediction model is

$$f(t) = \sum_{i=1}^{m} \left[\omega_i(t) f_i(t) + \varepsilon_i\right], \sum_{i=1}^{m} \omega_i(t) = 1, \tag{8.27}$$

where ε_i stands for random perturbations. Equation (8.27) is called a combined prediction model with changing weights.

Assume that at time moment t, the observational value is $x(t)$, $t = 1, 2, ..., n$. Then, the sum of squared error of the predictions from the combined prediction model in equ. (8.27) with changing weights is given by

$$J = \sum_{t=1}^{n} (x(t) - f(t))^2 = \sum_{t=1}^{n} e^2(t). \tag{8.28}$$

The optimal weights are those that minimize the sum of squared error $J = \sum_{t=1}^{n} e^2(t)$ of predictions. Denote

$$\omega(t) = (\omega_1(t), \omega_2(t), ..., \omega_m(t)), \eta_m = (1, 1, ..., 1).$$

And, let

$$e_i(t) = x(t) - f_i(t), \; i = 1, 2, ..., m, \; t = 1, 2, ..., n,$$

stand for the prediction error of the ith model at time t. Then,

$$e = \begin{bmatrix} e_1(1) & e_2(1) & \cdots & e_m(1) \\ e_1(2) & e_2(2) & \cdots & e_m(2) \\ \cdots & \cdots & \cdots & \cdots \\ e_1(n) & e_2(n) & \cdots & e_m(n) \end{bmatrix}$$

is called the matrix of prediction errors. And, $E = e^T e$ is called the information matrix generated by prediction errors.

From

$$\begin{aligned} e(t) &= x(t) - f(t) \\ &= \sum_{i=1}^{m} \omega_i(t)\, x(t) - \sum_{i=1}^{m} \omega_i(t)\, f_i(t) \\ &= \sum_{i=1}^{m} \omega_i(t)\, [x(t) - f_i(t)] \\ &= \sum_{i=1}^{m} \omega_i(t)\, e_i(t), \end{aligned}$$

it follows that the problem of optimizing the weight vector $\omega(t)$ can be reduced to the following problem of nonlinear programming:

$$\begin{cases} \min J = \omega(t)\, E\omega(t)^T \\ s.t. \begin{cases} \omega(t) \cdot \eta_m^T = 1 \\ \omega_i(t) \geq 0 \end{cases} \end{cases}$$

for $i = 1, 2, ..., m, t = 1, 2, ..., n$. The optimal solution of equ. (8.28) is

$$\omega(t) = \frac{E^{-1}\eta_m^T}{\eta_m E^{-1}\eta_m^T}, \tag{8.29}$$

which is the optimal weight vector for the combined prediction model with varying weights. The corresponding sum of squared errors of prediction is

$$\min J = \frac{1}{\eta_m E^{-1}\eta_m^T}. \tag{8.30}$$

8.5.3 Combined Predictions with Fixed Structure

A so-called combined prediction with fixed structure represents a combined prediction such that its parameters and internal system structure of the model employed do not change over time. That is, as time moves forward, the number of models, types of the models, and the weights of the models participating in the prediction do not change.

Methods of Combined Weights

Assume that $f_1, f_2, ..., f_m$ are prediction values produced from m different models, and ω_i, $i = 1, 2, ..., m$, stands for the weight of the ith model. Then,

$$f = \sum_{i=1}^{m} \omega_i f_i, \sum_{i=1}^{m} \omega_i = 1 \tag{8.31}$$

is the so-called combined prediction model with fixed weights.

The values of the weights ω_i, $i = 1, 2, ..., m$, in equ. (8.31) can be chosen in many different ways. And each specific choice of these weights leads to a specially combined model.

1. The Method of Average

Let

$$\omega_i = \frac{1}{m}, i = 1, 2, ..., m. \tag{8.32}$$

By taking the average, all models participating in the grand combined prediction model are treated equally and evenly. This is a relatively easy way to decide on the values of the model weights. This method is often used when the researcher does not have a clue as to which model participating in the final prediction has more advantage over the others.

2. The Method of Standardized Difference

Let

$$\omega_i = \frac{s - s_i}{s} \cdot \frac{1}{m-1},\ s = \sum_{i=1}^{m} s_i,\ i = 1, 2, ..., m, \tag{8.33}$$

where s_i stands for the standardized difference of the ith model. When using this method, the model with the least standardized difference will have the largest weight in the combined prediction model. That is, the researcher employs the ability to predict as the weight of each participating model. As one can expect, the standardized difference of the prediction results from the grand combined prediction model is the smallest among all those from combined models with other choices of participating models' weights.

3. The Method of Binomial Coefficients

When using the method of binomial coefficients, more weights are given to those predicted values in the middle of the predicted sequence of values. More specifically, first order f_i, $i = 1, 2, ..., m$, from the smallest to the largest. Without loss of generality, let us assume these values are ordered as follows,

$$f_1 \le f_2 \le \cdots \le f_m.$$

Then let

$$\omega_i = \frac{C_{m-1}^{i-1}}{2^{m-1}}, i = 1, 2, ..., m. \tag{8.34}$$

4. The Method of Deviation Coefficients

Assume that there are n prediction points, f_i^j the predicted value at point j made by model i, $i = 1, 2, ..., m$, $j = 1, 2, ..., n$, and f^j the average of the m predicted values at point j, $j = 1, 2, ..., n$.

Take

$$\omega_i = \frac{d - d_i}{d} \cdot \frac{1}{m-1}, i = 1, 2, ..., m, \tag{8.35}$$

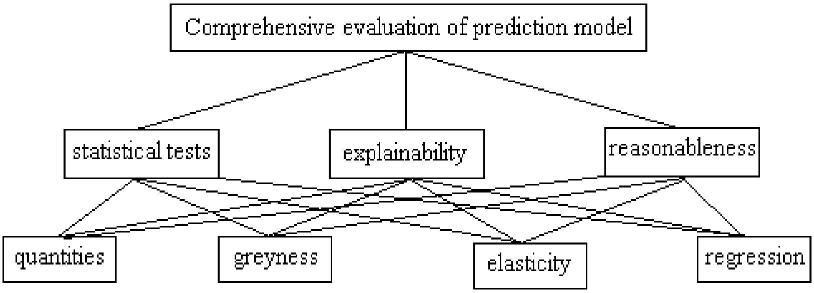

FIGURE 8.1. A graphical representation of the method of AHP

where

$$d_i = \frac{1}{n}\sqrt{\sum_{j=1}^{n}\left(f_i^j - f^j\right)^2}, i = 1, 2, ..., m,$$

are the deviation coefficients, and

$$d = \sum_{i=1}^{m} d_i.$$

5. The Method of AHP

The AHP (about high personnel) method is designed to absorb the available experiences and evaluations of relevant experts in prediction and the special field of knowledge involved. At first, these relevant experts can delete those models that are obviously not related to the problem of concern. Then, they evaluate the remaining models individually according to a set of predetermined criteria.

When using this method, the level of goals consists of evaluation criteria employed to check on individual participating models. The level of standards consists of a set of predefined requirements, including such values as degrees of acceptance for statistical tests, explainability of participating models, feasibility of predicted values, and potential realizabilities. In order to make the experts' evaluation easier, the set of standards can be further refined so that the bottom level consists of all the participating models. After computing and processing the evaluation results, one can determine the individual weights to be assigned to the participating models. At the end, the participating model with the largest weight can be employed as the final prediction model, or these weights can be used to form the grand combined prediction model, in order to produce the desired comprehensive predictions. For a graphical representation of the method of AHP, please see Figure 8.1.

6. The Method of Combined Regions

Let

$$a = \cap_{i=1}^{r} (a_i \pm \Delta a_i),$$

where Δa_i stands for the confidence interval obtained from the ith participating model at point T. This method employs the intersection of all predicted confidence intervals from each participating model as the ultimate result. However, in practice, this intersection is often empty. In this case, one can delete a predicted maximum and minimum value until the intersection is no longer empty. If by doing so, the number of remaining a_is is still greater than a predetermined value, such as $\frac{r}{2}$, then one can select the a-value as the final result. Otherwise, one will need to reevaluate each of the participating prediction models.

When Δa_i does not exist, due to the fact that some participating models could not produce confidence interval estimates, one can still delete a maximum and a minimum value one by one until the remaining distribution of predicted intervals satisfies a set of predetermined requirements. In this case, the predicted result can be used only if the number of remaining participating models is greater than a predetermined number.

Example 8.5.1. In this example, we look at how our prediction was made for the future need for human talent of a special district.

After applying statistical tests and feasibility tests, the remaining prediction models left for us to consider are listed in Table 8.8 below.

Table 8.8. Prediction models and predicted values in 10,000 people

Prediction Model	Prediction for 1995	Prediction for 2000
Macroeconomic model	98.22	132.01
Grey model 1	88.29	111.27
Grey model 2	113.41	159.07
Elasticity model 1	105.61	142.96
Elasticity model 2	100.13	143.67
Regression model	99.23	132.90

In the following, we apply each of the methods mentioned earlier to combine the predicted values.

1. The method of average

In this case, we have $\omega_i = \frac{1}{6}, i = 1, 2, ..., 6$. So, we have the following predictions for 1995 and 2000:

$$a_{1995} = 100.74,\ a_{2000} = 137.01.$$

2. The method of binomial coefficients

After ordering the individual model prediction values a_i, the values for the year of 1995 can be written as follows,

$$(88.29, 98.22, 99.23, 100.13, 105.61, 113.41),$$

and the values for the year of 2000 can be written as follows,

$$(111.27, 132.01, 132.90, 142.96, 143.67, 159.07)\,.$$

Based on equ. (8.34), it follows that the vector of weights is given as,

$$\omega = (0.03125, 0.15625, 0.3125, 0.3125, 0.15625, 0.03125)\,.$$

Therefore, the predictions out of our combined model are given by

$$a_{1995} = 100.38 and a_{2000} = 137.72.$$

3. The method of deviation coefficients

Based on relevant definitions, we can compute (using the ordering of models as listed in Table 8.8):

$$d = (1.8839, 9.5736, 8.6138, 2.4922, 2.2507, 1.7543)$$

and

$$\omega = (0.1858, 0.1279, 0.1352, 0.1812, 0.1831, 0.1867)\,.$$

So, we have our final predilections as follows,

$$a_{1995} = 100.08 and a_{2000} = 137.20.$$

4. The method of AHP

Let us identify a good statistical test result as requirement C_1, good model explainability as requirement C_2, and feasibility of predictions as requirement C_3. In order to simplify our discussion here, let us combine the two selected grey models as one class of models, two elasticity models as one, too. So, we have four models A_1, A_2, A_3, and A_4 to combine. More specifically, A_1, A_2, A_3, and A_4, respectively, stand for the selected econometric model, grey model, elasticity model and regressional model. So now, we can construct our AHP model as follows.

First, we have the following evaluation results

C_1	A_1	A_2	A_3	A_4	ω
A_1	1	7	9	1	0.43
A_2	1/7	1	3	1/3	0.08
A_3	1/9	1/3	1	1/9	0.04
A_4	1	8	9	1	0.45

where C. I. = 0.03,

C_2	A_1	A_2	A_3	A_4	ω
A_1	1	5	2	3	0.49
A_2	1/5	1	1/2	1/3	0.09
A_3	1/2	2	1	1/2	0.17
A_4	1/3	3	2	1	0.25

where C. I. = 0.04,

C_3	A_1	A_2	A_3	A_4	ω
A_1	1	9	7	2	0.56
A_2	1/9	1	1/3	1/5	0.05
A_3	1/7	3	1	1/4	0.10
A_4	1/2	8	4	1	0.30

where C. I. = 0.03, and

O	C_1	C_2	C_3	ω
C_1	1	1/5	1/2	0.12
C_2	5	1	3	0.65
C_3	2	1/3	1	0.23

where C. I. = 0.00.

Next, by combining these evaluation results, we obtain the weights $\omega = (0.50, 0.08, 0.14, 0.29)$. Therefore, we have our predictions as follows,

$$a_{1995} = 100.33,\ a_{2000} = 137.84.$$

5. The method of combined regions.

Because some models do not produce confidence interval estimates, we will apply predicted values in our combination. From Table 8.8, it follows that all extreme values were from grey models. So, let us delete these models for our further consideration. From the remaining models, the middle values will be our predictions. That is, we have

$$a_{1995} = 101.68,\ a_{2000} = 137.84.$$

Now, we summarize the results of the previous five methods of combination in Table 8.9.

Table 8.9. Results of different combination methods

Criteria			Selected Models			
Symbol	Name	ω	A_1	A_2	A_3	A_4
C_1	Stat. test	0.12	0.43	0.08	0.04	0.45
C_2	Explainability	0.65	0.49	0.09	0.17	0.25
C_3	Feasibility	0.23	0.56	0.05	0.10	0.30
Final	Combination	Weights	0.50	0.08	0.14	0.29

This table indicates that after using combined models, all predicted values converge to a small interval with obviously improved prediction accuracy.

9

Grey Prediction

9.1 Test of Grey Prediction Models

Prediction is an action based on discussions and studies of the past to tell about the future. Grey prediction is based on some theoretical treatment of the original data and establishment of grey models of the data to discover and to control the development laws of the system of interest so that scientific quantitative predictions about the future of the system can be made. All the grey models introduced in Chapter 7 can be employed as prediction models. For each specific problem, which particular model should be used as the prediction model depends on a sufficient usage of conclusions of relevant qualitative analysis. The choice of models varies from one case to another. The feasibility and qualification of a model in use need to be checked with various criteria. Only the models passing all the checks of different criteria can be used as prediction models.

Definition 9.1.1. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence of raw data, the corresponding model simulated sequence is

$$\widehat{X}^{(0)} = \left(\widehat{x}^{(0)}(1), \widehat{x}^{(0)}(2), ..., \widehat{x}^{(0)}(n)\right),$$

the sequence of errors is

$$\begin{aligned}\varepsilon^{(0)} &= (\varepsilon(1), \varepsilon(2), ..., \varepsilon(n)) \\ &= \left(x^{(0)}(1) - \widehat{x}^{(0)}(1), x^{(0)}(2) - \widehat{x}^{(0)}(2), ..., x^{(0)}(n) - \widehat{x}^{(0)}(n)\right),\end{aligned}$$

and the sequence of relative errors is

$$\begin{aligned}\Delta &= (\Delta_1, \Delta_2, ..., \Delta_n) \\ &= \left(\left|\frac{\varepsilon(1)}{x^{(0)}(1)}\right|, \left|\frac{\varepsilon(2)}{x^{(0)}(2)}\right|, ..., \left|\frac{\varepsilon(n)}{x^{(0)}(n)}\right|\right).\end{aligned}$$

1. For $k = 1, 2, ..., n$, $\Delta_k = \left|\frac{\varepsilon(k)}{x^{(0)}(k)}\right|$ *is called the relative simulation error at point k, and*

$$\overline{\Delta} = \frac{1}{n}\sum_{k=1}^{n}\Delta_k$$

the mean relative simulation error.

2. The number $1 - \overline{\Delta}$ *is called the mean relative accuracy, and* $1 - \Delta_n$ *the filtering accuracy.*

3. For a given α*, when both* $\overline{\Delta} < \alpha$ *and* $\Delta_n < \alpha$ *hold true, the model is said to be error-satisfactory.*

Definition 9.1.2. *Assume that* $X^{(0)}$ *is a sequence of original data,* $\widehat{X}^{(0)}$ *a corresponding simulated sequence, and* ε *the absolute degree of incidence of* $X^{(0)}$ *and* $\widehat{X}^{(0)}$*. If for any chosen* $\varepsilon_0 > 0$*, there exists* $\varepsilon > \varepsilon_0$*, then the model is said to be incidence-satisfactory.*

Definition 9.1.3. *Assume that* $X^{(0)}$ *is a sequence of original data,* $\widehat{X}^{(0)}$ *a corresponding model simulation sequence,* $\varepsilon^{(0)}$ *the error sequence,*

$$\overline{x} = \frac{1}{n}\sum_{k=1}^{n}x^{(0)}(k)$$

the mean of $X^{(0)}$,

$$S_1^2 = \frac{1}{n}\sum_{k=1}^{n}\left[x^{(0)}(k) - \overline{x}\right]^2$$

the variance of $X^{(0)}$

$$\overline{\varepsilon} = \frac{1}{n}\sum_{k=1}^{n}\varepsilon(k)$$

the mean error, and

$$S_2^2 = \frac{1}{n}\sum_{k=1}^{n}\left[\varepsilon(k) - \overline{\varepsilon}\right]^2$$

the variance of errors.

1. $C = \dfrac{S_2}{S_1}$ *is called the ratio of mean square deviations. For a fixed* $C_0 > 0$*, when* $C < C_0$*, the model is said to be the ratio of mean square deviation satisfactory.*

2. The following

$$p = P\left(|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 S_1\right)$$

is called a probability of small error. For a fixed $p_0 > 0$*, when* $p > p_0$*, the model is said to be small error probability satisfactory.*

The previous definitions provide us with three ways to check models in use. They judge model accuracies through observations on errors. Here, the smaller the mean relative error $\overline{\Delta}$ and the filtering error are the better; the greater the degree ε of incidence the better; the smaller the ratio C of mean square deviations the better, because small C implies that S_2 is smaller than S_1. That is, the variance of errors of simulation is small and the variance of the original data is large. It indicates that simulation errors are relative centered around a point with small amplitude in data fluctuation, and that the original data are relatively more scattered with larger amplitude in data fluctuation. So, in order to obtain good simulation results, we need to have S_2 as small as possible compared to S_1. The greater the small error probability p is the better. For each group of chosen values of a, ε_0, C_0, p_0, a level of simulation accuracy of the established model is determined. The commonly used levels of accuracy are given in Table 9.1, which is given here as a reference.

Table 9.1. Definition of critical values and their error checks

Critical Values	Relative Error	Degree of Incidences	Ratio of Mean Square Dev.	Small Error Probability
level	a	ε_0	C_0	p_0
1	0.01	0.90	0.35	0.95
2	0.05	0.80	0.5	0.80
3	0.10	0.70	0.65	0.70
4	0.20	0.60	0.80	0.60

In general, the most commonly used are the critical values of relative error.

9.2 Predictions of Sequences

Prediction of sequences is the prediction about future behavior of systems variables. The most commonly used model in grey systems theory for predictions of sequences is GM(1, 1). Based on the practical circumstances, other grey systems models can also be employed. On the foundation of qualitative analysis, define appropriate sequence operators and build GM models on the sequences obtained by applying the sequence operators. With accuracy checks, the models can be used to make predictions.

Example 9.2.1. Let us look at the prediction on revenues of the non-governmental enterprises at Changge County, Henan Province, The People's Republic of China.[1]

We have the raw data

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{4} = (10155, 12588, 23480, 35388)\,.$$

Now, we introduce a second-order weakening operator D^2 (for more details, see Chapter 4),

$$X^{(0)}D = \left(x^{(0)}(1)d, x^{(0)}(2)d, x^{(0)}(3)d, x^{(0)}(4)d\right),$$

where

$$x^{(0)}(k)d = \frac{1}{4-k+1}[x^{(0)}(k) + x^{(0)}(k+1) + \cdots + x^{(0)}(4)],$$

$k = 1, 2, 3, 4$; and

$$X^{(0)}D^2 = \left(x^{(0)}(1)d^2, x^{(0)}(2)d^2, x^{(0)}(3)d^2, x^{(0)}(4)d^2\right),$$

where

$$x^{(0)}(k)d^2 = \frac{1}{4-k+1}[x^{(0)}(k)d + x^{(0)}(k+1)d + \cdots + x^{(0)}(4)d],$$

$k = 1, 2, 3, 4$. Then, we have

$$X^{(0)}D^2 = (27260, 29547, 32411, 35388)\,,$$

which is written as

$$X^{(0)}D^2 = X = (x(1), x(2), x(3), x(4))\,.$$

Now, the 1-AGO sequence of $X^{(0)}$ is

$$X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{4} = (27260, 56807, 89218, 124606)\,.$$

Assume

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b.$$

[1] The relevant data came from Changge County Statistics Bureau.

Based on the least squares method, we obtain the estimated values for a and b as follows,

$$\widehat{a} = -0.089995, \widehat{b} = 25790.28.$$

So, the resultant GM(1, 1) model is given by

$$\frac{dx^{(1)}}{dt} - 0.089995x^{(1)} = 25790.28,$$

with its time response sequence being

$$\begin{cases} \widehat{x}^{(1)}(k+1) &= 313834e^{0.089995k} - 286574 \\ \widehat{x}^{(0)}(k+1) &= \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k). \end{cases}$$

From these results, we obtain the simulated sequence

$$\widehat{X} = (\widehat{x}(i))_{i=1}^{4} = (27260, 29553, 32337, 35381)$$

with the sequence of errors

$$\varepsilon^{(0)} = \left(\varepsilon^{(0)}(i)\right)_{i=1}^{4} = (0, -6, 74, 7),$$

the sequence of relative errors

$$\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4) = (0, 0.0002, 0.00228, 0.0002),$$

the mean relative error

$$\begin{aligned}\overline{\Delta} &= \tfrac{1}{4}\sum_{k=1}^{4}\Delta_k = \tfrac{1}{4}(0 + 0.0002 + 0.00228 + 0.0002) \\ &= 0.00067 = 0.067\% < 0.01,\end{aligned}$$

and the filtering error

$$\Delta_4 = 0.0002 = 0.02\% < 0.01.$$

So, the accuracy of our simulation is in level one.

Now, we compute the absolute degree ε of grey incidences of X and $\widehat{X}$.

$$\begin{aligned}|s| &= \left|\sum_{k=2}^{3}[x(k) - x(1)] + \tfrac{1}{2}[x(4) - x(1)]\right| \\ &= \left|2287 + 5151 + \tfrac{1}{2}\cdot 8128\right| = 11502, \\ |\widehat{s}| &= \left|\sum_{k=2}^{3}[\widehat{x}(k) - \widehat{x}(1)] + \tfrac{1}{2}[\widehat{x}(4) - \widehat{x}(1)]\right| \\ &= \left|2293 + 5077 + \tfrac{1}{2}\cdot 8121\right| = 11430.5,\end{aligned}$$

$$\begin{aligned}|\widehat{s}-s| &= |\sum_{k=2}^{3}\{[\widehat{x}(k)-\widehat{x}(1)]-[x(k)-x(1)]\} \\ &\quad +\tfrac{1}{2}\{[\widehat{x}(4)-\widehat{x}(1)]-[x(4)-x(1)]\}| \\ &= \left|\{2293-2287\}+\{5077-5151\}+\tfrac{1}{2}\cdot\{8121-8128\}\right| \\ &= \left|6+(-74)+(-\tfrac{7}{2})\right| = 71.5.\end{aligned}$$

So,

$$\varepsilon = \frac{1+|s|+|\widehat{s}|}{1+|s|+|\widehat{s}|+|s-\widehat{s}|} = 0.997 > 0.90.$$

That is, the degree of incidence is in level one.

Compute the ratio of mean square deviations C:

$$\overline{x} = \frac{1}{4}\sum_{k=1}^{4} x(k) = 31151.5,$$

$$S_1^2 = \frac{1}{4}\sum_{k=1}^{4}[x(k)-\overline{x}]^2 = 37252465,$$

$$S_1 = 6103.48,$$

$$\overline{\varepsilon} = \frac{1}{4}\sum_{k=1}^{4}\varepsilon(k) = 18.75,$$

$$S_2^2 = \frac{1}{4}\sum_{k=1}^{4}[\varepsilon(k)-\overline{\varepsilon}]^2 = 4154.75,$$

$$S_2 = 64.46.$$

It follows that

$$C = \frac{S_2}{S_1} = \frac{64.46}{6103.48} = 0.01 < 0.35,$$

which is in the level one.

Compute the small error probability. From

$$0.6745S_1 = 0.6745\cdot 6103.48 = 4116.80$$

and

$$|\varepsilon(1)-\overline{\varepsilon}| = 18.75, |\varepsilon(2)-\overline{\varepsilon}| = 24.75,$$

$$|\varepsilon(3)-\overline{\varepsilon}| = 55.25, |\varepsilon(4)-\overline{\varepsilon}| = 11.75,$$

it follows that

$$P(|\varepsilon(k)-\overline{\varepsilon}| < 0.6745S_1) = 1 > 0.95,$$

which belongs to level one.

So, with our accuracy checks in place, we can apply the grey model

$$\begin{cases} \widehat{x}^{(1)}(k+1) &= 313834e^{0.089995k} - 286574 \\ \widehat{x}^{(0)}(k+1) &= \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k) \end{cases}$$

to make predictions. Here, we list five predicted values as follows:

$$\widehat{X}^{(0)} = \left(\widehat{x}^{(0)}(i)\right)_{i=5}^{9} = (38714, 42359, 46348, 50712, 55488)\,.$$

9.3 Interval Predictions

For chaotic data sequences, where no matter which model is used, it is difficult for the simulation outcome to pass accuracy tests; we will not be able to produce exact prediction values. In this case, we can consider deriving a range for the future changes, and predict the interval of possible values.

Definition 9.3.1. *Assume that $X(t)$ is the zigzagged line of a sequence, and that $f_\ell(t)$ and $f_u(t)$ are continuous smooth curves. If for any t, the following always holds true,*

$$f_\ell(t) \leq X(t) \leq f_u(t),$$

then $f_\ell(t)$ is called a lower bound (function) of $X(t)$, and $f_u(t)$ a upper bound (function) of $X(t)$; and

$$S = \{(t, \overline{X}(t)) | \overline{X}(t) \in [f_\ell(t), f_u(t)]\}$$

is called the value band of $X(t)$.

Definition 9.3.2. *1. If the lower and upper boundary functions of the value band S of $X(t)$ are the same kind of function, then S is called a uniform band.*

2. When S is a uniform band with the lower bound function $f_\ell(t)$ and upper bound function $f_u(t)$ being exponential, the band S is called a uniform exponential band, or exponential band for short.

3. When S is a uniform band with the lower bound function $f_\ell(t)$ and upper bound function $f_u(t)$ being linear, S is called a uniform linear band, or linear band for short.

4. If $t_1 < t_2$ always implies that

$$f_u(t_1) - f_\ell(t_1) < f_u(t_2) - f_\ell(t_2),$$

then S is called a trumpet-like band.

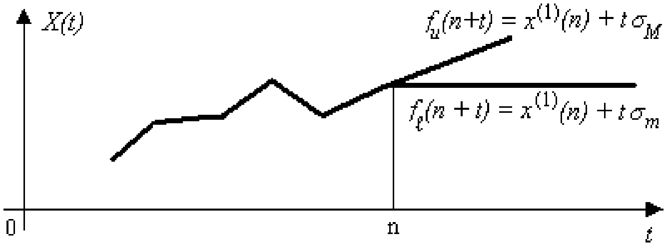

FIGURE 9.1. The prediction range of $X^{(1)}$

Definition 9.3.3. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence of original data with the 1-AGO sequence

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\right).$$

Let

$$\sigma_M = \max_{1\leq k\leq n}\{x^{(0)}(k)\}, \sigma_m = \min_{1\leq k\leq n}\{x^{(0)}(k)\}$$

and take the lower bound function $f_\ell(n+t)$ and the upper bound function $f_u(n+t)$ of $X^{(1)}$ as follows, respectively,

$$f_\ell(n+t) = x^{(1)}(n) + t\sigma_m,$$

and

$$f_u(n+t) = x^{(1)}(n) + t\sigma_M.$$

Then

$$S = \{(t, X(t))|t > n, X(t) \in [f_\ell(t), f_u(t)]\}$$

is called a proportional band.

Proposition 9.3.1. *Each proportional band is a straight-line trumpet-like band.*

In fact, the lower bound function and upper bound function of each proportional band are straight-lines, which are proportionally increasing with time, with slopes σ_m and σ_M respectively. The prediction range of $X^{(1)}$ is shown in Figure 9.1.

Let $X^{(0)}$ be a sequence of original data with time variable t. The curves, one of which connects all the local minimum points and the other connects all the local maximum points in the zigzagged curve of the sequence $X^{(0)}$, are called the lower bound curve and the upper bound curve of the sequence $X^{(0)}$, respectively.

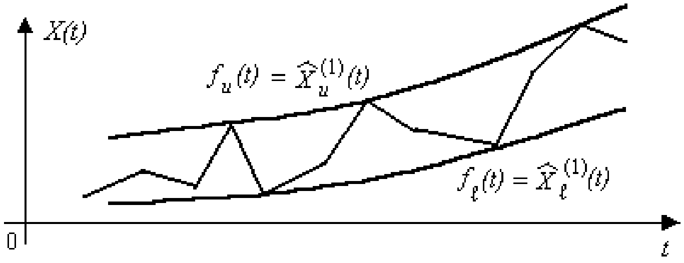

FIGURE 9.2. How a wrapping band looks like

Definition 9.3.4. *Assume that $X^{(0)}$ is a sequence of raw data, $X_\ell^{(0)}$ the sequence corresponding to the lower bound curve of $X^{(0)}$, and $X_u^{(0)}$ the sequence corresponding to the upper bound curve of $X^{(0)}$. If*

$$\widehat{x}_\ell^{(1)}(k+1) = \left[x_\ell^{(0)}(1) - \frac{b_\ell}{a_\ell}\right] \cdot e^{-a_\ell k} + \frac{b_\ell}{a_\ell}$$

and

$$\widehat{x}_u^{(1)}(k+1) = \left[x_u^{(0)}(1) - \frac{b_u}{a_u}\right] \cdot e^{-a_u k} + \frac{b_u}{a_u}$$

are, respectively, the GM(1, 1) time response sequences of $X_\ell^{(0)}$ and $X_u^{(0)}$, then

$$S = \left\{(t, X(t)) | X(t) \in [\widehat{x}_\ell^{(1)}(t), \widehat{x}_u^{(1)}(t)]\right\}$$

is called a wrapping band.

A wrapping band is given here in Figure 9.2.

Definition 9.3.5. *Assume that $X^{(0)}$ is a sequence of raw data, m different subsequences of $X^{(0)}$ can be used to establish m different GM(1, 1) models. Assume that the corresponding parameters are*

$$\widehat{a}_i = \left[\begin{array}{cc} a_i & b_i \end{array}\right]^T, \ i = 1, 2, ..., m.$$

Let

$$-\sigma_M = \max_{1 \le k \le n}\{-a_i\}, -\sigma_m = \min_{1 \le k \le n}\{-a_i\},$$

and

$$\widehat{x}_\ell^{(1)}(k+1) = \left[x_\ell^{(0)}(1) - \frac{b_m}{a_m}\right] \cdot e^{-a_m k} + \frac{b_m}{a_m},$$

$$\widehat{x}_u^{(1)}(k+1) = \left[x_u^{(0)}(1) - \frac{b_M}{a_M}\right] \cdot e^{-a_M k} + \frac{b_M}{a_M}.$$

Then

$$S = \left\{(t, X(t)) | X(t) \in [\widehat{x}_\ell^{(1)}(t), \widehat{x}_u^{(1)}(t)]\right\}$$

is called a development band.

Proposition 9.3.2. *Wrapping and development bands are all exponential bands.*

Definition 9.3.6. *Assume that*

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\right)$$

is a sequence of raw data, and $f_\ell(t)$ and $f_u(t)$ are the lower bound and upper bound functions of the 1-AGO sequence $X^{(1)}$. Then for any $k > 0$,

$$\widehat{x}_\ell^{(1)}(n+k) = f_\ell(n+k)$$

is called the lowest predicted value,

$$\widehat{x}_u^{(1)}(n+k) = f_u(n+k)$$

is called the highest predicted value, and

$$\widehat{x}^{(1)}(n+k) = \frac{1}{2}[f_\ell(n+k) + f_u(n+k)]$$

is called the basic predicted value.

Example 9.3.1. The following sequence gives the retailing amounts of cotton cloth in Henan Province, the People's Republic of China,

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{6} = (4.9445, 5.5828, 5.3441, 5.2669, 4.5640, 3.6524),$$

where the dimension of $x^{(0)}(k)$ is 0.1 billion meters, $x^{(0)}(1) = 4.9445$ is the datum for the year of 1978, and $x^{(0)}(2) = 5.5828$ for the year of 1979, ... Make a proportional band prediction based on this set of data.

Solution: Let

$$\sigma_M = \max_{1\le k\le 6}\{x^{(0)}(k)\} = 5.5828,$$

and

$$\sigma_m = \min_{1\le k\le 6}\{x^{(0)}(k)\} = 3.6524.$$

From

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i),$$

$k = 1, 2, ..., 6$, we obtain the 1-AGO sequence of $X^{(0)}$ as follows.

$$\begin{aligned} X^{(1)} &= \left(x^{(1)}(i)\right)_{i=1}^{6} \\ &= (4.9445, 10.5273, 15.8714, 21.1383, 25.7023, 29.3547). \end{aligned}$$

So

$$f_u(6+k) = x^{(1)}(6) + k\sigma_M = 29.3547 + 5.5828k$$

and

$$f_\ell(6+k) = x^{(1)}(6) + k\sigma_m = 29.3547 + 3.6524k.$$

When $k = 1, 2, 3$, the highest predicted values are

$$\widehat{x}_u^{(1)}(7) = f_u(6+1) = x^{(1)}(6) + 1 \cdot \sigma_M = 34.9375,$$

$$\widehat{x}_u^{(1)}(8) = f_u(6+2) = x^{(1)}(6) + 2 \cdot \sigma_M = 40.5203,$$

$$\widehat{x}_u^{(1)}(9) = f_u(6+3) = x^{(1)}(6) + 3 \cdot \sigma_M = 46.1031;$$

the lowest predicted values are

$$\widehat{x}_\ell^{(1)}(7) = f_\ell(6+1) = x^{(1)}(6) + 1 \cdot \sigma_m = 33.0071,$$

$$\widehat{x}_\ell^{(1)}(8) = f_\ell(6+2) = x^{(1)}(6) + 2 \cdot \sigma_m = 36.6595,$$

$$\widehat{x}_\ell^{(1)}(9) = f_\ell(6+3) = x^{(1)}(6) + 3 \cdot \sigma_m = 40.3119;$$

and the basic predicted values are

$$\widehat{x}^{(1)}(7) = \tfrac{1}{2}[\widehat{x}_u^{(1)}(7) + \widehat{x}_\ell^{(1)}(7)] = 33.9723,$$

$$\widehat{x}^{(1)}(8) = \tfrac{1}{2}[\widehat{x}_u^{(1)}(8) + \widehat{x}_\ell^{(1)}(8)] = 38.5899,$$

$$\widehat{x}^{(1)}(9) = \tfrac{1}{2}[\widehat{x}_u^{(1)}(9) + \widehat{x}_\ell^{(1)}(9)] = 43.2075.$$

Example 9.3.2. The following sequence gives the data of per capita consumption of cotton cloths of peasants in Henan Province,

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{6} = (5.43, 3.90, 3.93, 4.43, 3.97, 2.77)\,,$$

where the dimension of $x^{(0)}(k)$ is meter, for $k = 1$, 2, ..., 6, $x^{(0)}(1) = 5.43$ is the datum for the year of 1978, $x^{(0)}(2) = 3.90$ for the year of 1979, ..., $x^{(0)}(6) = 2.77$ for the year of 1983. Do a wrapping band prediction.

Solution: The zigzagged line $X^{(0)}$, its upper wrapping curve $f_u(t)$, and the lower wrapping curve $f_\ell(t)$ are shown in Figure9.3.

The upper wrapping sequence, corresponding to $f_u(t)$, is given as

$$X_u^{(0)} = \left(x_u^{(0)}(i)\right)_{i=1}^{6} = (5.43, 4.97, 4.67, 4.43, 3.33, 3.90)\,,$$

and its 1-AGO sequence is

$$X_u^{(1)} = \left(x_u^{(1)}(i)\right)_{i=1}^{6} = (5.43, 10.40, 15.06, 91.50, 23.63, 27.53)\,.$$

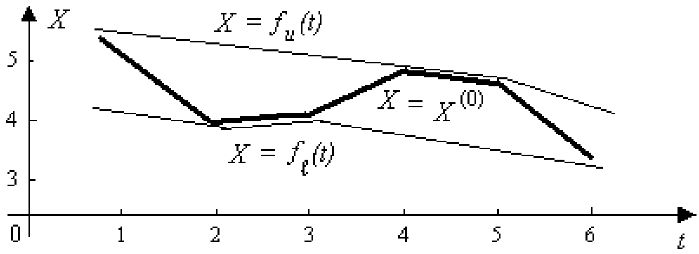

FIGURE 9.3. A zigzagged line and its upper and lower wrapping functions

GM(1, 1) time response sequence of $X_u^{(1)}$ is

$$\widehat{x}_u^{(1)}(k+1) = -84.83 \cdot e^{-0.06k} + 90.26.$$

So, the inverse accumulating restoration sequence $X_u^{(0)}$ of $X_u^{(1)}$ is

$$\begin{aligned}\widehat{x}_u^{(0)}(k+1) &= \widehat{x}_u^{(1)}(k+1) - \widehat{x}_u^{(1)}(k) \\ &= -84.83 \cdot e^{-0.06k} + 84.83 \cdot e^{-0.06k+0.06} \\ &= 84.83 \cdot (e^{0.06} - 1) \cdot e^{-0.06k} \\ &= 5.254 \cdot e^{-0.06k}.\end{aligned}$$

So, we obtain the highest predicted values

$$\widehat{x}_u^{(0)}(7) = 3.66, \widehat{x}_u^{(0)}(8) = 3.45, \widehat{x}_u^{(0)}(9) = 3.25.$$

The lower wrapping sequence, corresponding to $f_\ell(t)$, is given by

$$X_\ell^{(0)} = \left(x_\ell^{(0)}(i)\right)_{i=1}^{6} = (3.87, 3.90, 3.53, 3.33, 3.07, 3.77)\,,$$

and its 1-AGO sequence $X_\ell^{(1)}$ is

$$X_\ell^{(1)} = \left(x_\ell^{(1)}(i)\right)_{i=1}^{6} = (3.87, 7.77, 11.30, 14.83, 17.70, 20.47)\,.$$

So, we have the GM(1, 1) time response sequence

$$\begin{cases}\widehat{x}_\ell^{(1)}(k+1) &= -48.33 \cdot e^{-0.084k} + 52.20 \\ \widehat{x}_\ell^{(0)}(k+1) &= \widehat{x}_\ell^{(1)}(k+1) - \widehat{x}_\ell^{(1)}(k).\end{cases}$$

That is

$$\widehat{x}_\ell^{(0)}(k+1) = 4.24 \cdot e^{-0.084k}.$$

So, it follows that the lowest predicted values are:

$$\widehat{x}_\ell^{(0)}(7) = 2.56, \widehat{x}_\ell^{(0)}(8) = 2.35, \widehat{x}_\ell^{(0)}(9) = 2.16.$$

Therefore, we have the basic predicted values:

$$\widehat{x}^{(0)}(7) = \frac{1}{2}[\widehat{x}_u^{(0)}(7) + \widehat{x}_\ell^{(0)}(7)] = 3.112,$$

$$\widehat{x}^{(0)}(8) = \frac{1}{2}[\widehat{x}_u^{(0)}(8) + \widehat{x}_\ell^{(0)}(8)] = 2.90,$$

$$\widehat{x}^{(0)}(9) = \frac{1}{2}[\widehat{x}_u^{(0)}(9) + \widehat{x}_\ell^{(0)}(9)] = 2.705.$$

Example 9.3.3. Let us see the wrapping GM(1, 1) prediction for the output of 20 major experimental materials used at Henan Agriculture University.[2]

The following gives various seasonal output data during 1990 $\sim$ 1996 of 20 major experimental materials,

$$X_{ij}^{(0)} = \left(x_{ij}^{(0)}(1), x_{ij}^{(0)}(2), x_{ij}^{(0)}(3), x_{ij}^{(0)}(4), x_{ij}^{(0)}(5), x_{ij}^{(0)}(6), x_{ij}^{(0)}(7)\right),$$

$i = 1, 2, ..., 20$; $j = 1, 2, 3, 4$, where

Table 9.2. The variables defined

Name	Unit	Code	Name	Unit	Code
Sodium hydroxide	Bottle	X_1	100 ml beaker	Number	X_{11}
Sodium chloride	Bottle	X_2	250 ml beaker	Number	X_{12}
Hydroch- loric acid	Bottle	X_3	500 ml beaker	Number	X_{13}
Sulphuric acid	Bottle	X_4	2000 ml beaker	Number	X_{14}
Agar	Kg	X_5	100 ml flask	Number	X_{15}
Carbon te- trachloride	Bottle	X_6	125 ml rea- gent bottle	Number	X_{16}
Bitoluene	Bottle	X_7	500 ml rea- gent bottle	Number	X_{17}
Absolute alcohol	Bottle	X_8	60 ml jar	Number	X_{18}
PH test paper	Box	X_9	60 ml drop- ping bottle	Number	X_{19}
Potassium hydroxide	Bottle	X_{10}	9 cm cul- ture dish	Number	X_{20}

[2] For more details please refer to the work of Zhao, D.Y [1989].

Find the relevant upper wrapping sequence

$$X_{ij_u}^{(0)} = \left(x_{ij_u}^{(0)}(1), x_{ij_u}^{(0)}(2), x_{ij_u}^{(0)}(3), x_{ij_u}^{(0)}(4), x_{ij_u}^{(0)}(5), x_{ij_u}^{(0)}(6), x_{ij_u}^{(0)}(7)\right),$$

$i = 1, 2, ..., 20; j = 1, 2, 3, 4$, and the lower wrapping sequence

$$X_{ij_\ell}^{(0)} = \left(x_{ij_\ell}^{(0)}(1), x_{ij_\ell}^{(0)}(2), x_{ij_\ell}^{(0)}(3), x_{ij_\ell}^{(0)}(4), x_{ij_\ell}^{(0)}(5), x_{ij_\ell}^{(0)}(6), x_{ij_\ell}^{(0)}(7)\right),$$

$i = 1, 2, ..., 20; j = 1, 2, 3, 4$, to establish a upper wrapping GM(1, 1) model

$$\widehat{x}_{ij_u}^{(1)}(k+1) = \left[x_{ij_u}^{(0)}(1) - \frac{b_{ij_u}}{a_{ij_u}}\right] \cdot e^{-a_{ij_u}k} + \frac{b_{ij_u}}{a_{ij_u}}$$

and a lower wrapping GM(1, 1) model

$$\widehat{x}_{ij_\ell}^{(1)}(k+1) = \left[x_{ij_\ell}^{(0)}(1) - \frac{b_{ij_\ell}}{a_{ij_\ell}}\right] \cdot e^{-a_{ij_\ell}k} + \frac{b_{ij_\ell}}{a_{ij_\ell}},$$

$i = 1, 2, ..., 20; j = 1, 2, 3, 4$. Their restored values through inverse accumulating are given respectively as follows.

$$\widehat{x}_{ij_u}^{(0)}(k+1) = \widehat{x}_{ij_u}^{(1)}(k+1) - \widehat{x}_{ij_u}^{(1)}(k)$$

and

$$\widehat{x}_{ij_\ell}^{(0)}(k+1) = \widehat{x}_{ij_\ell}^{(1)}(k+1) - \widehat{x}_{ij_\ell}^{(1)}(k),$$

$i = 1, 2, ..., 20, j = 1, 2, 3, 4$.

Here, there are 160 different models. To save some space, the specific models, the upper and lower wrapping sequences of the output sequences of the seasonal data of the 20 major experimental materials are omitted here. Because the basic predicted values

$$\widehat{x}_{ij}^{(0)}(k) = \frac{1}{2}[\widehat{x}_{ij_u}^{(0)}(k) + \widehat{x}_{ij_\ell}^{(0)}(k)],$$

$i = 1, 2, ..., 20; j = 1, 2, 3, 4$, are calculated from those of $\widehat{x}_{ij_u}^{(0)}(k)$ and $\widehat{x}_{ij_\ell}^{(0)}(k)$, they are also omitted here. We only list the highest and lowest predicted values for various seasons of the 20 major experimental materials for the year 1997 in Table 9.3.

Table 9.3. Predicted high and low values for each variable

Seasons	1		2		3		4	
Variables	hi	lo	hi	lo	hi	lo	hi	lo
X_1	20	85	32	70	32	101	15	89
X_2	20	45	10	64	6	45	20	60
X_3	32	190	61	180	20	200	40	150
X_4	112	190	60	160	15	124	18	218
X_5	12	76	20	55	5	39	6	89
X_6	15	75	8	65	16	60	20	65
X_7	12	130	25	110	9	89	12	74
X_8	8	48	20	220	32	140	20	132
X_9	132	398	126	658	112	626	198	598
X_{10}	6	40	10	33	10	47	5	50
X_{11}	18	331	11	172	36	200	18	126
X_{12}	76	388	35	230	25	74	15	634
X_{13}	94	277	21	215	24	680	30	225
X_{14}	34	81	5	64	10	132	20	61
X_{15}	30	320	27	289	50	250	56	480
X_{16}	11	205	15	238	10	160	8	200
X_{17}	9	366	10	230	8	321	25	227
X_{18}	10	204	12	216	25	362	12	229
X_{19}	12	981	84	1123	71	1411	91	1771
X_{20}	61	500	25	426	22	386	11	72

9.4 Disaster Predictions

Essentially, disaster prediction is a prediction for abnormal values. Then, the first question is: What kinds of values are abnormal? In general, people use their experience and subjective criteria to determine what values are normal and what are not. The task for disaster predictions is to pinpoint the time moment(s) for one or several abnormal values to occur so that relevant parties can have enough time to make preparations for disasters to come.

Definition 9.4.1. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data. For a given upper abnormal (or catastrophe) value ξ, *the sub-sequence of* X,

$$\begin{aligned} X_\xi &= (x[q(1)], x[q(2)], ..., x[q(m)]) \\ &= \{x[q(i)] |\ x[q(i)] \geq \xi, i = 1, 2, ..., m\}, \end{aligned}$$

is called an upper catastrophe sequence.

Definition 9.4.2. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data. For a given lower abnormal (or catastrophe) value ζ, *the sub-sequence of* X,

$$\begin{aligned} X_\zeta &= (x[q(1)], x[q(2)], ..., x[q(\ell)]) \\ &= \{x[q(i)] | x[q(i)] \le \zeta, i = 1, 2, ..., \ell\}, \end{aligned}$$

is called a lower catastrophe sequence.

The upper and lower abnormal sequences are called *catastrophe sequences*. Because different catastrophe sequences require different approaches to handle related details, in the following discussions, we do not distinguish between the upper and lower catastrophe sequences.

Definition 9.4.3. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data, and

$$X_\xi = (x[q(1)], x[q(2)], ..., x[q(m)]) \subset X$$

a catastrophe sequence of X. Then

$$Q^{(0)} = (q(1), q(2), ..., q(m))$$

is called a catastrophe date sequence.

The so-called disaster prediction is about finding patterns, if any, through the study of catastrophe date sequences in order to predict future dates of occurrences of catastrophes. In grey systems theory, each disaster prediction is realized or done through establishing GM(1, 1) models for relevant catastrophe date sequences.

Definition 9.4.4. *Assume that*

$$Q^{(0)} = (q(1), q(2), ..., q(m))$$

is a catastrophe date sequence, and its 1-AGO sequence is

$$Q^{(1)} = \left(q^{(1)}(1), q^{(1)}(2), ..., q^{(1)}(m)\right),$$

where

$$q^{(1)}(i) = \sum_{j=1}^{i} q(j),$$

$i = 1, 2, ..., m$. The sequence mean generated based on consecutive neighbors of $Q^{(1)}$ is

$$Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), ..., z^{(1)}(m)\right),$$

where $z^{(1)}(1) = q^{(1)}(1)$, and

$$z^{(1)}(i) = \frac{1}{2}[q^{(1)}(i) + q^{(1)}(i-1)],$$

$i = 2, 3, ..., m$. Then,

$$q^{(0)}(k) + az^{(1)}(k) = b$$

is called a catastrophe $GM(1,1)$ model.

Proposition 9.4.1. *Assume that $\widehat{a} = \begin{bmatrix} a & b \end{bmatrix}^T$ is the least squares estimate of the parameters in a catastrophe $GM(1,1)$ model. Then the $GM(1,1)$ ordinality response sequence of the catastrophe date sequence is*

$$\begin{cases} \widehat{q}^{(1)}(k+1) & = \left[q(1) - \frac{b}{a}\right] \cdot e^{-ak} + \frac{b}{a} \\ \widehat{q}(k+1) & = \widehat{q}^{(1)}(k+1) - \widehat{q}^{(1)}(k). \end{cases}$$

That is

$$\begin{aligned} \widehat{q}(k+1) &= \left[q(1) - \tfrac{b}{a}\right] \cdot e^{-ak} - \left[q(1) - \tfrac{b}{a}\right] \cdot e^{-a(k-1)} \\ &= (1 - e^{a}) \cdot \left[q(1) - \tfrac{b}{a}\right] \cdot e^{-ak}. \end{aligned}$$

Definition 9.4.5. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data with n being the present time moment. For a given abnormal value ξ, the corresponding catastrophe date sequence is given as

$$Q^{(0)} = (q(1), q(2), ..., q(m)),$$

where $q(m)(\leq n)$ represents the date of the last catastrophe that had occurred. Then, $\widehat{q}(m+1)$ is the predicted date for the next disaster to occur. For any $k > 0$, $\widehat{q}(m+k)$ is the predicted date for the kth disaster to occur in the future.

Example 9.4.1. The following sequence gives the annual average precipitations (in the unit of mm) of a certain region

$$\begin{aligned} X &= (x(i))_{i=1}^{17} \\ &= (390.6, 412, 320, 559.2, 380.8, 542.4, 553, 310, 561, \\ &\quad 300, 632, 540, 406.2, 313.8, 576, 587.6, 318.5), \end{aligned}$$

where $x(1), x(2), ..., x(17)$ are respectively the data for the years of 1980, 1981, ..., 1996. Take $\xi = 320$ mm as a lower abnormal (drought) value. Do a drought prediction for this special region of our study.

Solution: Let $\xi = 320$. We obtain the following lower catastrophe sequence

$$\begin{aligned} X_\xi &= (x(3), x(8), x(10), x(14), x(17)) \\ &= (320, 310, 300, 313.8, 318.5), \end{aligned}$$

with the corresponding catastrophe date sequence

$$Q^{(0)} = (q(i))_{i=1}^{5} = (3, 8, 10, 14, 17),$$

and its 1-AGO sequence $Q^{(1)}$

$$Q^{(1)} = \left(q^{(1)}(i)\right)_{i=1}^{5} = (3, 11, 21, 35, 52).$$

The sequence mean generated based on consecutive neighbors of $Q^{(1)}$ is given by

$$Z^{(1)} = \left(z^{(1)}(i)\right)_{i=1}^{5} = (3, 7, 16, 28, 43.5).$$

Let

$$q^{(0)}(k) + az^{(1)}(k) = b.$$

From

$$B = \begin{bmatrix} -7 & 1 \\ -16 & 1 \\ -28 & 1 \\ -43.5 & 1 \end{bmatrix} \text{ and } Y = \begin{bmatrix} 8 \\ 10 \\ 14 \\ 17 \end{bmatrix},$$

it follows that

$$\widehat{a} = \left[\begin{array}{cc} a & b \end{array}\right]^T = \left[B^T B\right]^{-1} B^T Y = \left[\begin{array}{cc} -0.25361 & 6.258339 \end{array}\right]^T.$$

Therefore, the GM(1, 1) ordinality response sequence of the catastrophe date sequence is

$$\begin{cases} \widehat{q}^{(1)}(k+1) &= 27.677 \cdot e^{0.25361k} - 24.677 \\ \widehat{q}(k+1) &= \widehat{q}^{(1)}(k+1) - \widehat{q}^{(1)}(k); \end{cases}$$

that is

$$\begin{aligned} \widehat{q}(k+1) &= 27.677 \cdot e^{0.25361k} - 27.677 \cdot e^{0.25361(k-1)} \\ &= 27.677 \cdot (1 - e^{-0.25361}) \cdot e^{0.25361k} \\ &= 6.1998 \cdot e^{0.25361k}. \end{aligned}$$

So, we can obtain a simulated sequence for $Q^{(0)}$ as follows,

$$\widehat{Q}^{(0)} = (\widehat{q}(i))_{i=1}^{5} = (6.1998, 7.989, 10.296, 13.268, 17.098)\,.$$

From

$$\varepsilon(k) = q(k) - \widehat{q}(k),$$

$k = 1, 2, 3, 4, 5$, we obtain the error sequence as follows,

$$\varepsilon^{(0)} = (\varepsilon(i))_{i=1}^{5} = (-3.1998, 0.011, -0.296, 0.732, -0.098)\,.$$

And from

$$\Delta_k = \left|\frac{\varepsilon(k)}{q(k)}\right|,$$

$k = 2, 3, 4, 5$, it follows that the sequence of relative errors is

$$\Delta = (\Delta_2, \Delta_3, \Delta_4, \Delta_5) = (0.1\%, 2.96\%, 5.1\%, 0.6\%)\,.$$

From this sequence, we calculate the average relative error

$$\overline{\Delta} = \frac{1}{4}\sum_{k=2}^{5}\Delta_k = 2.19\%,$$

with $1-\overline{\Delta} = 97.81\%$ as the average relative accuracy, and $1 - \Delta_5 = 99.4\%$ as the filtering accuracy. So, we can use

$$\widehat{q}(k+1) = 6.1998 \cdot e^{0.25361k}$$

to do our predictions. Because

$$\widehat{q}(5+1) = \widehat{q}(6) \approx 22,$$

and

$$\widehat{q}(6) - \widehat{q}(5) \approx 22 - 17 = 5,$$

we predict that in four years, counting from the time of the last drought, in 2001 there might be a drought. In order to improve the accuracy of our prediction, we can take several different abnormal values to build various models to make predictions.

9.5 Seasonal Disaster Predictions

In this section, we look at situations where seasonal patterns exist naturally.

Definition 9.5.1. *Assume that* $\Omega = [a, b]$ *is the overall time interval of a study. If*

$$\omega_i = [a_i, b_i] \subset \Omega = [a, b],$$

$i = 1, 2, ..., s$, *satisfy*

1. $\Omega = \bigcup_{i=1}^{s} \omega_i$;
2. $\omega_i \cap \omega_j = \emptyset$, *for any* $j \neq i$,

then each $\omega_i, i = 1, 2, ..., s$, *is called a season in* Ω, *or a time interval or time zone.*

For example, when $\Omega = [1, 365]$ represents a year with February 1 as the starting point 1, then

$$\omega_1 = [1, 89], \omega_2 = [90, 181],$$

$$\omega_3 = [182, 273], \omega_4 = [274, 365]$$

would represent the spring, summer, autumn, and winter seasons of a year. If the starting point 1 is January 1, then

$$\omega_1 = [1, 31], \omega_2 = [32, 59], \omega_3 = [60, 90],$$

$$\omega_4 = [91, 120], \omega_5 = [121, 151], \omega_6 = [152, 181],$$

$$\omega_7 = [182, 212], \omega_8 = [213, 243], \omega_9 = [244, 273],$$

$$\omega_{10} = [274, 304], \omega_{11} = [305, 334], \omega_{12} = [335, 365]$$

will be the 12 months in a year.

Definition 9.5.2. *Assume that* $\omega_i \subset \Omega$ *is a season, and*

$$X = (x(1), x(2), ..., x(n)) \subset \omega_i$$

a sequence of raw data. For a fixed abnormal value ξ, *the corresponding catastrophe sequence*

$$X_\xi = (x[q(1)], x[q(2)], ..., x[q(m)])$$

is called a seasonal catastrophe sequence. Accordingly,

$$Q^{(0)} = (q(1), q(2), ..., q(m))$$

is called a seasonal catastrophe date sequence.

Proposition 9.5.1. *Assume that* $\omega_i = [a_i, b_i] \subset \Omega$, $a_i > 0$, *and that*

$$X = (x(1), x(2), ..., x(n)) \subset \omega_i = [a_i, b_i]$$

is a positive sequence of raw data. Let

$$y(k) = x(k) - a_i,$$

$k = 1, 2, ..., n$. *The distinguishing rate of the data in the sequence*

$$Y = (y(1), y(2), ..., y(n))$$

is greater than that of the original sequence X.

Proof: Suppose that

$$x(k), x(k+1)$$

are a pair of consecutive neighbors in X and

$$y(k), y(k+1)$$

the corresponding entries in Y. Then

$$\left|\frac{y(k+1) - y(k)}{y(k)}\right| = \left|\frac{x(k+1) - a_i - x(k) + a_i}{x(k) - a_i}\right| = \left|\frac{x(k+1) - x(k)}{x(k) - a_i}\right|.$$

From $a_i > 0$, it follows that $x(k) - a_i < x(k)$. Therefore,

$$\left|\frac{y(k+1) - y(k)}{y(k)}\right| = \left|\frac{x(k+1) - x(k)}{x(k) - a_i}\right| > \left|\frac{x(k+1) - x(k)}{x(k)}\right|.$$

That is, the relative differences of entries in Y are greater than those of the entries in X. So, the distinguishing rate has been improved.

For example, we take $\omega_i = [500, 520]$, and a sequence of raw data as

$$X = (x(i))_{i=1}^{5} = (502, 506, 509, 514, 518),$$

and

$$Y = (y(i))_{i=1}^{5} = (2, 6, 9, 14, 18).$$

Then

$$\left|\frac{x(3) - x(2)}{x(2)}\right| = \left|\frac{509 - 506}{506}\right| = 0.0059,$$

and

$$\left|\frac{y(3) - y(2)}{y(2)}\right| = \left|\frac{9 - 6}{6}\right| = 0.5.$$

Similarly, the differences between other pairs of entries can be discussed.

Seasonal disaster prediction is the study of occurrences of abnormal values appearing in a specified time interval. Based on Proposition 9.5.1, the starting point of the season can be simplified as to zero so that one can concentrate on the study of the sequence Y with an improved distinguishing rate between data values.

A seasonal disaster prediction can be conducted according to the following steps.

Step 1: Collect the sequence of raw data

$$X = (x(1), x(2), ..., x(n)).$$

Step 2: Study the range of change of the sequence of raw data, and determine the season $\omega_i = [a_i, b_i]$ of interest.

Step 3: Let $y(k) = x(k) - a_i$, and transform the original sequence into

$$Y = (y(1), y(2), ..., y(n))$$

in order to improve the distinguishing rate between the data values.

Step 4: Choose an abnormal value ξ and find the seasonal catastrophe sequence

$$Y_\xi = (y[q(1)], y[q(2)], ..., y[q(m)])$$

and the seasonal catastrophe date sequence

$$Q^{(0)} = (q(1), q(2), ..., q(m)).$$

Step 5: Establish the catastrophe GM(1, 1) model:

$$q^{(0)}(k) + az^{(1)}(k) = b.$$

Step 6: Test the simulation accuracy and make predictions.

Seasonal disaster prediction can be applied to studying occurrences of disasters during a certain season of a year. For example, the spring rain in Yunnan Province, the People's Republic of China, appears in the spring. The early frost in Shanxi Province and the Chilly-Dew Wind in the north area of Guangxi Province appear at the junction of the autumn and the winter seasons. The Wheat-Dry-Hot Wind of Henan Province comes at the end of May and early June. The bollworms of Henan Province generally appear in June. The flood season in southern Henan Province is often in the summer. What is worth noticing here is that the sequences of raw data in many disaster predictions are the sequences of dates in the ordinary sense, and the corresponding catastrophe date sequences are the times when the abnormalities occur.

Example 9.5.1. Let us look at how we predicted the Wheat-Dry-Hot Wind in the eastern area of Pingyu County of Henan Province.

Each Dry-Hot Wind is a serious natural disaster that appears near the time when wheat is almost mature for harvest. This wind can cause the wheat in the field to dry out too early causing reduced wheat yield by 20% to 40%. If it is possible to predict the occurrence of a Dry-Hot Wind in advance, some appropriate remedial measures can be applied, such as planting early-maturing varieties, or spraying in advance anti-wind and ripening fertilizers, etc., in order to decrease the loss.

In the eastern area of Pingyu County, there are four towns: Yangbu, Shuangmiao, Gaoyangdian, and Hedian. The dates Dry-Hot Wind occurred during the years from 1975 to 1988 are given in the following Table 9.4.

Table 9.4. Dates when Dry-Hot Winds occurred

Date	Month	Year	Date	Month	Year
3	6	1975	27	5	1982
25	5	1976	4	6	1983
7	6	1977	24	5	1984
1	6	1978	31	5	1985
29	5	1979	28	5	1986
26	5	1980	25	5	1987
5	6	1981	25	5	1988

If an occurrence of a Dry-Hot Wind before May 30 is considered a disaster, make a prediction based on Table 9.4.

Solution: Step 1: Starting on January 1, we can obtain a sequence of raw data from Table 9.4 as follows,

$$\begin{aligned} X &= (x(i))_{i=1}^{14} \\ &= (185, 176, 189, 183, 180, 177, 187, \\ &\quad 178, 186, 175, 182, 179, 176, 176). \end{aligned}$$

Step 2: Take $\omega = $ [May 20, June 10] $= [171, 192]$. Then $X \subset \omega$.

Step 3: Let $y(k) = x(k) - 171, k = 1, 2, ..., 14$. Then X can be changed to

$$\begin{aligned} Y &= (y(i))_{i=1}^{14} \\ &= (14, 5, 18, 12, 9, 6, 16, 7, 15, 4, 11, 8, 5, 5). \end{aligned}$$

Step 4: Let $\xi = $ May 29 $-$ May 20 $= 180 - 171 = 9$ be a lower abnormal value. So, the seasonal catastrophe sequence is

$$\begin{aligned} Y_\xi &= \{y[q(k)] | y[q(k)] \leq 9\} \\ &= (y[q(1)], y[q(2)], \cdots, y[q(8)]) \\ &= (y(2), y(5), y(6), y(8), y(10), y(12), y(13), y(14)) \,. \end{aligned}$$

Therefore,

$$Q^{(0)} = (q(i))_{i=1}^{8} = (2, 5, 6, 8, 10, 12, 13, 14) \,.$$

Step 5: Assume that

$$q^{(0)}(k) + az^{(1)}(k) = b.$$

Then

$$\widehat{a} = \left[\begin{array}{cc} a & b \end{array}\right]^T = \left[B^T B\right]^{-1} B^T Y = \left[\begin{array}{cc} -0.1588 & 5.017 \end{array}\right]^T.$$

Therefore, the catastrophe GM(1, 1) ordinality response sequence is

$$\begin{cases} \widehat{q}^{(1)}(k+1) & = 33.59 \cdot e^{0.1588k} - 31.59 \\ \widehat{q}(k+1) & = \widehat{q}^{(1)}(k+1) - \widehat{q}^{(1)}(k). \end{cases}$$

That is,

$$\widehat{q}(k+1) \quad = 33.59 \cdot (1 - e^{-0.1588}) \cdot e^{0.1588k} = 4.93 \cdot e^{0.1588k}.$$

Step 6: From

$$\widehat{q}(k+1) = 4.93 \cdot e^{0.1588k},$$

it follows that the simulated sequence of $Q^{(0)}$ is

$$\begin{aligned} \widehat{Q}^{(0)} &= (\widehat{q}(1), \widehat{q}(2), \widehat{q}(3), \widehat{q}(4), \widehat{q}(5), \widehat{q}(6), \widehat{q}(7), \widehat{q}(8)) \\ &= (4.93, 5.77, 6.77, 7.93, 9.30, 10.91, 12.78, 14.98)\,. \end{aligned}$$

From

$$\varepsilon(k) = q(k) - \widehat{q}(k),$$

$k = 1, 2, ..., 8$, it follows that

$$\begin{aligned} \varepsilon^{(0)} &= (\varepsilon(1), \varepsilon(2), \varepsilon(3), \varepsilon(4), \varepsilon(5), \varepsilon(6), \varepsilon(7), \varepsilon(8)) \\ &= (-2.93, -0.77, -0.77, 0.07, 0.70, 1.09, 0.22, -0.98)\,. \end{aligned}$$

Again from

$$\Delta_k = \left| \frac{\varepsilon(k)}{q(k)} \right|,$$

$k = 2, 3, ..., 8$, it follows that

$$\Delta \quad = (\Delta_i)_{i=2}^{8} = (0.154, 0.128, 0.009, 0.07, 0.091, 0.017, 0.07)$$

with the average relative error

$$\overline{\Delta} = \frac{1}{7} \sum_{k=2}^{8} \Delta_k = 0.077,$$

the average relative accuracy

$$1 - \overline{\Delta} = 92.3\%,$$

and the filtering accuracy

$$1 - \Delta_8 = 93\%.$$

Here, our simulation accuracy is close to the level two. Hence, our model obtained above can be used to make predictions. By using our models, we have

$$\widehat{q}(9) \approx 17, \widehat{q}(10) \approx 20,$$

$$\widehat{q}(9) - q(8) \approx 17 - 14 = 3,$$

$$\widehat{q}(10) - \widehat{q}(9) \approx 20 - 17 = 3,$$

$$1988 + 3 = 1991, 1991 + 3 = 1994.$$

That is, in 1991 and 1994, Dry-Hot Wind could become a disaster, (these prediction results were done before October 1988), and in 1989, there would not be a Dry-Hot Wind before May 29. The actual record showed that a Dry-Hot Wind occurred on June 2, 1989. And, so, no disaster was reported.

If combined with methods in meteorology and climatology, the output and reliability of the method of disaster prediction in grey systems theory can be further improved.

9.6 Stock-Market-Like Predictions

When the sequence of raw data vibrates widely with a relatively large amplitude, it is often difficult to find an appropriate simulation model. In this case, if the prediction on the ranges of change, as described in Section 9.3, are not satisfactory, we can make predictions on the wavy curve of the future development of the data based on the wavy curve of the known data sequence. This kind of prediction is called a stock-market-like prediction.

Definition 9.6.1. *Assume that*

$$X = (x(1), x(2), ..., x(n))$$

is a sequence of raw data. Then

$$x_k = x(k) + (t - k)[x(k+1) - x(k)]$$

is called a k-zigzagged line of the sequence X, and

$$\{x_k = x(k) + (t - k)[x(k+1) - x(k)] | k = 1, 2, ..., n-1\}$$

the zigzagged line of the sequence X, still denoted X. That is,

$$X = \{x_k = x(k) + (t - k)[x(k+1) - x(k)] | k = 1, 2, ..., n-1\}.$$

Definition 9.6.2. *Assume that*

$$\sigma_M = \max_{1\le k\le n}\{x(k)\}, \sigma_m = \min_{1\le k\le n}\{x(k)\}.$$

1. For any $\xi \in [\sigma_m, \sigma_M], X = \xi$ *is called* ξ*-contour (line);*

2. The solution $(t_i, x(t_i))$, ($i = 1, 2, ...$), *of the system of equations*

$$\begin{cases} X = \{x_k = x(k) + (t-k)[x(k+1) - x(k)] | k = 1, 2, ..., n-1\} \\ X = \xi \end{cases}$$

is called a ξ*-contour point.*

ξ-contour points are the intersection points of the zigzagged line of X and the ξ-*contour line*.

Proposition 9.6.1. *If there exists an* ξ*-contour point on the* i*-zigzagged line of* X*, then its coordinates of the point are*

$$\left(i + \frac{\xi - x(i)}{x(i+1) - x(i)}, \xi\right).$$

Proof: The equation of the i-zigzagged line is

$$X = x(i) + (t_i - i)[x(i+1) - x(i)].$$

Solving

$$\begin{cases} X = x(i) + (t_i - i)[x(i+1) - x(i)] \\ X = \xi \end{cases}$$

gives

$$t_i = i + \frac{\xi - x(i)}{x(i+1) - x(i)}. \square$$

Definition 9.6.3. *Assume that*

$$X_\xi = (P_1, P_2, ..., P_m)$$

is a sequence of ξ*-contour points, where* P_i *is located on the* t_i*th zigzagged line segment with coordinates*

$$\left(t_i + \frac{\xi - x(t_i)}{x(t_i+1) - x(t_i)}, \xi\right).$$

Let

$$q(i) = t_i + \frac{\xi - x(t_i)}{x(t_i+1) - x(t_i)},$$

$i = 1, 2, ..., m$. *Then*

$$Q^{(0)} = (q(1), q(2), ..., q(m))$$

is called a sequence of ξ-contour moments.

Establishing a GM(1, 1) for the sequence of ξ-contour moments can produce predicted values for future ξ-contour moments:

$$\widehat{q}(m+1), \widehat{q}(m+2), ..., \widehat{q}(m+k).$$

Definition 9.6.4. *Assume that*

$$\xi_0 = \sigma_m,$$

$$\xi_i = \sigma_m + \frac{i}{s}(\sigma_M - \sigma_m), \; i = 1, 2, \cdots, s-1,$$

$$\xi_s = \sigma_M.$$

Then $X = \xi_i, i = 0, 1, 2, ..., s$, are called contour lines with equal interval, and otherwise, called contour lines of non-equal interval.

When taking contour lines, more attention needs to be given to the assurance that the sequence of the corresponding contour points satisfies the requirements for building GM(1, 1) models. In general, we can take either contour line of the equal interval or contour lines of non-equal intervals, depending on the situation involved.

Definition 9.6.5. *Assume that $X = \xi_i, i = 1, 2, ..., s$, are s different contour lines,*

$$Q_i^{(0)} = (q_i(1), q_i(2), ..., q_i(m_i))$$

$i = 1, 2, ..., s$, sequences of contour moments, and

$$\widehat{q}_i(m_i+1), \widehat{q}_i(m_i+2), ..., \widehat{q}_i(m_i+k_i)$$

$i = 1, 2, ..., s$, GM(1, 1) predicted values for ξ_i-contour moments. If there exist $i \neq j$ such that

$$\widehat{q}_i(m_i+\ell_i) = \widehat{q}_j(m_j+\ell_j),$$

then $\widehat{q}_i(m_i+\ell_i)$ and $\widehat{q}_j(m_j+\ell_j)$ are called a pair of useless predicted moments.

Proposition 9.6.2. *Assume that*

$$\widehat{q}_i(m_i+1), \widehat{q}_i(m_i+2), ..., \widehat{q}_i(m_i+k_i),$$

$i = 1, 2, ..., s$, are the GM(1, 1) predicted values for ξ_i-contour moments. Delete all useless moments in

$$\widehat{q}_1(m_1+1), \widehat{q}_1(m_1+2), ..., \widehat{q}_1(m_1+k_1);$$

$$\widehat{q}_2(m_2+1), \widehat{q}_2(m_2+2), ..., \widehat{q}_2(m_2+k_2);$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\widehat{q}_i(m_i+1), \widehat{q}_i(m_i+2), ..., \widehat{q}_i(m_i+k_i);$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\widehat{q}_s(m_s+1), \widehat{q}_s(m_s+2), ..., \widehat{q}_s(m_s+k_s)$$

and rank the remaining moments from the smallest to the greatest as follows,

$$\widehat{q}(1) < \widehat{q}(2) < \cdots < \widehat{q}(n_s),$$

where $n_s \leq k_1+k_2+\cdots+k_s$. If $X = \xi_{\widehat{q}(k)}$ is the contour line corresponding to $\widehat{q}(k)$, then the predicted wavy curve of $X^{(0)}$ is

$$\begin{aligned} X &= \widehat{X}^{(0)} \\ &= \left\{ \xi_{\widehat{q}(k)} + [t - \widehat{q}(k)] \cdot (\xi_{\widehat{q}(k+1)} - \xi_{\widehat{q}(k)}) | k = 1, 2, ..., n \right\}. \end{aligned}$$

Example 9.6.1. Let us look at our stock-market-like prediction done for the annual runoff amount of the upper reaches of Fen River Reservoir in Shanxi Province, the People's Republic of China.

The total volume of water that can be kept in the Fen River Reservoir, which was built in 1958, is 0.72 billion m^3. The upper reaches of the reservoir include Ningwu County, Jingle County, Lan County, and Loufan County with a total drainage area of 52680 km^2. The curve in Figure 9.4 gives the annual runoff amounts from the year 1951 to the year 1980.

Let us take

$$\xi_1 = 2, \xi_2 = 2.5, \xi_3 = 3, \xi_4 = 4, \xi_5 = 5,$$

$$\xi_6 = 6, \xi_7 = 7, \xi_8 = 8, \xi_9 = 8.5,$$

with 0.1 billion m^3 as the dimension.

Then, the sequences of ξ_i-contour moments are given respectively as follows. For $\xi_1 = 2$,

$$Q_1^{(0)} = \{q_1(k)\}_1^5 = (15, 21.5, 22.1, 24.4, 25.2),$$

FIGURE 9.4. Annual runoff amounts in upper reaches of Fen River Reservoir

for $\xi_2 = 2.5$,

$$Q_2^{(0)} = \{q_2(k)\}_1^8 = (13, 14.8, 15.2, 22.1, 22.2, 23.9, 25.6, 29.9)\,,$$

for $\xi_3 = 3$,

$$\begin{aligned} Q_3^{(0)} &= \{q_3(k)\}_1^{13} \\ &= (1.7, 2.2, 5, 7.1, 12.4, 13.1, 14.7, \\ &\quad 15.5, 20.7, 22.3, 23.7, 26, 29.6), \end{aligned}$$

for $\xi_4 = 4$,

$$\begin{aligned} Q_4^{(0)} &= \{q_4(k)\}_1^{19} \\ &= (1, 2.6, 4.8, 5.5, 6.4, 7.4, 9.8, 10.7, 11.4, 13.5, \\ &\quad 14.4, 16.1, 17.9, 18.1, 19.9, 22.6, 23.4, 26.4, 29.1), \end{aligned}$$

for $\xi_5 = 5$,

$$\begin{aligned} Q_5^{(0)} &= \{q_5(k)\}_1^{13} \\ &= (3, 4.6, 7.8, 9.6, 13.9, 14.1, 16.1, \\ &\quad 17.8, 19, 22.8, 23.1, 26.7, 28.4), \end{aligned}$$

for $\xi_6 = 6$,

$$Q_6^{(0)} = \{q_6(k)\}_1^6 = (3.3, 4.5, 8.1, 9.4, 16.2, 17.7)\,,$$

for $\xi_7 = 7$,

$$Q_7^{(0)} = \{q_7(k)\}_1^6 = (3.5, 4.3, 8.5, 9.3, 16.3, 17.6)\,,$$

for $\xi_8 = 8$,

$$Q_8^{(0)} = \{q_8(k)\}_1^6 = (3.7, 4.2, 8.8, 9, 16.5, 17.5)\,,$$

and for $\xi_9 = 8.5$,

$$Q_9^{(0)} = \{q_9(k)\}_1^5 = (3.8, 4.1, 9.1, 16.6, 17.4)\,.$$

Apply accumulating generation once on $Q_i^{(0)}, i = 1, 2, ..., 9$. Then, the GM(1, 1) response sequences of $Q_i^{(1)}$, i = 1, 2, ..., 9, are respectively given as follows.

$$\widehat{q}_1^{(1)}(k+1) = 359.86 \cdot e^{0.06k} - 344.86,$$

$$\widehat{q}_2^{(1)}(k+1) = 128 \cdot e^{0.11k} - 115.7,$$

$$\widehat{q}_3^{(1)}(k+1) = 45.65 \cdot e^{0.13k} - 43.95,$$

$$\widehat{q}_4^{(1)}(k+1) = 54.15 \cdot e^{0.1k} - 53.15,$$

$$\widehat{q}_5^{(1)}(k+1) = 68.68 \cdot e^{0.12k} - 65.98,$$

$$\widehat{q}_6^{(1)}(k+1) = 16.1 \cdot e^{0.3k} - 12.8,$$

$$\widehat{q}_7^{(1)}(k+1) = 15.52 \cdot e^{0.3k} - 13.02,$$

$$\widehat{q}_8^{(1)}(k+1) = 16.7 \cdot e^{0.3k} - 13,$$

$$\widehat{q}_9^{(1)}(k+1) = 14.59 \cdot e^{0.37k} - 10.79.$$

Let

$$\widehat{q}_i(k+1) = \widehat{q}_i^{(1)}(k+1) - \widehat{q}_i^{(1)}(k).$$

We then can obtain the predicted sequences of ξ_i-contour moments, for i = 1, 2, ..., 9, as follows.

$$\widehat{Q}_1^{(0)} = (\widehat{q}_1(i))_{i=6}^{9} = (33.79, 35.79, 37.92, 40.16)\,,$$

$$\widehat{Q}_2^{(0)} = (\widehat{q}_2(i))_{i=8}^{11} = (33.26, 37.2, 41.6, 46.52)\,,$$

$$\widehat{Q}_3^{(0)} = (\widehat{q}_3(i))_{i=14}^{16} = (36.91, 42.17, 48.19)\,,$$

$$\widehat{Q}_4^{(0)} = (\widehat{q}_4(i))_{i=20}^{23} = (34.02, 37.59, 41.52, 45.87)\,,$$

$$\widehat{Q}_5^{(0)} = (\widehat{q}_5(i))_{i=14}^{17} = (34.17, 38.38, 43.11, 48.42)\,,$$

$$\widehat{Q}_6^{(0)} = (\widehat{q}_6(7), \widehat{q}_6(8)) = (34.39, 46.47)\,,$$

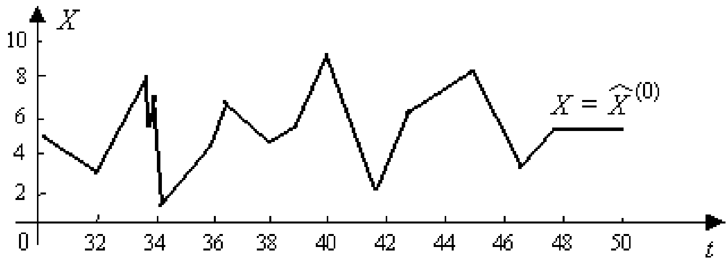

FIGURE 9.5. Predicted curve for the annual runoff amounts at upper reaches of Fen River Reservoir

$$\widehat{Q}_7^{(0)} = (\widehat{q}_7(7), \widehat{q}_7(8)) = (34.13, 45.96)\,,$$

$$\widehat{Q}_8^{(0)} = (\widehat{q}_8(7), \widehat{q}_8(8)) = (33.97, 45.68)\,,$$

and

$$\widehat{Q}_9^{(0)} = (\widehat{q}_9(6)) = (40.17)\,.$$

Based on these predicted values, we can draw the predicted curve for the annual runoff amounts at the upper reaches of Fen River Reservoir (see Figure 9.5 for more details).

9.7 Systems Predictions

For a system with several mutually related factors and many behavioral variables, no single model can truly reflect the development pattern of the system. So, we must consider establishing a system of models in order to make effective predictions for the system of interest.

Definition 9.7.1. *Assume that*

$$X_1^{(0)}, X_2^{(0)}, ..., X_m^{(0)}$$

are sequences of raw data for the state variables of a system, and

$$U_1^{(0)}, U_2^{(0)}, ..., U_s^{(0)}$$

sequences of data for the control variables of the system. Then,

$$\left\{\begin{array}{rcl}
\dfrac{dx_1^{(1)}}{dt} & = & a_{11}x_1^{(1)} + a_{12}x_2^{(1)} + \cdots + a_{1m}x_m^{(1)} + \\
 & & +b_{11}u_1^{(1)} + b_{12}u_2^{(1)} + \cdots + b_{1s}u_s^{(1)} \\
\dfrac{dx_2^{(1)}}{dt} & = & a_{21}x_1^{(1)} + a_{22}x_2^{(1)} + \cdots + a_{2m}x_m^{(1)} + \\
 & & +b_{21}u_1^{(1)} + b_{22}u_2^{(1)} + \cdots + b_{2s}u_s^{(1)} \\
\cdots\cdots & & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\dfrac{dx_m^{(1)}}{dt} & = & a_{m1}x_1^{(1)} + a_{m2}x_2^{(1)} + \cdots + a_{mm}x_m^{(1)} + \\
 & & +b_{m1}u_1^{(1)} + b_{m2}u_2^{(1)} + \cdots + b_{ms}u_s^{(1)} \\
\dfrac{du_1^{(1)}}{dt} & = & c_1u_1^{(1)} + d_1 \\
\dfrac{du_2^{(1)}}{dt} & = & c_2u_2^{(1)} + d_2 \\
\cdots\cdots & & \cdots\cdots \\
\dfrac{du_s^{(1)}}{dt} & = & c_su_s^{(1)} + d_s
\end{array}\right.$$

is called a system of prediction models.

In fact, a system of prediction models consists of m GM$(1, m+s)$ and s GM$(1, 1)$ differential equations.

Definition 9.7.2. *The matrix form of a system of prediction models is given as*

$$\left\{\begin{array}{l}
\dot{X} = AX + BU \\
\dot{U} = CU + D,
\end{array}\right.$$

where

$$X = [x_1, x_2, ..., x_m]^T,$$

$$U = [u_1, u_2, ..., u_s]^T,$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{ms} \end{bmatrix},$$

$$C = \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_s \end{bmatrix}, \; D = \begin{bmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_s \end{bmatrix}.$$

with X being the state vector, U the control vector, A the state matrix, B the control matrix, C the development matrix, and D the grey action vector.

Proposition 9.7.1. *The time response sequences of the system of prediction models as given in Definition 9.7.2 are*

$$\begin{aligned}
\widehat{x}_1^{(1)}(k+1) \;\; &= \;\; \left\{x_1^{(1)}(0) + \frac{1}{a_{11}}\left[\sum_{j=2}^{m} a_{1j}x_j^{(1)}(k+1)\right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{s} b_{1i}u_i^{(1)}(k+1)\right]\right\} \cdot e^{a_{11}k} \\
&\quad - \frac{1}{a_{11}}\left[\sum_{j=2}^{m} a_{1j}x_j^{(1)}(k+1) + \sum_{i=1}^{s} b_{1i}u_i^{(1)}(k+1)\right], \\
\widehat{x}_2^{(1)}(k+1) \;\; &= \left\{x_2^{(1)}(0) + \frac{1}{a_{22}}\left[\sum_{j\neq 2} a_{2j}x_j^{(1)}(k+1)\right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{s} b_{2i}u_i^{(1)}(k+1)\right]\right\} \cdot e^{a_{22}k} \\
&\quad - \frac{1}{a_{22}}\left[\sum_{j\neq 2} a_{2j}x_j^{(1)}(k+1) + \sum_{i=1}^{s} b_{2i}u_i^{(1)}(k+1)\right], \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\widehat{x}_m^{(1)}(k+1) \;\; &= \;\; \left\{x_m^{(1)}(0) + \frac{1}{a_{mm}}\left[\sum_{j\neq m} a_{mj}x_j^{(1)}(k+1)\right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{s} b_{mi}u_i^{(1)}(k+1)\right]\right\} \cdot e^{a_{mm}k} \\
&\quad - \frac{1}{a_{mm}}\left[\sum_{j\neq m} a_{mj}x_j^{(1)}(k+1) + \sum_{i=1}^{s} b_{mi}u_i^{(1)}(k+1)\right],
\end{aligned}$$

$$\widehat{u}_1^{(1)}(k+1) = \left[u_1^{(1)}(0) + \frac{d_1}{c_1}\right] \cdot e^{c_1 k} - \frac{d_1}{c_1},$$

$$\widehat{u}_2^{(1)}(k+1) = \left[u_2^{(1)}(0) + \frac{d_2}{c_2}\right] \cdot e^{c_2 k} - \frac{d_2}{c_2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\widehat{u}_s^{(1)}(k+1) = \left[u_s^{(1)}(0) + \frac{d_s}{c_s}\right] \cdot e^{c_s k} - \frac{d_s}{c_s},$$

where the response sequences of the state variables are approximate.

Example 9.7.1. Let us look at our prediction model for the grain production system of Shanxi Province, the People's Republic of China.

We have the system state variables,

1. x_1 : total grain yield of Shanxi Province (in 0.1 billion kg), and
2. x_2 : average yield per *mu* (0.15 *mu* = 1 acre) of Shanxi Province (in kg),

and system control variables

1. u_1 : total sown area in Shanxi Province (in 10,000 *mu*),
2. u_2 : average amount of fertilizers applied in each *mu* of land in Shanxi Province (in kg); and
3. u_3 : the ratio of irrigated area and the total sown area (%).

According to an analysis done previously, the total grain production has something to do with the total sown area and average yield per *mu* of land. And, the average yield per *mu* is closely related to the amount of fertilizers applied and the ratio of irrigated area and the total sown area. So, we obtain our system of prediction models as follows.

$$\begin{cases} \dfrac{dx_1^{(1)}}{dt} = a_{11}x_1^{(1)} + a_{12}x_2^{(1)} + b_{11}u_1^{(1)} \\ \dfrac{dx_2^{(1)}}{dt} = a_{22}x_2^{(1)} + b_{22}u_2^{(1)} + b_{23}u_3^{(1)} \\ \dfrac{du_1^{(1)}}{dt} = c_1u_1^{(1)} + d_1 \\ \dfrac{du_2^{(1)}}{dt} = c_2u_2^{(1)} + d_2 \\ \dfrac{du_3^{(1)}}{dt} = c_3u_3^{(1)} + d_3. \end{cases}$$

From the historical data of x_1, x_2, and u_1, u_2, u_3, we can obtain the least squares estimates for the system's state matrix A, control matrix B, development matrix C, and grey action vector D as follows.

$$\widehat{A} = \begin{bmatrix} -1.97209 & 0.8118 \\ 0 & -1.52385 \end{bmatrix}, \widehat{B} = \begin{bmatrix} 0.0033 & 0 & 0 \\ 0 & 2.09693 & 8.0016 \end{bmatrix},$$

$$\widehat{C} = \begin{bmatrix} -0.0058 & 0 & 0 \\ 0 & 0.0819 & 0 \\ 0 & 0 & 0.0014 \end{bmatrix}, \widehat{D} = \begin{bmatrix} 5000.7174 \\ 10.1785 \\ 28.2795 \end{bmatrix}.$$

So, the system of prediction models is

$$\begin{cases} \dfrac{dx_1^{(1)}}{dt} = -1.97209x_1^{(1)} + 0.8118x_2^{(1)} + 0.0033u_1^{(1)} \\ \dfrac{dx_2^{(1)}}{dt} = -1.52385x_2^{(1)} + 2.09693u_2^{(1)} + 8.0016u_3^{(1)} \\ \dfrac{du_1^{(1)}}{dt} = -0.0058u_1^{(1)} + 5000.7174 \\ \dfrac{du_2^{(1)}}{dt} = 0.0819u_2^{(1)} + 10.1785 \\ \dfrac{du_3^{(1)}}{dt} = 0.0014u_3^{(1)} + 28.2795 \end{cases}$$

with the response sequences

$$\begin{aligned} \widehat{x}_1^{(1)}(k+1) &= \{x_1^{(1)}(0) - \frac{1}{1.97209}[0.8118x_2^{(1)}(k+1) \\ &\quad + 0.0033u_1^{(1)}(k+1)]\} \cdot e^{-1.97209k} \\ &\quad + \frac{1}{1.97209}[0.8118x_2^{(1)}(k+1) + 0.0033u_1^{(1)}(k+1)], \end{aligned}$$

$$\begin{aligned}\widehat{x}_2^{(1)}(k+1) &= \{x_2^{(1)}(0) - \frac{1}{1.52385}[2.09693u_2^{(1)}(k+1) \\ &\quad + 8.0016u_3^{(1)}(k+1)]\} \cdot e^{-1.52385k} \\ &\quad + \frac{1}{1.52385}[2.09693u_2^{(1)}(k+1) + 8.0016u_3^{(1)}(k+1)],\end{aligned}$$

$$\widehat{u}_1^{(1)}(k+1) = \left[u_1^{(1)}(0) - \frac{5000.7174}{0.0058}\right] \cdot e^{-0.0058k} + \frac{7000.7174}{0.0058},$$

$$\widehat{u}_2^{(1)}(k+1) = \left[u_2^{(1)}(0) - \frac{10.1785}{0.0819}\right] \cdot e^{0..0819k} - \frac{10.1785}{0.0819},$$

$$\widehat{u}_3^{(1)}(k+1) = \left[u_3^{(1)}(0) - \frac{28.2795}{0.0014}\right] \cdot e^{0.0014k} - \frac{28.2795}{0.0014}.$$

The predicted values from our system of prediction models are given in the following Table 9.5.

Table 9.5. Predicted values

	1990	1995	2000
u_1	4402	4301.15	4178
u_2	55.62	83.76	126
u_3	29.1	29.3	29.5
x_1	98.96	113.72	135.5
x_2	225.36	263.19	319.5

When the control variables take different values, we can obtain different predicted values for the system. For example, when $u_1 = 4492$ in ten thousand *mu*, $u_2 = 32$ kg, $u_3 = 28.6\%$, or let u_1 be 4000, 4250, 4500 in ten thousand *mu*, u_2 be 50 kg, 75 kg, $u_3 = 29\%$, respectively; the corresponding predicted values for the grain productions are listed in Table 9.6.

Table 9.6. Control values and the relevant predictions

Control		Variable	State			Variable
u_1	u_2	u_3	x_1			x_2
4492	32	28.6	87.6	87.45	87.45	194.5
4000			97.77	97.7	97.7	
4250	50	29	83.18	98.12	98.12	221
4500			98.6	98.53	98.53	
4000			111.94	111.86	111.86	
4250	75	29	112.35	112.28	112.275	255
4500			112.77	112.70	112.70	

From Table 9.6, it can be seen that if the sown area u_1 is controlled between 4000 to 4500 in ten thousand *mu*, the amount u_2 of fertilizers applied in each *mu* is between 50 kg to 75 kg, and the ratio of the irrigated area u_3 is 29%, then the total grain production after 1990 would be stabilized between 9.77 to 11.2 billion kg.

9.8 Practical Applications

In this final section of this chapter, we use one real-life example to try out our theory developed in this chapter.

Example 9.8.1. Let us do a grey prediction for flood and drought disasters for Henan Province.

Henan is a province, which, historically, mainly lives on agriculture revenues. The entire province is located in a warm temperature zone and northern Asian tropical zone. Its annual average precipitation is about 600 to 1200 mm, which increases from southeast to northwest. It is mainly affected by monsoon climates with a combined characteristic of the south and the north. There exist great differences in weather conditions between various locations in the province. Its climatic conditions are good for the growth of a great many different varieties of agricultural plants. On the other hand, circulations of monsoons bring about a great frequency of disastrous weather conditions, such as drought, flood, Dry-Hot Wind, sandstorm, hail, frost, etc., which occur alternatively, causing serious consequences for grain production and human's living in the province. Especially, drought and flood bring more serious disasters to much greater areas. Here, what is worth mentioning is drought, which can cover the entire province and occurs almost every year. For example, during the time period of 600 plus years from 1300 to 1911, droughts of the province size occurred in 88 years. From 1949 to 1989, a period of 41 years, there appeared major droughts in 13 years, each of which covered an area of more than 10,000,000 *mu* (0.15 *mu* = 1 acre), where in four years, there appeared disastrous droughts affecting more than 25,000,000 *mu*. Even though the disastrous areas of floods have been smaller than those suffering from droughts, the actual consequences have not been any less than those of droughts. For example, during the 600 plus years from 1300 to 1911, Henan had suffered from major floods 69 times. During the time period of 41 years from 1949 to 1989, there appeared 13 major floods, covering an area greater than 10,000,000 *mu*, with disastrous areas greater than 25,000,000 *mu* in 6 years. Here, in this example, we do a prediction, using the available statistical records of Henan Province regarding floods, droughts, and related sizes of areas affected for the years from 1949 to 1989, about potential occurrences and disastrous levels of flood and drought for the last 10 years of the 20th century, based on grey systems methodology and principles of disaster predictions.

By using the sizes of areas affected by floods and droughts, we divide floods and droughts into four grey classes:

- Light disaster: disastrous area is less than 5,000,000 mu;
- Medium disaster: disastrous area is between 5,000,000 mu and 10,000, 000 mu;
- Serious disaster: disastrous area is between 10,000,000 mu and 25,000, 000 mu;
- Extraordinary disaster: disastrous area is greater than 25,000,000 mu.

Here, our task is to predict for medium, serious, and extraordinary disasters. In our catastrophe models, we denote medium, serious, and extraordinary disasters as the third class, the second class, and the first class disasters, respectively. And, $q_i^1(k)$, $i = 1$, 2, 3, is used to stand for the year when the ith class flood disaster occurred the kth time after 1949, and $q_i^2(k), i = 1$, 2, 3, for the year when the ith class drought disaster occurred the kth time after 1949. Based on the available statistical records, we have the following 6 catastrophe sequences.

$$
\begin{aligned}
q_1^1 &= \left(q_1^1(i)\right)_{i=1}^{6} = (1956, 1957, 1963, 1964, 1982, 1984)\,, \\
q_2^1 &= \left(q_2^1(i)\right)_{i=1}^{7} = (1953, 1954, 1965, 1975, 1976, 1979, 1985)\,, \\
q_3^1 &= \left(q_3^1(i)\right)_{i=1}^{9} \\
&= (1950, 1952, 1955, 1958, 1960, 1962, 1977, 1980, 1983)\,, \\
q_1^2 &= \left(q_1^2(i)\right)_{i=1}^{4} = (1961, 1978, 1986, 1988)\,, \\
q_2^2 &= \left(q_2^2(i)\right)_{i=1}^{9} \\
&= (1959, 1960, 1962, 1965, 1966, 1981, 1982, 1985, 1987)\,, \\
q_3^2 &= \left(q_3^2(i)\right)_{i=1}^{6} = (1968, 1972, 1976, 1977, 1979, 1980)\,.
\end{aligned}
$$

In order to improve the distinguishing rates between data values in the sequences without loss of convenience to compare the actual year, let

$$d_i^j(k) = q_i^j(k) - 1900,$$

$j = 1, 2$; $i = 1, 2, 3$. So, we obtain the following transformed catastrophe sequences.

$$
\begin{aligned}
d_1^1 &= \left(d_1^1(i)\right)_{i=1}^{6} = (56, 57, 63, 64, 82, 84)\,, \\
d_2^1 &= \left(d_2^1(i)\right)_{i=1}^{7} = (53, 54, 65, 75, 76, 79, 85)\,, \\
d_3^1 &= \left(d_3^1(i)\right)_{i=1}^{9} = (50, 52, 55, 58, 60, 62, 77, 80, 83)\,,
\end{aligned}
$$

$$d_1^2 = \left(d_1^2(i)\right)_{i=1}^4 = (61, 78, 86, 88)\,,$$

$$d_2^2 = \left(d_2^2(i)\right)_{i=1}^9 = (59, 60, 62, 65, 66, 81, 82, 85, 87)\,,$$

$$d_3^2 = \left(d_3^2(i)\right)_{i=1}^6 = (68, 72, 76, 77, 79, 80)\,.$$

By establishing a GM(1, 1) model for each d_i^j, $j = 1, 2$; $i = 1, 2, 3$, we obtain the following restored sequences of the GM(1, 1) ordinality response sequences for various disasters.

$$\widehat{d}_1^1(k+1) = 50.21 \cdot e^{0.105k},$$

$$\widehat{d}_2^1(k+1) = 55.01 \cdot e^{0.07516k},$$

$$\widehat{d}_3^1(k+1) = 46.55 \cdot e^{0.07384k},$$

$$\widehat{d}_1^2(k+1) = 74.57 \cdot e^{0.0588k},$$

$$\widehat{d}_2^2(k+1) = 55.61 \cdot e^{0.0596k},$$

$$\widehat{d}_3^2(k+1) = 71.21 \cdot e^{0.02457k}.$$

After conducting model validity tests, we find that the simulated relative errors, degrees of incidences, the ratios of mean square deviations, and small error probabilities all satisfy requirements. Hence, we can use the model as a prediction model.

From

$$\widehat{d}_1^1(7) = 94.27, \widehat{d}_1^1(8) = 104.71,$$

it follows that

$$[\widehat{d}_1^1(7)] = [94.27] = 94, [\widehat{d}_1^1(8)] = [104.71] = 105.$$

So,

$$[\widehat{q}_1^1(7)] = [\widehat{d}_1^1(7)] + 1900 = 1994$$

and

$$[\ \widehat{q}_1^1(8)] = [\ \widehat{d}_1^1(8)] + 1900 = 2005.$$

It follows that in the 20th century, there might be an extraordinary flood disaster in 1994, which would affect more than 25,000,00 *mu* of land in Henan Province. The next occurrence of extraordinary flood disaster would be in the year of 2005. Similarly, we can obtain predicted years for serious flood disasters with disastrous areas between 10,000,000 *mu* to 25,000,000 *mu* and medium-sized flood disasters with disastrous areas between 5,000,000 *mu* to 10,000,000 *mu* as follows.

$$[\,\widehat{q}_2^1(8)] = [\,\widehat{d}_2^1(8)] + 1900 = 1993,$$

$$[\,\widehat{q}_2^1(9)] = [\,\widehat{d}_2^1(9)] + 1900 = 2000,$$

and

$$[\,\widehat{q}_3^1(10)] = [\,\widehat{d}_3^1(10)] + 1900 = 1991,$$

$$[\,\widehat{q}_3^1(11)] = [\,\widehat{d}_3^1(11)] + 1900 = 1997.$$

The predicted years for various kinds of drought disasters are given as follows.

$$[\,\widehat{q}_1^2(5)] = 1994, [\,\widehat{q}_1^2(6)] = 2000;$$

$$[\,\widehat{q}_2^2(10)] = 1995, [\,\widehat{q}_2^2(11)] = 2001;$$

$$[\,\widehat{q}_3^2(7)] = 1990, [\,\widehat{q}_3^2(8)] = 1993,$$

$$[\,\widehat{q}_3^2(9)] = 1996, [\,\widehat{q}_3^2(10)] = 1998.$$

All predicted values of our model here are listed in Table 7.7.

Table 9.7. Predicted values of the chosen model

	Flood	Disaster	Drought		Disaster
Medium	1991	1997	1990	1993	1996
Serious	1993	2000	1995		
Extraordinary	1994			1994	2000

This model and related simulated values have been obtained in the year 1989. The predicted outcomes for the flood and drought disasters for the year 1994 and the drought disaster for the year 1995 have agreed well with the actual circumstances in these years. In 1994, the drought affected about 60,000,000 *mu* and the flood about 23,000,000 *mu*. And in 1995, the drought affected about 18,000,000 *mu*.

10
Grey Decisions

10.1 Introduction

Deciding what actions to take in order to achieve a scheduled target based on the actual circumstances in the environment is so-called decision making. The essential meaning of decision making is to “make a decision” or to “choose an appropriate reaction.” Decision making is not only an important part in various administrative activities at all levels, but also permeates everyone’s work, study, and life from the start to end. Understanding about decision making can be divided into two categories, one with a more general scope and the other a more narrow scope. Generally speaking, decision making means the entire process of raising questions, collecting supporting materials, determining goals, making plans, analyzing and evaluating situations, choosing plans, implementing, feedback and modifying the plans, etc. The more narrow category for decision making implies the step and the only step in the entire decision making process about choosing plans, which is often known as “having the final say.” There also exist people who understand decision making as choosing plans under the uncertainty of known information. In this case, the decision making depends heavily on the personal experience, attitude, and determination of the decision maker, who would have to be responsible for certain risks associated with the final decision. Grey decision making is one where the model in use contains grey elements, or where a general decision making model is combined with some grey models. The emphasis of grey decision making is on the study of the problem of choosing plans.

In the discussion below, we call the problem of concern, the phenomenon of focus, the current situation of a system's behavior, etc., an event, which will be the start of each of our decision making.

Definition 10.1.1. *Events, countermeasures, objectives, and effects are called the four key elements of decision making.*

Definition 10.1.2. *The totality of all events within a range of research is called the set of events of the research, denoted*

$$A = \{a_1, a_2, ..., a_n\},$$

where $a_i, i = 1, 2, ..., n$, *is the* ith *event. The corresponding totality of all possible countermeasures is called the countermeasure set, denoted*

$$B = \{b_1, b_2, ..., b_m\},$$

where $b_j, j = 1, 2, ..., m$, *is the* jth *countermeasure.*

Definition 10.1.3. *Assume that*

$$A = \{a_1, a_2, ..., a_n\}$$

is the set of events of a research and

$$B = \{b_1, b_2, ..., b_m\}$$

the countermeasure set. Then, the Cartesian product

$$A \times B = \{(a_i, b_j) | a_i \in A, b_j \in B\}$$

is called the situation set, denoted $S = A \times B$. *For any* $a_i \in A$ *and* $b_j \in B$, *the pair* (a_i, b_j) *is called a situation, denoted* $s_{ij} = (a_i, b_j)$.

For example, in the decision making on what to plant in agriculture, weather conditions can be used as the set of events with a normal year denoted as a_1, a drought year as a_2, and a flood year as a_3. Then, the set of events is

$$A = \{\text{normal year, drought year, flood year}\} = \{a_1, a_2, a_3\}.$$

And, different strains of crops can be seen as countermeasures with corn denoted as b_1, Chinese sorghum as b_2, soybeans as b_3, sesame b_4, potatoes and yams as b_5, $\cdots$; then the countermeasure set is given as

$$\begin{aligned} B &= \{\text{corn, Chinese sorghum, soybean, sesame, potatos and yams, ...}\} \\ &= \{b_1, b_2, b_3, b_4, b_5, ...\}. \end{aligned}$$

Therefore, the situation set is

$$\begin{aligned} S &= A \times B \\ &= \{s_{11}, s_{12}, s_{13}, s_{14}, s_{15}, ...; \\ &\quad s_{21}, s_{22}, s_{23}, s_{24}, s_{25}, ...; \\ &\quad s_{31}, s_{32}, s_{33}, s_{34}, s_{35}, ...\}, \end{aligned}$$

where $s_{ij} = (a_i, b_j)$, For example

$$\begin{aligned} s_{11} &= (a_1, b_1) = (\text{normal year, corn}) \, , \\ s_{21} &= (a_2, b_1) = (\text{drought year, corn}) \, , \\ s_{31} &= (a_3, b_1) = (\text{flood year, corn}) \, , \end{aligned}$$

and

$$s_{12} = (a_1, b_2) = (\text{normal year, Chinese sorghum}) \, .$$

Here, the events and countermeasures are simple. So, the situations constructed are relatively simple too. In a practical decision making, the events of consideration are often complicated, consisting of many kinds of simple events, and the countermeasures are complicated too. Hence, the resultant situations can be extremely complicated.

For now, let us continue to use the previous planting decision making as our example.

As a matter of fact, the set of events is the organic body consisting of weather, soil, irrigation, fertilizer, agricultural chemical, work force, technology, etc. And, the countermeasures are not simply the individual strains of crops, but various proportional combinations of many different strains of crops. Let us write

- The event: "normal year, loam, 50% effective irrigation area, fertilizer and agricultural chemical essentially meet the need, sufficient work force, medium level of technology, etc." as a_1,
- The event: "drought year, black earth, 50% effective irrigation area, sufficient fertilizer and work force, lack of agricultural chemical, medium level of technology, etc." as a_2;

. . . .

then, we have the set of events

$$A = \{a_1, a_2, \cdots\} \, .$$

Let us write

- The countermeasure: "30% corn + 10% Chinese sorghum + 20% soybeans + 15% sesame + 15% potatoes and yams + 10% others" as b_1, and
- The countermeasure: "10% corn + 20% Chinese sorghum + 30% soybeans + 30% sesame + 10% others" as b_2;

. . . .

then, we have the countermeasure set

$$B = \{b_1, b_2, ...\}.$$

Now, the situation $s_{11} = (a_1, b_1)$ is that, under the condition of a normal year, loam, 50% effective irrigation area, that fertilizer and agricultural chemical essentially meet the need, sufficient workforce, medium level of technology, etc., plant 30% corn, 10% Chinese sorghum + 20% soybeans + 15% sesame + 15% potatoes and yams + 10% others.

Let us look at the example of teaching scheduling. The collection of all course offerings of a fixed semester at a certain school can be seen as the set of events, all teaching faculty of this school, and various teaching methods, such as laboratory, interns, multimedia, etc., are seen as the set of countermeasures. Based on the circumstances, one teacher can teach several courses, or several teachers teach one course together. The work load could be 100% teaching, or 60% teaching, 20% laboratory, 10% interns, and 10% work with multimedia and others.

For a fixed situation $s_{ij} \in S$ under a set of prescheduled targets or objectives, one needs to evaluate the effects. According to the evaluation, choices will need to be made. This is the decision making. In the following sections, we address several different grey decision making methods.

10.2 Grey Target Decisions

In this section, let us look at such decision processes where the target of each decision making is grey.

Definition 10.2.1. *Assume that*

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}$$

is the situation set, $u_{ij}^{(k)}$ the effect value of situation s_{ij} with objective k, and R the set of all real numbers. Then,

$$u^{(k)} : \quad S \to R$$

$$s_{ij} \mapsto u_{ij}^{(k)}$$

is called the effect mapping of S with objective k.

Definition 10.2.2. *1. If $u_{ij}^{(k)} = u_{ih}^{(k)}$, then the countermeasures b_j and b_h are said to be equivalent with respect to event a_i with objective k, denoted $b_j \cong b_h$, and the set*

$$B_i^{(k)} = \{b_j | b_j \in B, b_j \cong b_h\}$$

is called the effect equivalence class of the countermeasure b_h with respect to event a_i with objective k.

2. Assume that k is such an objective satisfying that the greater the effect value the better. If $u_{ij}^{(k)} > u_{ih}^{(k)}$, then the countermeasure b_j is said to be superior to the countermeasure b_h with respect to event a_i with objective k, denoted $b_j \succ b_h$. The set

$$B_{ih}^{(k)} = \{b_j | b_j \in B, b_j \succ b_h\}$$

is called the superior class of the countermeasure b_h with respect to event a_i with objective k.

Similarly, we can define the concept of superior classes of countermeasures for the cases of (1) the closer to a fixed moderate value the effect value is the better, and (2) the smaller the effect value the better.

Definition 10.2.3. *1. If $u_{ih}^{(k)} = u_{gh}^{(k)}$, then the events a_i and a_g are said to be equivalent with respect to countermeasure b_h with objective k, denoted $a_i \cong a_g$. The set*

$$A_h^{(k)} = \{a_i | a_i \in A, a_i \cong a_g\}$$

is called the effect equivalence class of the event a_g with respect to countermeasure b_h with objective k.

2. Assume that k is an objective such that the greater the effect value the better. If $u_{ih}^{(k)} > u_{gh}^{(k)}$, then the event a_i is said to be superior to the event a_g with respect to countermeasure b_h with objective k, denoted $a_i \succ a_g$. The set

$$A_{gh}^{(k)} = \{a_i | a_i \in A, a_i \succ a_g\}$$

is called the superior class of the event a_g with respect to countermeasure b_h with objective k.

Similarly, we can define the concept of superior classes of events for the cases of (1) the closer to a fixed moderate value the effect value is the better, and (2) the smaller the effect value the better.

Definition 10.2.4. *1. If $u_{ij}^{(k)} = u_{gh}^{(k)}$, then the situation s_{ij} is said to be equivalent to the situation s_{gh} with objective k, denoted $s_{ij} \cong s_{gh}$. The set*

$$S^{(k)} = \{s_{ij} | s_{ij} \in S, s_{ij} \cong s_{gh}\}$$

is called the effect equivalence class of the situation s_{gh} *with objective* k.

2. Assume that k *is an objective satisfying that the greater the effect value the better. If* $u_{ij}^{(k)} > u_{gh}^{(k)}$, *then the situation* s_{ij} *is said to be superior to the situation* s_{gh} *with objective* k, *denoted* $s_{ij} \succ s_{gh}$. *The set*

$$S_{gh}^{(k)} = \{s_{ij} | s_{ij} \in S, s_{ij} \succ s_{gh}\}$$

is called the effect superior class of the situation s_{gh} *with objective* k.

Similarly, we can define the concept of superior classes of situations for the cases of (1) the closer to a fixed moderate value the effect value is the better, and (2) the smaller the effect value the better.

Proposition 10.2.1. *Assume that*

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\} \neq \emptyset$$

and

$$U^{(k)} = \left\{u_{ij}^{(k)} | s_{ij} \in S\right\}$$

the effect set with objective k, *and* $\left\{S^{(k)}\right\}$ *is the set of effect equivalence classes of situations with objective* k. *Then, the mapping*

$$u^{(k)} : \quad \left\{S^{(k)}\right\} \to U^{(k)}$$

$$S^{(k)} \mapsto u_{ij}(k)$$

is one-to-one and onto.

Definition 10.2.5. *Assume that* $d_1^{(k)}$ *and* $d_2^{(k)}$ *are the upper and the lower threshold values for situation effects with objective* k. *Then,*

$$S^1 = \left\{r | d_1^{(k)} \leq r \leq d_2^{(k)}\right\}$$

is called the grey target of one-dimensional decision making with objective k, *and* $u_{ij}^{(k)} \in [d_1^{(k)}, d_2^{(k)}]$ *a satisfactory effect with objective* k, *the corresponding* s_{ij} *the desirable situation with objective* k, *and* b_j *the desirable countermeasure with respect to event* a_i *with objective* k.

Proposition 10.2.2. *Assume that* $u_{ij}^{(k)}$ *is the effect value of the situation* s_{ij} *with objective* k. *If* $u_{ij}^{(k)} \in S^1$, *that is,* s_{ij} *is a desirable situation with objective* k, *then for any* $s \in S_{ij}^{(k)}$, s *is also desirable. That is, when* s_{ij} *is desirable, all situations in its effect superior class are all desirable.*

What is addressed above is about the case of a single objective. Similarly, we can discuss decision making grey targets of situations with multiple objectives.

Definition 10.2.6. *Assume that* $d_1^{(1)}$ *and* $d_2^{(1)}$ *are the threshold values of situation effects for objective* 1, *and* $d_1^{(2)}$ *and* $d_2^{(2)}$ *the threshold values of situation effects for objective 2. Then,*

$$S^2 = \left\{(r^{(1)}, r^{(2)})|d_1^{(1)} \leq r^{(1)} \leq d_2^{(1)}, d_1^{(2)} \leq r^{(2)} \leq d_2^{(2)}\right\}$$

is called a grey target of two-dimensional decision making. If the effect vector of the situation s_{ij} *satisfies* $u_{ij} = (u_{ij}^{(1)}, u_{ij}^{(2)}) \in S^2$, *then* s_{ij} *is said to be a desirable situation with objectives 1 and 2,* b_j *a desirable countermeasure of the event* a_i *with objectives 1 and 2.*

Definition 10.2.7. *Assume that* $d_1^{(1)}$ *and* $d_2^{(1)}, d_1^{(2)}$ *and* $d_2^{(2)}, ..., d_1^{(s)}$ *and* $d_2^{(s)}$ *are the threshold values of situation effects with objective 1, 2, ...,* s, *respectively. Then, the following region of the* s*-dimensional Euclidean space*

$$\begin{aligned} S^s = \quad & \{(r^{(1)}, r^{(2)}, ..., r^{(s)})|d_1^{(1)} \leq r^{(1)} \leq d_2^{(1)}, \\ & d_1^{(2)} \leq r^{(2)} \leq d_2^{(2)}, ..., d_1^{(s)} \leq r^{(s)} \leq d_2^{(s)}\} \end{aligned}$$

is called a grey target of s*-dimensional decision making. If the effect vector of the situation* s_{ij} *satisfies*

$$u_{ij} = \left(u_{ij}^{(1)}, u_{ij}^{(2)}, ..., u_{ij}^{(s)}\right) \in S^s,$$

where $u_{ij}^{(k)}, k = 1, 2, ..., s$, *is the effect value of the situation* s_{ij} *with objective* k, *then* s_{ij} *is said to be a desirable situation with objective 1, 2,* $\cdots, s$, b_j *a desirable countermeasure of the event* a_i *with objective 1, 2, ...,* s.

Grey targets of a decision making are essentially the region for the location of desirable effects in terms of relative optimization. In many cases, because achieving the absolute optimum is often impossible, we are happy with reaching a satisfactory result. Of course, according to the need, we can gradually shrink the grey target for our decision making need until it degenerates into a point, which is the optimum effect with the corresponding situation as the optimum situation, and the corresponding countermeasure as the optimum countermeasure. Therefore, in the following, our discussion is around solid spherical targets.

Definition 10.2.8. *The following*

$$\begin{aligned} R^s = \quad & \{(r^{(1)}, r^{(2)}, ..., r^{(s)})|(r^{(1)} - r_0^{(1)})^2 + (r^{(2)} - r_0^{(2)})^2 \\ & + \cdots + (r^{(s)} - r_0^{(s)})^2 \leq R^2\} \end{aligned}$$

is called the s*-dimensional spherical grey target with center*

$$r_0 = (r_0^{(1)}, r_0^{(2)}, ..., r_0^{(s)})$$

and radius R, and

$$r_0 = (r_0^{(1)}, r_0^{(2)}, ..., r_0^{(s)})$$

the optimum effect vector.

Definition 10.2.9. *Assume that*

$$r_1 = (r_1^{(1)}, r_1^{(2)}, ..., r_1^{(s)}) \in R^s.$$

Then

$$\begin{aligned} |r_1 - r_0| &= [(r_1^{(1)} - r_0^{(1)})^2 + (r_1^{(2)} - r_0^{(2)})^2 \\ &\quad + \cdots + (r^{(s)} - r_0^{(s)})^2]^{\frac{1}{2}} \end{aligned}$$

is called the bull's-eye-distance of the vector r_1.

The values of bull's-eye-distances reflect the superiority of effect vectors.

Definition 10.2.10. *Assume that s_{ij} and s_{gh} are two different situations,*

$$u_{ij} = \left(u_{ij}^{(1)}, u_{ij}^{(2)}, ..., u_{ij}^{(s)}\right)$$

and

$$u_{gh} = \left(u_{gh}^{(1)}, u_{gh}^{(2)}, ..., u_{gh}^{(s)}\right)$$

are *the effect vectors of the s_{ij} and s_{gh}, respectively. If*

$$|u_{ij} - r_0| \geq |u_{gh} - r_0|,$$

then the situation s_{gh} is said to be superior to the situation s_{ij}, denoted $s_{gh} \succ s_{ij}$. When an equal sign holds true here, the situations s_{gh} and s_{ij} are said to be equivalent, denoted $s_{gh} \cong s_{ij}$.

Definition 10.2.11. *If for any $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$, it is always true that $u_{ij} \neq r_0$, then we say that the optimum situation does not exist, or that the event does not have any optimum countermeasure.*

Definition 10.2.12. *If the optimum situation does not exist, but there exist g and h such that for any $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$, it is always true that*

$$|u_{ij} - r_0| \geq |u_{gh} - r_0|;$$

that is, for any $s_{ij} \in S$, $s_{gh} \succ s_{ij}$, then s_{gh} is called a quasi-optimum situation, a_g a quasi-optimum event, and b_h a quasi-optimum countermeasure.

For the sake of convenience for our discussion, we can take the origin as the center or the bull's-eye of the grey target, which can be done by simply applying an appropriate transformation to effect vectors. In this case, the bull's-eye-distance is transformed as the 2-norm of effect vectors of decision making.

Theorem 10.2.1. *Assume that*

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}$$

is the situation set, and

$$R^s = \{(r^{(1)}, r^{(2)}, ..., r^{(s)}) | (r^{(1)} - r_0^{(1)})^2 + (r^{(2)} - r_0^{(2)})^2 + \cdots + (r^{(s)} - r_0^{(s)})^2 \leq R^2\}$$

the s-dimensional spherical grey target. Then, S becomes an ordered set with "superiority" being the ordering relation.

Theorem 10.2.2. *In the situation set* $(S, \succ)$, *there must exist a quasi-optimum situation.*

Proof. This is a restatement of Zorn's Lemma in set theory. □

Example 10.2.1. Assume that "raining" is event a_1, and "using an umbrella," "wearing a raincoat," and "wearing a bamboo hat" are the countermeasures b_1, b_2, and b_3, respectively. Try to make a grey target decision with three objectives: effective, economical, and convenient.

Solution: We denote being effective as objective 1, being economical as objective 2, and being convenient as objective 3. So, the three situations

$$s_{11} = (a_1, b_1) = (\text{rain, umbrella}),$$

$$s_{12} = (a_1, b_2) = (\text{rain, raincoat}),$$

$$s_{13} = (a_1, b_3) = (\text{rain, bamboo hat})$$

will have different effects with different objectives.

1. For objective 1, it is to look at the situation from the angle of being dry, umbrella is the best, raincoat the second, and bamboo hat the worst. We now define the best effect as level one, denoted as $u_{11}^{(1)} = 1$. Similarly, we have $u_{12}^{(1)} = 2$ and $u_{13}^{(1)} = 3$.
2. For objective 2, using a bamboo hat is the most economical choice, umbrella the second, and raincoat the third. So, we have $u_{11}^{(2)} = 2$, $u_{12}^{(2)} = 3$, and $u_{13}^{(2)} = 1$.
3. For the objective 3, a bamboo hat is the most convenient, umbrella the second, and raincoat the last. Therefore, we have $u_{11}^{(3)} = 2$, $u_{12}^{(3)} = 3$, and $u_{13}^{(3)} = 1$.

From these discussions, we have obtained the effect vectors for the three situations:

$$u_{11} = \left(u_{11}^{(1)}, u_{11}^{(2)}, u_{11}^{(3)}\right) = (1, 2, 2),$$

$$u_{12} = \left(u_{12}^{(1)}, u_{12}^{(2)}, u_{12}^{(3)}\right) = (2, 3, 3),$$

and

$$u_{13} = \left(u_{13}^{(1)}, u_{13}^{(2)}, u_{13}^{(3)}\right) = (3, 1, 1)\,.$$

Take the center of the spherical grey target as $r_0 = (1, 1, 1)$ and the radius as $\sqrt{2}$. That is,

$$R^3 = \{(r^{(1)}, r^{(2)}, r^{(3)}) | (r^{(1)} - 1)^2 + (r^{(2)} - 1)^2 + (r^{(3)} - 1)^2 \leq 2\}.$$

Then $u_{11} \in R^3$. That is, u_{11} is a satisfactory effect vector, the corresponding s_{11} a desirable situation. That is, when raining, using an umbrella is a desirable countermeasure.

Furthermore, we can compute the bull's-eye-distances:

$$\begin{aligned}
|u_{11} - r_0| &= [(u_{11}^{(1)} - r_0^{(1)})^2 + (u_{11}^{(2)} - r_0^{(2)})^2 + (u_{11}^{(3)} - r_0^{(3)})^2]^{\frac{1}{2}} \\
&= [(1-1)^2 + (2-1)^2 + (2-1)^2]^{\frac{1}{2}} = \sqrt{2}, \\
|u_{12} - r_0| &= [(u_{12}^{(1)} - r_0^{(1)})^2 + (u_{12}^{(2)} - r_0^{(2)})^2 + (u_{12}^{(3)} - r_0^{(3)})^2]^{\frac{1}{2}} \\
&= [(2-1)^2 + (3-1)^2 + (3-1)^2]^{\frac{1}{2}} = 3, \\
|u_{13} - r_0| &= [(u_{13}^{(1)} - r_0^{(1)})^2 + (u_{13}^{(2)} - r_0^{(2)})^2 + (u_{13}^{(3)} - r_0^{(3)})^2]^{\frac{1}{2}} \\
&= [(3-1)^2 + (1-1)^2 + (1-1)^2]^{\frac{1}{2}} = 2.
\end{aligned}$$

Because $\sqrt{2} < 2 < 3$, we conclude that $s_{11} \succ s_{13} \succ s_{12}$. That is, s_{11} is a quasi-optimum situation.

If we take the first-level effects of various objectives as 0, the second effects as 1, the third effects as 2, then we can obtain a spherical grey target such that its center is located at the origin.

If the event a_1 is changed to a complex event such as either raining and riding a bicycle or raining and carrying a heavy object, etc., then the situation effect vectors will change accordingly. Also, the relevant desirable situations, and quasi-optimum situations can be changed with the definition of event.

In Example 10.2.1, even though there does not exist an optimum situation, we have found a desirable quasi-optimum situation. This is the intelligent meaning or flexibility of grey target decision making. For example, a company sends a representative to a distinct location to negotiate and to sign a contract to buy some products. He or she might be given the order that "When the total price is around \$50,000, the quality, variety, standards, and prices are appropriate, sign the contract." Or, he or she can also be given the order: "You can sign the contract only if the total cost is \$50,000, the products are high-quality brand names, and unit price is $\times\times$ dollars." The first statement gives a grey target, where the representative is given some degree of freedom so that he or she will succeed relatively easily in the negotiation, whereas the second statement describes the bull's-eye of the target without giving the representative any room to succeed in the deal. If for the situation of rain, only the countermeasure of being most effective, most economical, and most convenient will be considered, then staying home will be the only solution.

10.3 Grey Incidence Decisions

The bull's-eye-distance of a situation effect vector is a standard measuring the superiority among situations, and the concept of degrees of incidence between a situation effect vector and optimum effect vector is another method to evaluate the superiority of situations.

Definition 10.3.1. *Assume that*

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}$$

is the situation set, and

$$u_{i_0 j_0} = \left(u_{i_0 j_0}^{(1)}, u_{i_0 j_0}^{(2)}, ..., u_{i_0 j_0}^{(s)} \right)$$

the optimum effect vector. If the situation corresponding to $u_{i_0 j_0}$ satisfies $s_{i_0 j_0} \notin S$, then $u_{i_0 j_0}$ is called the imagined optimum $effect\ vector$, and the corresponding $s_{i_0 j_0}$ the imagined optimum situation.

Proposition 10.3.1. *Assume that*

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}$$

is a situation set and the corresponding effect vector of a situation s_{ij} is

$$u_{ij} = \left(u_{ij}^{(1)}, u_{ij}^{(2)}, ..., u_{ij}^{(s)} \right),$$

$i = 1, 2, ..., n;\ j = 1, 2, ..., m.$

1. When k is an objective satisfying that the greater the effect value the better, take

$$u_{i_0 j_0}^{(k)} = \max_{\substack{1 \le i \le n \\ 1 \le j \le m}} \left\{ u_{ij}^{(k)} \right\};$$

2. When k is an objective satisfying that it is good when the effect value is close to a moderate value u_0, take

$$u_{i_0 j_0}^{(k)} = u_0;$$

3. When k is an objective satisfying that the smaller the effect value is the better, take

$$u_{i_0j_0}^{(k)} = \min_{\substack{1\le i\le n \\ 1\le j\le m}} \left\{u_{ij}^{(k)}\right\}.$$

Then

$$u_{i_0j_0} = \left(u_{i_0j_0}^{(1)}, u_{i_0j_0}^{(2)}, ..., u_{i_0j_0}^{(s)}\right)$$

is the imagined optimum effect vector.

Proposition 10.3.2. *Assume that*

$$S = \{s_{ij} = (a_i, b_j)|a_i \in A, b_j \in B\}$$

is the situation set,

$$u_{ij} = \left(u_{ij}^{(1)}, u_{ij}^{(2)}, ..., u_{ij}^{(s)}\right),$$

the effect vector of situation s_{ij}, $i = 1, 2, ..., n$; $j = 1, 2, ..., m$,

$$u_{i_0j_0} = \left(u_{i_0j_0}^{(1)}, u_{i_0j_0}^{(2)}, ..., u_{i_0j_0}^{(s)}\right)$$

the imagined optimum effect vector, and $\varepsilon_{ij}, i = 1, 2, ..., n$; $j = 1, 2, ..., m$, the degree of incidence between u_{ij} and $u_{i_0j_0}$. If $\varepsilon_{i_1j_1}$ satisfies that for any $i \in \{1, 2, ..., n\}$ with $i \neq i_1$ and $j \in \{1, 2, ..., n\}$ with $j \neq j_1$, it is always true that $\varepsilon_{i_1j_1} \geq \varepsilon_{ij}$, then $u_{i_1j_1}$ is said to be a quasi-optimum effect vector and $s_{i_1j_1}$ a quasi-optimum situation.

Grey incidence decisions can be performed as follows.

Step 1: Determine the set of events

$$A = \{a_1, a_2, ..., a_n\}$$

and the set of countermeasures

$$B = \{b_1, b_2, ..., b_m\}$$

so that the set of situations

$$S = \{s_{ij} = (a_i, b_j)|a_i \in A, b_j \in B\}$$

can be constructed.

Step 2: Choose the objectives 1,2, ..., s for the decision making.

Step 3: Gather the effect value $u_{ij}^{(k)}$ of the situation s_{ij} with objective k, $i = 1, 2, ..., n$; $j = 1, 2, ..., m$. So, it follows that

$$\begin{aligned} u^{(k)} = \ & (u_{11}^{(k)}, u_{12}^{(k)}, ..., u_{1m}^{(k)}; u_{21}^{(k)}, u_{22}^{(k)}, ..., u_{2m}^{(k)}; \\ & \quad \cdots\cdots\cdots\cdots\cdots; u_{n1}^{(k)}, u_{n2}^{(k)}, ..., u_{nm}^{(k)}), \end{aligned}$$

$k = 1, 2, ..., s$.

Step 4: Compute the average image of the situation effect sequence $u^{(k)}$ with objective k. Without loss of generality, we still use

$$u^{(k)} = (u_{11}^{(k)}, u_{12}^{(k)}, ..., u_{1m}^{(k)}; u_{21}^{(k)}, u_{22}^{(k)}, ..., u_{2m}^{(k)}; \cdots\cdots\cdots\cdots; u_{n1}^{(k)}, u_{n2}^{(k)}, ..., u_{nm}^{(k)}),$$

$k = 1, 2, \cdots, s$, to represent the average image.

Step 5: Based on the result in Step 4, write out the effect vector of situation s_{ij}

$$u_{ij} = \left(u_{ij}^{(1)}, u_{ij}^{(2)}, ..., u_{ij}^{(s)}\right),$$

$i = 1, 2, ..., n$; $j = 1, 2, ..., m$.

Step 6: Compute the imagined optimum effect vector

$$u_{i_0 j_0} = \left(u_{i_0 j_0}^{(1)}, u_{i_0 j_0}^{(2)}, ..., u_{i_0 j_0}^{(s)}\right).$$

Step 7: Compute the absolute degree of grey incidence ε_{ij} between u_{ij} and $u_{i_0 j_0}$, $i = 1, 2, ..., n$; $j = 1, 2, ..., m$.

Step 8: Based on $\max\limits_{\substack{1 \le i \le n \\ 1 \le j \le m}} \{\varepsilon_{ij}\} = \varepsilon_{i_1 j_1}$, it follows that $u_{i_1 j_1}$ is a quasi-optimum effect vector and $s_{i_1 j_1}$ a quasi-optimum situation .

Example 10.3.1. Let us look at how a grey incidence decision about a certain city's reconstruction of main traffic roads is made.

Solution: Step 1: The reconstruction of the main traffic roads is our event a_1. So, the set of events is $A = \{a_1\}$. And, denote

- The plan of separating the traffic into several roads as countermeasure b_1,
- The plan of building an expressway as countermeasure b_2,
- The plan of building a multi-level traffic line as countermeasure b_3,
- The plan of building a railway as countermeasure b_4,
- Laying a railway track on the current road foundation as countermeasure b_5, and
- The plan of building 3-dimensional traffic as countermeasure b_6;

then, we have the set of countermeasures

$$B = \{b_1, b_2, b_3, b_4, b_5, b_6\}.$$

Now, we have the set of all possible situations as follows.

$$\begin{aligned} S &= \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\} \\ &= \{s_{11}, s_{12}, s_{13}, s_{14}, s_{15}, s_{16}\}. \end{aligned}$$

Step 2: Determine 10 different objectives. Denote

- The traffic capacity (%) as objective 1,
- The total project expense (in 10,000 yuan) as objective 2,
- The total dismantling and relocation cost (in 10,000 yuan) as objective 3,
- Traffic amount (cars/hour) as objective 4,
- Speed of through traffic (km/hour) as objective 5,
- The new construction quality (%) as objective 6,
- Public charges (qualitative) as objective 7,
- Safety (qualitative) as objective 8,
- Synthesized index (no dimension) as objective 9, and
- Difficulty level (qualitative) of the project as objective 10.

Step 3: Gather the situation effect sequence $u^{(k)}$ with the objective k, for $k = 1, 2, ..., 10$. For the situation effect sequence with respect to the traffic capacity, we have

$$u^{(1)} = \left(u_i^{(1)}\right)_{i=11}^{16} = (88, 36, 62, 36, 36, 62).$$

For the situation effect sequence with respect to the total project expense, we have

$$u^{(2)} = \left(u_i^{(2)}\right)_{i=11}^{16} = (26550, 46880, 33430, 46160, 44760, 25490).$$

For the situation effect sequence with respect to the total dismantling and relocation cost, we have

$$u^{(3)} = \left(u_i^{(3)}\right)_{i=11}^{16} = (17700, 2620, 11880, 495, 495, 11800).$$

For the situation effect sequence with respect to traffic amount, we have

$$u^{(4)} = \left(u_i^{(4)}\right)_{i=11}^{16} = (2200, 800, 2000, 800, 800, 3500).$$

For the situation effect sequence with respect to speed of through traffic, we have

$$u^{(5)} = \left(u_i^{(5)}\right)_{i=11}^{16} = (25, 60, 30, 80, 60, 50)\,.$$

For the situation effect sequence with respect to the new construction quality, we have

$$u^{(6)} = \left(u_i^{(6)}\right)_{i=11}^{16} = (0.51, 0.75, 0.58, 0.7, 0.75, 0.63)\,.$$

For the situation effect sequence with respect to public charges, we have

$$\begin{aligned} u^{(7)} &= \left(u_i^{(7)}\right)_{i=11}^{16} \\ &= (\text{great, small, very great, very great, very great, great})\,. \end{aligned}$$

For the situation effect sequence with respect to safety, we have

$$\begin{aligned} u^{(8)} &= \left(u_i^{(8)}\right)_{i=11}^{16} \\ &= (\text{very bad, good, bad, very good, bad, good})\,. \end{aligned}$$

For the situation effect sequence with respect to the synthesized index, we have

$$u^{(9)} = \left(u_i^{(9)}\right)_{i=11}^{16} = (2.25, 3, 2.5, 3.25, 3, 3,)\,.$$

For the situation effect sequence with respect to the difficulty level of the project, we have

$$\begin{aligned} u^{(10)} &= \left(u_i^{(10)}\right)_{i=11}^{16} \\ &= (\text{general, very hard, hard, hardest, very hard, hard})\,. \end{aligned}$$

We now take the three qualitative objectives 7, 8, and 10 as quantitative as follows,

$$u^{(7)} = (0.5, 0.33, 0.67, 0.67, 0.67, 0.5)\,,$$

$$u^{(8)} = (0.67, 0.33, 0.5, 0.2, 0.5, 0.33)\,,$$

and

$$u^{(10)} = (0.2, 0.6, 0.4, 0.85, 0.6, 0.4)\,.$$

Step 4: Compute the average images of the situation effect sequences with objective k, and still use the same symbols to represent the average images.

$$u^{(1)} = (1.66, 0.68, 1.17, 0.68, 0.68, 1.17),$$

$$u^{(2)} = (0.71, 1.26, 0.90, 1.24, 1.20, 0.69),$$

$$u^{(3)} = (2.36, 0.35, 1.58, 0.07, 0.07, 1.57),$$

$$u^{(4)} = (1.31, 0.48, 1.19, 0.48, 0.48, 2.08),$$

$$u^{(5)} = (0.49, 1.18, 0.59, 1.57, 1.18, 0.98),$$

$$u^{(6)} = (0.78, 1.15, 0.89, 1.08, 1.15, 0.97),$$

$$u^{(7)} = (0.89, 0.59, 1.20, 1.20, 1.20, 0.89),$$

$$u^{(8)} = (1.60, 0.79, 1.19, 0.48, 1.19, 0.79),$$

$$u^{(9)} = (0.80, 1.06, 0.88, 1.15, 1.06, 1.06),$$

and

$$u^{(10)} = (0.39, 1.18, 0.79, 1.67, 1.18, 0.79).$$

Step 5: From the results in Step 4, the effect vector u_{ij} of the situation s_{ij} can be obtained as follows, $i = 1$; $j = 1, 2, ..., 6$.

$$\begin{aligned}
u_{11} &= \left(u_{11}^{(i)}\right)_{i=1}^{10} \\
&= (1.66, 0.71, 2.36, 1.31, 0.49, 0.78, 0.89, 1.60, 0.80, 0.39), \\
u_{12} &= \left(u_{12}^{(i)}\right)_{i=1}^{10} \\
&= (0.68, 1.26, 0.35, 0.48, 1.18, 1.15, 0.59, 0.79, 1.06, 1.18), \\
u_{13} &= \left(u_{13}^{(i)}\right)_{i=1}^{10} \\
&= (1.17, 0.90, 1.58, 1.19, 0.59, 0.89, 1.20, 1.19, 0.88, 0.79), \\
u_{14} &= \left(u_{14}^{(i)}\right)_{i=1}^{10} \\
&= (0.68, 1.24, 0.07, 0.48, 1.57, 1.08, 1.20, 0.48, 1.15, 1.67), \\
u_{15} &= \left(u_{15}^{(i)}\right)_{i=1}^{10} \\
&= (0.68, 1.20, 0.07, 0.48, 1.18, 1.15, 1.20, 1.19, 1.06, 1.18),
\end{aligned}$$

and

$$\begin{aligned} u_{16} &= \left(u_{16}^{(i)}\right)_{i=1}^{10} \\ &= (1.17, 0.69, 1.57, 2.08, 0.98, 0.97, 0.89, 0.79, 1.06, 0.79)\,. \end{aligned}$$

Step 6: Find the imagined optimum effect vector. Because the objective of traffic capacity is the stronger the better, we have

$$u_{i_0 j_0}^{(1)} = \max_{1 \le j \le 6} \left\{u_{1j}^{(1)}\right\} = u_{11}^{(1)} = 1.66.$$

Because the objective of the total project expense is the lower the better, we have

$$u_{i_0 j_0}^{(2)} = \min_{1 \le j \le 6} \left\{u_{1j}^{(2)}\right\} = u_{16}^{(2)} = 0.69.$$

Because the objective of the total dismantling and relocation cost is the less the better, we have

$$u_{i_0 j_0}^{(3)} = \min_{1 \le j \le 6} \left\{u_{1j}^{(3)}\right\} = u_{14}^{(3)} = u_{15}^{(3)} = 0.07.$$

Because the objective of traffic amount is the greater the better, we have

$$u_{i_0 j_0}^{(4)} = \max_{1 \le j \le 6} \left\{u_{1j}^{(4)}\right\} = u_{16}^{(4)} = 2.08.$$

Because the objective of speed of through traffic is the greater the better, we have

$$u_{i_0 j_0}^{(5)} = \max_{1 \le j \le 6} \left\{u_{1j}^{(5)}\right\} = u_{14}^{(5)} = 1.57.$$

Because the objective of the new construction quality is the higher the better, we have

$$u_{i_0 j_0}^{(6)} = \max_{1 \le j \le 6} \left\{u_{1j}^{(6)}\right\} = u_{12}^{(6)} = u_{15}^{(6)} = 1.15.$$

Because the objective of public charges is the smaller the better, we have

$$u_{i_0 j_0}^{(7)} = \min_{1 \le j \le 6} \left\{u_{1j}^{(7)}\right\} = u_{12}^{(7)} = 0.59.$$

Because the objective of safety is the more the better, we have

$$u_{i_0 j_0}^{(8)} = \min_{1 \le j \le 6} \left\{u_{1j}^{(8)}\right\} = u_{14}^{(8)} = 0.48.$$

Because the objective of synthesized index is the greater the better, we have

$$u_{i_0 j_0}^{(9)} = \max_{1 \le j \le 6} \left\{u_{1j}^{(9)}\right\} = u_{14}^{(9)} = 1.15.$$

Because the objective of difficulty level (qualitative) of the project is the easier the better, we have

$$u_{i_0j_0}^{(10)} = \min_{1\leq j\leq 6}\left\{u_{1j}^{(10)}\right\} = u_{11}^{(10)} = 0.39.$$

Therefore, the imagined optimum effect vector is

$$\begin{aligned} u_{i_0j_0} &= \left(u_{i_0j_0}^{(i)}\right)_{i=1}^{10} \\ &= (1.66, 0.69, 0.07, 2.08, 1.57, 1.15, 0.59, 0.48, 1.15, 0.39)\,. \end{aligned}$$

Step 7: Compute the absolute degree ε_{ij} of grey incidence between u_{ij} and $u_{i_0j_0}$ $i = 1$; $j = 1, 2, ..., 6$. We have that

$$\varepsilon_{11} = 0.91, \varepsilon_{12} = 0.53, \varepsilon_{13} = 0.62,$$

$$\varepsilon_{14} = 0.53, \varepsilon_{15} = 0.53, \varepsilon_{16} = 0.58.$$

Step 8: From that

$$\max_{1\leq j\leq 6}\{\varepsilon_{1j}\} = \varepsilon_{11} = 0.91,$$

it follows that u_{11} is a quasi-optimum effect vector and s_{11} a quasi-optimum situation. That is, for our project of rerouting, the plan of separating the traffic into several roads is a desirable quasi-optimum countermeasure.

In Example 10.3.1, we have treated all 10 different objectives of the decision making equally. However, in many practical situations, the degrees of attention given to all the different objectives should be different. That is, some objectives will be seen to be more important than others. To meet this requirement of putting different weights on different objectives, in the process of decision making, the subjective willingness of the decision maker needs to be considered appropriately so that the conclusions of both quantitative analysis and qualitative analysis can be supplementary to each other in order to reduce mistakes in the decision making.

Example 10.3.2. For the rerouting project in Example 10.3.1, let us take a weight η_k for objective k, $k = 1$, 2, ..., 10, as follows.

$$\eta_1 = 0.15, \eta_2 = 0.15, \eta_3 = 0.10, \eta_4 = 0.10, \eta_5 = 0.08,$$

$$\eta_6 = 0.08, \eta_7 = 0.10, \eta_8 = 0.08, \eta_9 = 0.08, \eta_{10} = 0.08.$$

Applying $\eta_k \cdot u^{(k)}$ as the situation effect vector with objective k gives

$$\eta_1 \cdot u^{(1)} = 0.15 \cdot u^{(1)} = (0.249, 0.102, 0.1755, 0.102, 0.102, 0.1755)\,,$$

$$\eta_2 \cdot u^{(2)} = 0.15 \cdot u^{(2)} = (0.1065, 0.189, 0.135, 0.186, 0.18, 0.1035)\,,$$

$$\eta_3 \cdot u^{(3)} = 0.15 \cdot u^{(3)} = (0.236, 0.035, 0.158, 0.007, 0.007, 0.157)\,,$$

$$\eta_4 \cdot u^{(4)} = 0.1 \cdot u^{(4)} = (0.131, 0.048, 0.119, 0.048, 0.048, 0.208)\,,$$

$$\eta_5 \cdot u^{(5)} = 0.08 \cdot u^{(5)} = (0.0392, 0.0944, 0.0472, 0.1256, 0.0944, 0.0784)\,,$$

$$\eta_6 \cdot u^{(6)} = 0.08 \cdot u^{(6)} = (0.0624, 0.092, 0.0712, 0.0864, 0.092, 0.0776)\,,$$

$$\eta_7 \cdot u^{(7)} = 0.1 \cdot u^{(7)} = (0.089, 0.059, 0.120, 0.120, 0.120, 0.089)\,,$$

$$\eta_8 \cdot u^{(8)} = 0.08 \cdot u^{(8)} = (0.128, 0.0632, 0.0952, 0.0384, 0.0952, 0.0632)\,,$$

$$\eta_9 \cdot u^{(9)} = 0.08 \cdot u^{(9)} = (0.064, 0.0848, 0.0704, 0.092, 0.0848, 0.0848)\,,$$

and

$$\eta_{10} \cdot u^{(10)} = 0.08 \cdot u^{(10)} = (0.0312, 0.0944, 0.0632, 0.1336, 0.0944, 0.0632)\,.$$

Similar to Example 10.3.1, take

$$\eta_1 \cdot u_{i_0 j_0}^{(1)} = \max_{1 \le j \le 6} \left\{ \eta_1 \cdot u_{1j}^{(1)} \right\} = \eta_1 \cdot u_{11}^{(1)} = 0.249,$$

$$\eta_2 \cdot u_{i_0 j_0}^{(2)} = \min_{1 \le j \le 6} \left\{ \eta_2 \cdot u_{1j}^{(2)} \right\} = \eta_2 \cdot u_{16}^{(2)} = 0.1035,$$

$$\eta_3 \cdot u_{i_0 j_0}^{(3)} = \min_{1 \le j \le 6} \left\{ \eta_3 \cdot u_{1j}^{(3)} \right\} = \eta_3 \cdot u_{14}^{(3)} = 0.007,$$

$$\eta_4 \cdot u_{i_0 j_0}^{(4)} = \max_{1 \le j \le 6} \left\{ \eta_4 \cdot u_{1j}^{(4)} \right\} = \eta_4 \cdot u_{16}^{(4)} = 0.208,$$

$$\eta_5 \cdot u_{i_0 j_0}^{(5)} = \max_{1 \le j \le 6} \left\{ \eta_5 \cdot u_{1j}^{(5)} \right\} = \eta_5 \cdot u_{14}^{(5)} = 0.1256,$$

$$\eta_6 \cdot u_{i_0 j_0}^{(6)} = \max_{1 \le j \le 6} \left\{ \eta_6 \cdot u_{1j}^{(6)} \right\} = \eta_6 \cdot u_{12}^{(6)} = 0.092,$$

$$\eta_7 \cdot u_{i_0 j_0}^{(7)} = \min_{1 \le j \le 6} \left\{ \eta_7 \cdot u_{1j}^{(7)} \right\} = \eta_7 \cdot u_{12}^{(7)} = 0.059,$$

$$\eta_8 \cdot u_{i_0 j_0}^{(8)} = \min_{1 \le j \le 6} \left\{ \eta_8 \cdot u_{1j}^{(8)} \right\} = \eta_8 \cdot u_{14}^{(8)} = 0.0384,$$

$$\eta_9 \cdot u_{i_0 j_0}^{(9)} = \max_{1 \le j \le 6} \left\{ \eta_9 \cdot u_{1j}^{(9)} \right\} = \eta_9 \cdot u_{14}^{(9)} = 0.092,$$

and

$$\eta_{10} \cdot u_{i_0 j_0}^{(10)} = \min_{1 \le j \le 6} \left\{ \eta_{10} . u_{1j}^{(10)} \right\} = \eta_{10} \cdot u_{11}^{(10)} = 0.0312.$$

So, we have the following imagined optimum situation effect vector of the weighted objectives,

$$\begin{aligned} u_{i_0 j_0} &= (\eta_i \cdot u_{i_0 j_0}^{(i)})_{i=1}^{10} \\ &= (0.249, 0.1035, 0.007, 0.208, 0.1256, \\ &\quad 0.092, 0.059, 0.0384, 0.092, 0.0312). \end{aligned}$$

The situation effect vector, corresponding to the situation $s_{ij}, i = 1$; $j = 1, 2, ..., 6$, with weighted objectives, is given as follows.

$$\begin{aligned} u_{11} &= (\eta_i \cdot u_{11}^{(i)})_{i=1}^{10} \\ &= (0.249, 0.1065, 0.236, 0.131, 0.0392, \\ &\quad 0.0624, 0.089, 0.128, 0.064, 0.0312), \end{aligned}$$

$$\begin{aligned} u_{12} &= (\eta_i \cdot u_{12}^{(i)})_{i=1}^{10} \\ &= (0.102, 0.189, 0.035, 0.048, 0.0944, \\ &\quad 0.092, 0.059, 0.0632, 0.0848, 0.0944), \end{aligned}$$

$$\begin{aligned} u_{13} &= (\eta_i \cdot u_{13}^{(i)})_{i=1}^{10} \\ &= (0.1755, 0.135, 0.158, 0.119, 0.0472, \\ &\quad 0.0712, 0.12, 0.0952, 0.0704, 0.0632), \end{aligned}$$

$$\begin{aligned} u_{14} &= (\eta_i \cdot u_{14}^{(i)})_{i=1}^{10} \\ &= (0.102, 0.186, 0.007, 0.048, 0.1256, \\ &\quad 0.0864, 0.12, 0.0384, 0.092, 0.1336), \end{aligned}$$

$$\begin{aligned} u_{15} &= (\eta_i \cdot u_{15}^{(i)})_{i=1}^{10} \\ &= (0.102, 0.18, 0.007, 0.048, 0.0944, \\ &\quad 0.092, 0.12, 0.0952, 0.0848, 0.0944), \end{aligned}$$

and

$$\begin{aligned} u_{16} &= (\eta_i \cdot u_{16}^{(i)})_{i=1}^{10} \\ &= (0.1755, 0.1035, 0.157, 0.208, 0.0784, \\ &\quad 0.0776, 0.089, 0.0632, 0.0848, 0.0632). \end{aligned}$$

Computing the absolute degree ε_{ij} of grey incidence between u_{ij} and $u_{i_0 j_0}$, for $i = 1$; $j = 1, 2, ..., 6$, gives

$$\varepsilon_{11} = 0.97, \varepsilon_{12} = 0.67, \varepsilon_{13} = 0.73,$$

$$\varepsilon_{14} = 0.66, \varepsilon_{15} = 0.66, \varepsilon_{16} = 0.79.$$

From that

$$\max_{1 \le j \le 6} \{\varepsilon_{1j}\} = \varepsilon_{11} = 0.97,$$

we can see that u_{11} is a quasi-optimum effect vector, and s_{11} a quasi-optimum situation. Therefore, our conclusion is still that for our project of rerouting, the plan of separating the traffic into several roads is a desirable quasi-optimum countermeasure.

10.4 Grey Development Decisions

Grey development decisions are performed based on development tendencies existing in the situation of interest or expected future behaviors. This kind of decision does not emphasize the current effects of the situation, while putting attention on development of changes of the situation effects with time. The idea of a grey development decision can be applied in decision making of long-term development in such areas as science and technology, education, economy, social, population, natural ecology, etc., or practically in such areas as industrial projects and municipal planning. By looking at problems from the angle of development, we can plan things feasibly in order to avoid such a situation of constructing today and dismantling tomorrow so that a huge amount of manpower and construction materials could be saved.

Definition 10.4.1. *Assume that*

$$A = \{a_1, a_2, ..., a_n\}$$

is a set of events,

$$B = \{b_1, b_2, ..., b_m\}$$

a set of countermeasures, and

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}$$

the set of situations. Then,

$$u_{ij}^{(k)} = \left(u_{ij}^{(k)}(1), u_{ij}^{(k)}(2), ..., u_{ij}^{(k)}(h)\right)$$

is called the situation effect time sequence of s_{ij} *with objective* k.

Previously, we have discussed situations of a fixed time moment. In Definition 10.4.1, as time moves, the case of constant changing situation effects has been addressed.

Proposition 10.4.1. *Assume that the situation effect time sequence of* s_{ij} *with objective* k *is given by*

$$u_{ij}^{(k)} = \left(u_{ij}^{(k)}(1), u_{ij}^{(k)}(2), ..., u_{ij}^{(k)}(h)\right),$$

and

$$\widehat{a}_{ij}^{(k)} = \begin{bmatrix} a_{ij}^{(k)} & b_{ij}^{(k)} \end{bmatrix}^T$$

the least squares estimate of the parameters in $GM(1,1)$ *model of* $u_{ij}^{(k)}$. *Then, the restored sequence through inverse accumulating of the* $GM(1,1)$ *time response sequence of* $u_{ij}^{(k)}$ *is given by*

$$\widehat{u}_{ij}^{(k)}(\ell+1) = \left(1 - e^{a_{ij}^{(k)}}\right) \cdot \left[u_{ij}^{(k)}(1) - \frac{b_{ij}^{(k)}}{a_{ij}^{(k)}}\right] \cdot e^{-a_{ij}^{(k)}\ell}.$$

Definition 10.4.2. *Assume that the restored sequence through inverse accumulating of the* $GM(1,1)$ *time response sequence of the situation effect time sequence* $u_{ij}^{(k)}$ *of the situation* s_{ij} *with objective* k *is given by*

$$\widehat{u}_{ij}^{(k)}(\ell+1) = \left(1 - e^{a_{ij}^{(k)}}\right) \cdot \left[u_{ij}^{(k)}(1) - \frac{b_{ij}^{(k)}}{a_{ij}^{(k)}}\right] \cdot e^{-a_{ij}^{(k)}\ell}.$$

When k *is an objective satisfying that the greater the effect value is the better,*

1. if

$$\max_{\substack{1\le i\le n\\ 1\le j\le m}} \left\{-a_{ij}^{(k)}\right\} = -a_{i_0j_0}^{(k)},$$

then $s_{i_0j_0}$ *is called the optimum situation of development coefficients with objective* k*;*

2. if

$$\max_{\substack{1\le i\le n\\ 1\le j\le m}} \left\{\widehat{u}_{ij}^{(k)}(h+\ell)\right\} = \widehat{u}_{i_0j_0}^{(k)}(h+\ell),$$

then $s_{i_0j_0}$ *is called the optimum situation of predictions with objective* k.

Similarly, we can define the concepts of optimum situations of development coefficients and predictions for the cases of objectives satisfying that the smaller the effect value the better, and that the closer to a moderate value the effect value the better, respectively. To this end, for an objective, satisfying that the smaller the effect value is the better, "max" in Definition 10.4.2, 1 and 2 needs to be changed to "min". For an objective k, satisfying that the closer to a moderate value the effect value is the better, we can first take the means of the development coefficients and the predicted values, respectively, and then define the optimum situation based on the distances of the development coefficients and predicted values to their means.

Definition 10.4.3. *Assume all the conditions are as in Definition 10.4.2, if* k *is an objective satisfying that the closer to a moderate value the effect value is the better.*

1. When

$$\min_{\substack{1\le i\le n\\ 1\le j\le m}} \left\{[\frac{1}{n+m}\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij}^{(k)}] - a_{ij}^{(k)}\right\}$$

$$= [\frac{1}{n+m}\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij}^{(k)}] - a_{i_0j_0}^{(k)},$$

$s_{i_0j_0}$ *is called the optimum situation of development coefficients with objective* k.

2. When

$$\min_{\substack{1\le i\le n\\1\le j\le m}} \left\{ \widehat{u}_{ij}^{(k)}(h+\ell) - [\frac{1}{n+m}\sum_{j=1}^{m}\sum_{i=1}^{n}\widehat{u}_{ij}^{(k)}(h+\ell)] \right\}$$

$$= \widehat{u}_{i_0 j_0}^{(k)}(h+\ell) - [\frac{1}{n+m}\sum_{j=1}^{m}\sum_{i=1}^{n}\widehat{u}_{ij}^{(k)}(h+\ell)],$$

$s_{i_0 j_0}$ *is called the optimum situation of predictions with objective* k.

Example 10.4.1. Let us look at a grey development decision making made for a certain industrial company for its technical innovation.

Solution: Assume that the said technical innovation is the event a_1. So, the set of events is $A = \{a_1\}$. Denote the plan for yearly gradual and partial innovation as countermeasure b_1, innovation by stages as countermeasure b_2, and one-time innovation as countermeasure b_3. Then, we have the set of countermeasures $B = \{b_1, b_2, b_3\}$. Hence, we have obtained the situation set

$$\begin{aligned} S &= \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\} \\ &= \{s_{11}, s_{12}, s_{13}\}. \end{aligned}$$

Define the objective 1 as improving the company's benefits with net profits as effect values (in 10,000 yuan). With the objective 1, the situation effect time sequences of s_{ij} are given by

$$u_{11}^{(1)} = \left(u_{11}^{(1)}(i)\right)_{i=1}^{4} = (32, 43.5, 58.1, 70.2),$$

$$u_{12}^{(1)} = \left(u_{12}^{(1)}(i)\right)_{i=1}^{4} = (23.2, 39, 69.4, 82.6),$$

and

$$u_{13}^{(1)} = \left(u_{13}^{(1)}(i)\right)_{i=1}^{4} = (12, 13.5, 81, 102.1).$$

The least squares estimates of the parameters

$$\widehat{a}_{ij}^{(1)} = \begin{bmatrix} a_{ij}^{(k)} & b_{ij}^{(k)} \end{bmatrix}^T$$

of the GM(1, 1) model of $u_{ij}^{(1)}$ are respectively given as follows:

$$\widehat{a}_{11}^{(1)} = \begin{bmatrix} a_{11}^{(k)} & b_{11}^{(k)} \end{bmatrix}^T = \begin{bmatrix} -0.23 & 32.15 \end{bmatrix}^T,$$

$$\widehat{a}_{12}^{(1)} = \begin{bmatrix} a_{12}^{(k)} & b_{12}^{(k)} \end{bmatrix}^T = \begin{bmatrix} -0.32 & 29.87 \end{bmatrix}^T,$$

and

$$\widehat{a}_{13}^{(1)} = \left[\begin{array}{cc} a_{13}^{(k)} & b_{13}^{(k)} \end{array}\right]^T = \left[\begin{array}{cc} -0.58 & 18.45 \end{array}\right]^T.$$

Because objective 1 satisfies that the greater the effect value is the better, and because

$$\max_{1\le j\le 3}\left\{-\widehat{a}_{1j}^{(1)}\right\} = 0.58 = -\widehat{a}_{13}^{(1)},$$

s_{13} is the optimum situation of development coefficients with objective 1.

If further consideration of predicted values is needed, we have

$$\begin{aligned}\widehat{u}_{11}^{(1)}(4+\ell) &= \left(1-e^{a_{11}^{(1)}}\right)\cdot\left[u_{11}^{(1)}(1)-\frac{b_{11}^{(1)}}{a_{11}^{(1)}}\right]\cdot e^{-a_{11}^{(1)}(4+\ell-1)}\\ &= 35.296\cdot e^{0.23(4+\ell-1)},\\ \widehat{u}_{12}^{(1)}(4+\ell) &= \left(1-e^{a_{12}^{(1)}}\right)\cdot\left[u_{12}^{(1)}(1)-\frac{b_{12}^{(1)}}{a_{12}^{(1)}}\right]\cdot e^{-a_{12}^{(1)}(4+\ell-1)}\\ &= 31.916\cdot e^{0.32(4+\ell-1)},\end{aligned}$$

and

$$\begin{aligned}\widehat{u}_{13}^{(1)}(4+\ell) &= \left(1-e^{a_{13}^{(1)}}\right)\cdot\left[u_{13}^{(1)}(1)-\frac{b_{13}^{(1)}}{a_{13}^{(1)}}\right]\cdot e^{-a_{13}^{(1)}(4+\ell-1)}\\ &= 19.281\cdot e^{0.58(4+\ell-1)}.\end{aligned}$$

Let $\ell = 1$. We obtain

$$\widehat{u}_{11}^{(1)}(5) = 88.57, \widehat{u}_{12}^{(1)}(5) = 114.79, \widehat{u}_{13}^{(1)}(5) = 196.20.$$

So,

$$\max_{1\le j\le 3}\left\{\widehat{u}_{1j}^{(1)}(5)\right\} = 196.20 = \widehat{u}_{13}^{(1)}(5).$$

Therefore, s_{13} is the optimum situation of predictions with objective 1. From the angle of long-term development, this company needs to have a one-time innovation done.

In Example 10.4.1, we have obtained the same optimum situation for both development coefficients and predictions. Sometimes, one might face the case that the optimum situations for development coefficients and predictions are different. However, Theorem 10.4.1 below tells us that eventually, the optimum situations of development coefficients and predictions would surely approach a fixed situation.

Theorem 10.4.1. *Assume that k is an objective satisfying that the greater the effect value is the better, $s_{i_0 j_0}$ the optimum situation of development coefficients with objective k, that is,*

$$-a_{i_0 j_0}^{(k)} = \max_{\substack{1\le i\le n\\ 1\le j\le m}}\left\{-a_{ij}^{(k)}\right\},$$

and $\widehat{u}_{i_0j_0}^{(k)}(h+\ell+1)$ *is the predicted value for the situation effect of* $s_{i_0j_0}$. *Then, there must exist an* $\ell_0 > 0$ *such that*

$$\widehat{u}_{i_0j_0}^{(k)}(h+\ell_0+1) = \max_{\substack{1\le i\le n\\ 1\le j\le m}} \{\widehat{u}_{ij}^{(k)}(h+\ell_0+1)\}.$$

That is, in a sufficiently distant future, $s_{i_0j_0}$ *will become the optimum situation of predictions.*

Proof: 1. If the situation effect time sequences with objective k are all increasing, then for any $i \in \{1,\ 2,\ ...,n\}$ and $j \in \{1,\ 2,\ ...,m\}$, we have $-a_{ij}^{(k)} > 0$. When at least one of $i \neq i_0$ and $j \neq j_0$ holds true, $-a_{i_0j_0}^{(k)} > -a_{ij}^{(k)}$. So, there exists $\delta_{ij}^{(k)} > 0$ such that

$$-a_{i_0j_0}^{(k)} = -a_{ij}^{(k)} + \delta_{ij}^{(k)}.$$

Let

$$c_{ij}^{(k)} = \left(1 - e^{a_{ij}^{(k)}}\right) \cdot \left[u_{ij}^{(k)}(1) - \frac{b_{ij}^{(k)}}{a_{ij}^{(k)}}\right].$$

Then,

$$\begin{aligned}
\widehat{u}_{i_0j_0}^{(k)}(h+\ell+1) &= c_{i_0j_0}^{(k)} \cdot e^{-a_{i_0j_0}^{(k)}(h+\ell)} \\
&= c_{i_0j_0}^{(k)} \cdot e^{-a_{ij}^{(k)}(h+\ell)+\delta_{ij}^{(k)}(h+\ell)} \\
&= c_{ij}^{(k)} \cdot e^{-a_{ij}^{(k)}(h+\ell)} \cdot \frac{c_{i_0j_0}^{(k)}}{c_{ij}^{(k)}} \cdot e^{\delta_{ij}^{(k)}(h+\ell)} \\
&= \widehat{u}_{ij}^{(k)}(h+\ell+1) \cdot \frac{c_{i_0j_0}^{(k)}}{c_{ij}^{(k)}} \cdot e^{\delta_{ij}^{(k)}(h+\ell)}.
\end{aligned}$$

Based on the assumption that $\delta_{ij}^{(k)} > 0$, it follows that there is an ℓ_0, such that

$$\frac{c_{i_0j_0}^{(k)}}{c_{ij}^{(k)}} \cdot e^{\delta_{ij}^{(k)}(h+\ell_0)} > 1.$$

Therefore,

$$\widehat{u}_{i_0j_0}^{(k)}(h+\ell_0+1) > \widehat{u}_{ij}^{(k)}(h+\ell_0+1).$$

From the arbitrariness of $i \in \{1,2,...,n\}$ and $j \in \{1,2,...,m\}$, it follows that

$$\widehat{u}_{i_0j_0}^{(k)}(h+\ell_0+1) = \max_{\substack{1\le i\le n\\ 1\le j\le m}} \{\widehat{u}_{ij}^{(k)}(h+\ell_0+1)\}.$$

2. If all the situation effect time sequences with objective k are all decreasing, then for any $i \in \{1, 2, ..., n\}$ and $j \in \{1, 2, ..., m\}$, we have $-a_{ij}^{(k)} < 0$. Similarly, from $-a_{i_0j_0}^{(k)} > -a_{ij}^{(k)}$, we have $\delta_{ij}^{(k)} > 0$ such that

$$-a_{i_0j_0}^{(k)} = -a_{ij}^{(k)} + \delta_{ij}^{(k)}.$$

The rest of our discussion is similar to 1 and is omitted here.

For objectives, satisfying either that the smaller the effect value is the better or that the closer to a moderate value the effect value is the better, it can be shown that similar results as in Theorem 10.4.1 also hold true. Careful readers might have noticed that Theorem 10.4.1 does not state the case where there exist some increasing and decreasing sequences among situation effect time sequences at the same time. In fact, for objectives satisfying that the greater the effect value is the better, there is no need to consider decreasing situation effect time sequences. For objectives satisfying that the smaller the effect value is the better, in discussions, all increasing situation effect time sequences need to be deleted in advance. For the objectives satisfying that the closer to a moderate value the effect value is the better, one can consider only either increasing or decreasing situation effect time sequences dependent on the circumstances involved.

10.5 Grey Statistical Decisions

The method of grey statistical decision is applicable to the cases of collective decision making with many decision making units, such as the cases involving different departments, units, or individuals. The function of grey statistical decisions is to synthesize, to evaluate, and to make final decisions from among the set of different opinions and decision plans of the various parties involved.

Definition 10.5.1. *Assume that $a_i, i = 1, 2, ..., n$, are the units involved in the decision making process,*

$$A = \{a_1, a_2, ..., a_n\}$$

the decision making group, $b_j, j = 1, 2, ..., m$, the decision schemes,

$$B = \{b_1, b_2, ..., b_m\}$$

the set of decision schemes, x_{ij}, $i = 1, 2, ..., n$; $j = 1, 2, ..., m$, the evaluation value about scheme j of unit i, $f^k(\cdot), k = 1, 2, ..., s$, the whitenization weight function for grey class k, η_i, and *$i = 1, 2, ..., n$, the decision weight of unit* *i, satisfying* $\sum_{i=1}^{n}\eta_i = 1$. *Then*

$$\sigma_j^k = \frac{\sum_{i=1}^{n} f^k(x_{ij}) \cdot \eta_i}{\sum_{k=1}^{s}\sum_{i=1}^{n} f^k(x_{ij}) \cdot \eta_i},$$

$j = 1, 2, ..., m$; $k = 1, 2, ..., s$, *is called the decision coefficient for scheme* j *to belong to grey class* k.

Here, all decision units might have the same decision weight. That is, $\eta_i = \frac{1}{n}$, i = 1, 2, ..., n, or they might be different.

Definition 10.5.2. *The following*

$$\sigma_j = \left(\sigma_j^1, \sigma_j^2, ..., \sigma_j^s,\right),$$

$j = 1, 2, ..., m$, *is called the vector of the decision coefficients of scheme* j.

Definition 10.5.3. *If*

$$\max_{1 \le k \le s} \left\{\sigma_j^k\right\} = \sigma_j^{k^*},$$

then we say that scheme j *belongs to grey class* k^*.

A grey statistical decision can be made by going through similar steps as in Section 6.5.

Example 10.5.1. Let us now look at a grey statistical decision we made about the "Vitalizing the City through Science and Technology" plan of a certain city.

The governing body of the city, which, due to an agreement, we cannot name here, had organized several groups of experts to work out three practical schemes on how to "Vitalize × × × City through Progress in Science and Technology." Each of the schemes had its own characteristics. These schemes are denoted as b_1, b_2, and b_3, respectively. So, the set of decision schemes is $B = \{b_1, b_2, b_3\}$.

Next, we organized five groups of experts to evaluate the three schemes. That is, we had decision-making units a_1, a_2, a_3, a_4, and a_5 and the decision-making group $A = \{a_1, a_2, a_3, a_4, a_5\}$, the matrix of evaluation values done by the decision making group A on the decision schemes in B as follows.

$$C = [x_{ij}]_{5\times 3} = \begin{bmatrix} 80 & 60 & 40 \\ 60 & 50 & 50 \\ 75 & 70 & 60 \\ 90 & 80 & 80 \\ 50 & 70 & 60 \end{bmatrix}.$$

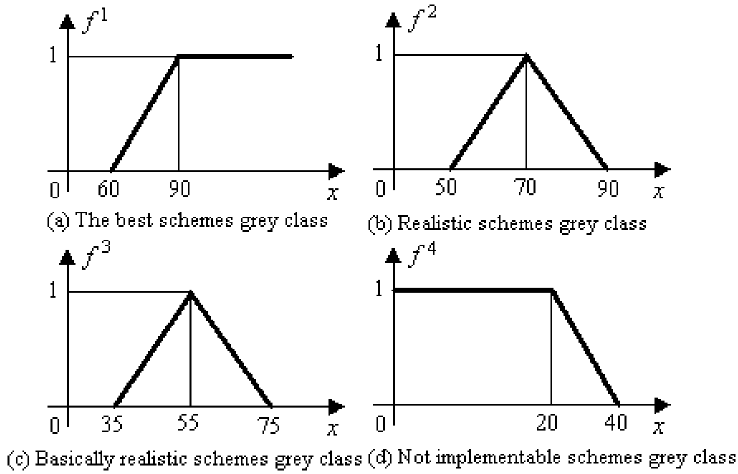

FIGURE 10.1. The whitenization weight functions of the grey classes

Let us in the following do a grey statistical evaluation and decision making based on the four grey classes: the best, realistic, basically realistic, and not implementable.

Solution: The whitenization weight function $f^k(\cdot)$ of grey class k, $k = 1, 2, 3, 4$, is given in Figure 10.1.

From the graphs of the whitenization weight functions in Figure 10.1, we have that

$$f^1(x) = \begin{cases} 0, & x < 60 \\ \dfrac{x-60}{30}, & 60 \le x \le 90 \\ 1, & x > 90; \end{cases}$$

$$f^2(x) = \begin{cases} 0, & x < 50 \\ \dfrac{x-50}{20}, & 50 \le x \le 70 \\ \dfrac{90-x}{20}, & 70 < x \le 90 \\ 0, & x > 90; \end{cases}$$

$$f^3(x) = \begin{cases} 0, & x < 35 \\ \dfrac{x-35}{20}, & 35 \le x \le 55 \\ \dfrac{75-x}{20}, & 55 < x \le 75 \\ 0, & x > 75; \end{cases}$$

and

$$f^4(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 20 \\ \dfrac{40-x}{20}, & 20 < x \le 40 \\ 0, & x > 40. \end{cases}$$

Assume that the decision-making weights of the individual decision units a_i, $i = 1, 2, 3, 4, 5$, are given as follows,

$$\eta_1 = 0.25, \eta_2 = 0.25, \eta_3 = 0.2, \eta_4 = 0.2, \eta_5 = 0.1.$$

Then, for scheme 1, when $k = 1$,

$$\begin{aligned} \sum_{i=1}^{5} f^1(x_{i1}) \cdot \eta_i &= f^1(80) \cdot 0.25 + f^1(60) \cdot 0.25 + f^1(75) \cdot 0.2 \\ &\quad + f^1(90) \cdot 0.2 + f^1(50) \cdot 0.1 \\ &= \tfrac{20}{30} \cdot 0.25 + 0 \cdot 0.25 + \tfrac{15}{30} \cdot 0.2 \\ &\quad + 1 \cdot 0.2 + 0 \cdot 0.1 = 0.47; \end{aligned}$$

when $k = 2$,

$$\begin{aligned} \sum_{i=1}^{5} f^2(x_{i1}) \cdot \eta_i &= f^2(80) \cdot 0.25 + f^2(60) \cdot 0.25 + f^2(75) \cdot 0.2 \\ &\quad + f^2(90) \cdot 0.2 + f^2(50) \cdot 0.1 \\ &= \tfrac{10}{20} \cdot 0.25 + \tfrac{10}{20} \cdot 0.25 + \tfrac{15}{20} \cdot 0.2 \\ &\quad + 0 \cdot 0.2 + 0 \cdot 0.1 = 0.4; \end{aligned}$$

when $k = 3$,

$$\begin{aligned}\sum_{i=1}^{5} f^3(x_{i1}) \cdot \eta_i &= f^3(80) \cdot 0.25 + f^3(60) \cdot 0.25 + f^3(75) \cdot 0.2 \\ &\quad + f^3(90) \cdot 0.2 + f^3(50) \cdot 0.1 \\ &= 0 \cdot 0.25 + \tfrac{15}{20} \cdot 0.25 + 0 \cdot 0.2 \\ &\quad + 0 \cdot 0.2 + \tfrac{15}{20} \cdot 0.1 = 0.2625;\end{aligned}$$

when $k = 4$,

$$\begin{aligned}\sum_{i=1}^{5} f^4(x_{i1}) \cdot \eta_i &= f^4(80) \cdot 0.25 + f^4(60) \cdot 0.25 + f^4(75) \cdot 0.2 \\ &\quad + f^4(90) \cdot 0.2 + f^4(50) \cdot 0.1 \\ &= 0 \cdot 0.25 + 0 \cdot 0.25 + 0 \cdot 0.2 \\ &\quad + 0 \cdot 0.2 + 0 \cdot 0.1 = 0.\end{aligned}$$

So, we have

$$\sum_{k=1}^{4}\sum_{i=1}^{5} f^k(x_{i1}) \cdot \eta_i = 0.47 + 0.4 + 0.2625 + 0 = 1.1325.$$

From that

$$\sigma_1^k = \frac{\sum_{i=1}^{5} f^k(x_{i1}) \cdot \eta_i}{\sum_{k=1}^{4}\sum_{i=1}^{5} f^k(x_{i1}) \cdot \eta_i},$$

$k = 1, 2, 3, 4$, it follows that

$$\sigma_1 = \left(\sigma_1^1, \sigma_1^2, \sigma_1^3, \sigma_1^4\right) = (0.42, 0.35, 0.23, 0).$$

For scheme 2, similar calculation can be done to produce that

$$\sum_{i=1}^{5} f^1(x_{i2}) \cdot \eta_i = 0.23, \sum_{i=1}^{5} f^2(x_{i2}) \cdot \eta_i = 0.525,$$

$$\sum_{i=1}^{5} f^3(x_{i2}) \cdot \eta_i = 0.45, \sum_{i=1}^{5} f^4(x_{i2}) \cdot \eta_i = 0,$$

and

$$\sum_{k=1}^{4}\sum_{i=1}^{5} f^k(x_{i2}) \cdot \eta_i = 1.205.$$

So,

$$\sigma_2 \quad = \left(\sigma_2^1, \sigma_2^2, \sigma_2^3, \sigma_2^4\right) = (0.19, 0.44, 0.37, 0)\,.$$

For scheme 3, again we can obtain similarly

$$\sum_{i=1}^{5} f^1(x_{i3}) \cdot \eta_i = 0.23, \sum_{i=1}^{5} f^2(x_{i3}) \cdot \eta_i = 0.525,$$

$$\sum_{i=1}^{5} f^3(x_{i3}) \cdot \eta_i = 0.45, \sum_{i=1}^{5} f^4(x_{i3}) \cdot \eta_i = 0,$$

and

$$\sum_{k=1}^{4}\sum_{i=1}^{5} f^k(x_{i3}) \cdot \eta_i = 0.855.$$

So,

$$\sigma_3 \quad = \left(\sigma_3^1, \sigma_3^2, \sigma_3^3, \sigma_3^4\right) = (0.15, 0.29, 0.56, 0)\,.$$

Because

$$\max_{1 \le k \le 4} \left\{\sigma_1^k\right\} = 0.42 = \sigma_1^1,$$

$$\max_{1 \le k \le 4} \left\{\sigma_2^k\right\} = 0.44 = \sigma_2^2,$$

and

$$\max_{1 \le k \le 4} \left\{\sigma_3^k\right\} = 0.56 = \sigma_3^3,$$

we can see that scheme 1 is the best, scheme 2 is realistic and implementable, and scheme 3 is basically realistic and implementable. Because all these three schemes are possible to be implemented, it is suggested that scheme 1 should be used as the foundation, and by merging all desirable features in schemes 2 and 3, we would obtain a more satisfactory and comprehensive plan for "vitalizing the city through science and technology."

When using the theory of grey statistical decisions, different meanings can be assigned to the decision-making group, decision schemes, and decision grey classes according to the circumstances under consideration. For example, when a company is making decisions on their products, the decision-making group consists of all relevant production units, such as different departments, workshops, etc. The decision schemes will include different products. And decision-making grey classes will stand for various investment requirements. In a decision making on teaching, the decision-making group can consist of all relevant classes, departments, and administrative offices. The decision schemes will mean different courses or programs. And the decision-making grey classes will represent different teaching or credit hours or educational goals. In a decision-making about agricultural productions, the decision-making group will consist of relevant farmers, related offices or organizations about agricultural technology, and/or relevant administrative organizations or offices for agricultural productions. The decision schemes will mean different crops. And the decision-making grey classes will represent various planting areas.

10.6 Grey Cluster Decisions

The thought of grey cluster decision is useful for synthetical evaluations about some objects based on several different criteria so that decisions can be made about whether an object meets some given standards. Grey cluster decisions are often applied to an area such as classification of people. For example, school students can be classified based on their capacity to receive information, their comprehension ability, and their individual potential so that different teaching methods can be applied and that different students can be enrolled in different programs. For example again, employees, technicians, and/or administrators of a business can be synthetically evaluated based on different sets of standards in order to determine whether they qualify for a certain office or position or rank or promotion, etc.

Definition 10.6.1. *Assume that there are n objects, m criteria, and s different decision classes. Let x_{ij}, $i = 1, 2, ..., n$; $j = 1, 2, ..., m$, be the observation value of object i with respect to criterion j, $f_j^k(\cdot)$, $j = 1, 2, ...,$ m, $k = 1, 2, ..., s$, the whitenization weight function for criterion j to be in grey class k, and η_j, $j = 1, 2, ..., m$, the synthetic decision weight of criterion j, satisfying that $\sum_{j=1}^{m} \eta_j = 1$. Then,*

$$\sigma_i^k = \sum_{j=1}^{m} f_j^k(x_{ij}) \cdot \eta_j$$

is called the decision coefficient for object i to be in grey class k.

Definition 10.6.2. *The following*

$$\sigma_i = \left(\sigma_i^1, \sigma_i^2, ..., \sigma_i^s\right),$$

$i = 1, 2, ..., n$, is called the vector of decision coefficients of object i.

Definition 10.6.3. *If*

$$\max_{1 \le k \le s} \left\{\sigma_i^k\right\} = \sigma_i^{k^*},$$

then we say that decision object i belongs to grey class k^*.

When solving practical problems, we often face a situation such that many objects belong to the same grey class, and the size of the class is limited. In this case, decisions can be made on which objects should be deleted from the class based on the magnitude of the decision coefficients of these objects.

Definition 10.6.4. *Assume that*

$$\max_{1\le k\le s}\left\{\sigma_{i_1}^{k}\right\}=\sigma_{i_1}^{k^*},\ \max_{1\le k\le s}\left\{\sigma_{i_2}^{k}\right\}=\sigma_{i_2}^{k^*},$$

and

$$\sigma_{i_1}^{k^*}>\sigma_{i_2}^{k^*}.$$

Then, we say that in grey class k^*, object i_1 is superior to object i_2.

Definition 10.6.5. *Assume that*

$$\max_{1\le k\le s}\left\{\sigma_{i_1}^{k}\right\}=\sigma_{i_1}^{k^*},$$

$$\max_{1\le k\le s}\left\{\sigma_{i_2}^{k}\right\}=\sigma_{i_2}^{k^*},$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\max_{1\le k\le s}\left\{\sigma_{i_h}^{k}\right\}=\sigma_{i_h}^{k^*},$$

and

$$\sigma_{i_1}^{k^*}>\sigma_{i_2}^{k^*}>\cdots>\sigma_{i_h}^{k^*}.$$

If the size of grey class k^* is ℓ objects, then $i_1, i_2, \cdots, i_\ell$ are called the objects accepted by grey class k^*, and $i_{\ell+1}, i_{\ell+2}$, ..., i_h the candidates of grey class k^*.

Example 10.6.1. A company has six middle-ranked technicians. Their observational values x_{ij}, $i = 1,2,3,4,5,6$; $j = 1,2,3,4,5$, with respect to the following five criteria: community service, work accomplished, level of mastery of subject matter, research achievements and publications, foreign language, are given in the following matrix.

$$C=[x_{ij}]_{6\times 5}=\begin{bmatrix}85 & 100 & 98 & 100 & 90\\ 88 & 100 & 60 & 10 & 74\\ 90 & 68 & 63 & 10 & 45\\ 87 & 89 & 63 & 15 & 56\\ 70 & 80 & 79 & 40 & 82\\ 75 & 76 & 77 & 32 & 79\end{bmatrix}.$$

Let us make a cluster decision based on the three grey classes: competent, qualified, not qualified. If two people can be promoted, determine those who should be promoted.

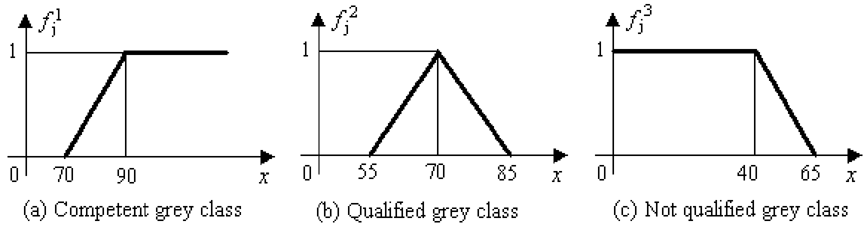

FIGURE 10.2. The whitenization weight functions of the five decision criteria with respect to the three grey classes

Solution: Assume that the whitenization weight functions of the five decision criteria with respect to the three grey classes are shown in Figure 10.2.

From Figure 10.2, it follows that

$$f_j^1(x) = \begin{cases} 0, & x < 70 \\ \dfrac{x-70}{20}, & 70 \le x \le 90 \\ 1, & x > 90; \end{cases}$$

$$f_j^2(x) = \begin{cases} 0, & x < 55 \\ \dfrac{x-55}{15}, & 55 \le x \le 70 \\ \dfrac{85-x}{15}, & 70 < x \le 85 \\ 0, & x > 85; \end{cases}$$

and

$$f_j^3(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 40 \\ \dfrac{65-x}{25}, & 40 < x \le 65 \\ 0, & x > 65, \end{cases}$$

where $j = 1, 2, 3, 4, 5$. That is, the whitenization weight functions with different criteria are identical.

Assume again that the decision weights of the criteria are given as follows:

$$\eta_1 = 0.22, \eta_2 = 0.22, \eta_3 = 0.22, \eta_4 = 0.20, \eta_5 = 0.14.$$

Then, we have

$$\begin{aligned}
\sigma_1^1 &= \sum_{j=1}^{5} f_j^1(x_{1j}) \cdot \eta_j \\
&= f_1^1(85) \cdot 0.22 + f_2^1(100) \cdot 0.22 + f_3^1(98) \cdot 0.22 \\
&\quad + f_4^1(100) \cdot 0.2 + f_5^1(90) \cdot 0.14 \\
&= \tfrac{15}{200} \cdot 0.22 + 1 \cdot 0.22 + 1 \cdot 0.2 \\
&\quad + 1 \cdot 0.2 + 1 \cdot 0.14 = 0.945; \\
\sigma_1^2 &= \sum_{j=1}^{5} f_j^2(x_{1j}) \cdot \eta_j \\
&= f_1^2(85) \cdot 0.22 + f_2^2(100) \cdot 0.22 + f_3^2(98) \cdot 0.22 \\
&\quad + f_4^2(100) \cdot 0.2 + f_5^2(90) \cdot 0.14 = 0;
\end{aligned}$$

and

$$\begin{aligned}
\sigma_1^3 &= \sum_{j=1}^{5} f_j^3(x_{1j}) \cdot \eta_j \\
&= f_1^3(85) \cdot 0.22 + f_2^3(100) \cdot 0.22 + f_3^3(98) \cdot 0.22 \\
&\quad + f_4^3(100) \cdot 0.2 + f_5^3(90) \cdot 0.14 = 0.
\end{aligned}$$

Therefore,

$$\sigma_1 = \left(\sigma_1^1, \sigma_1^2, \sigma_1^3\right) = (0.945, 0, 0).$$

Similarly, we can obtain

$$\begin{aligned}
\sigma_2 = \left(\sigma_2^1, \sigma_2^2, \sigma_2^3\right) &= (0.446, 0.176, 0.244), \\
\sigma_3 = \left(\sigma_3^1, \sigma_3^2, \sigma_3^3\right) &= (0.22, 0.308, 0.312), \\
\sigma_4 = \left(\sigma_4^1, \sigma_4^2, \sigma_4^3\right) &= (0.396, 0.127, 0.268), \\
\sigma_5 = \left(\sigma_5^1, \sigma_5^2, \sigma_5^3\right) &= (0.293, 0.409, 0.20),
\end{aligned}$$

and

$$\sigma_6 = \left(\sigma_6^1, \sigma_6^2, \sigma_6^3\right) = (0.297, 0.452, 0.20).$$

From

$$\max_{1\le k\le 3}\left\{\sigma_1^k\right\}=0.945=\sigma_1^1,\ \max_{1\le k\le 3}\left\{\sigma_2^k\right\}=0.446=\sigma_2^1,$$

$$\max_{1\le k\le 3}\left\{\sigma_3^k\right\}=0.312=\sigma_3^3,\ \max_{1\le k\le 3}\left\{\sigma_4^k\right\}=0.396=\sigma_4^1,$$

$$\max_{1\le k\le 3}\left\{\sigma_5^k\right\}=0.409=\sigma_5^2,\ \max_{1\le k\le 3}\left\{\sigma_6^k\right\}=0.452=\sigma_6^2,$$

it follows that technicians with codes 1, 2, and 4 belong to the grey class of being competent, technicians with codes 5 and 6 belong to the grey class of being qualified, and the technician with code 3 is not qualified for the job.

Since

$$\sigma_1^1=0.945>\sigma_2^1=0.446>\sigma_4^1=0.396,$$

if there exist only two opportunities for promotion, the people who should be promoted are technicians coded 1 and 2. Even though number 4 is in the grey class of being competent, due to the fact that only two can be promoted, he or she will have to be on the list of candidates.

10.7 Multiple-Target-Situation Decisions with a Synthesized Target

In this section, we study how to make decisions for multiple-target situations to meet a synthesized criterion.

Definition 10.7.1. *Assume that*

$$A=\{a_1,a_2,...,a_n\}$$

is the set of events,

$$B=\{b_1,b_2,...,b_m\}$$

the set of countermeasures,

$$S=A\times B=\{s_{ij}=(a_i,b_j)|a_i\in A,b_j\in B\}$$

the set of situations, and $u_{ij}^{(k)}$, $i=1,2,...,n,j=1,2,...,m$, *is the observational effect value of situation* $s_{ij}\in S$ *with objective* k. *Then,*

1. The following

$$r_{ij}^{(k)}=\frac{u_{ij}^{(k)}}{\max_i\max_j\left\{u_{ij}^{(k)}\right\}}$$

is called an upper effect measure;

2. The following

$$r_{ij}^{(k)} = \frac{\min_i \min_j \left\{ u_{ij}^{(k)} \right\}}{u_{ij}^{(k)}}$$

is called a $lower\ effect\ measure$*; and*

3. The following

$$r_{ij}^{(k)} = \frac{u_{i_0 j_0}^{(k)}}{u_{i_0 j_0}^{(k)} + |u_{ij}^{(k)} - u_{i_0 j_0}^{(k)}|}$$

is called a $moderate\ effect\ measure$*, where* $u_{i_0 j_0}^{(k)}$ *is a fixed moderate effect value with objective* k.

The concept of upper effect measure reflects the distance of the observational effect value from the maximum observational effect value. The concept of lower effect measure indicates the distance of the observational effect value from the minimum observational effect value. And the concept of moderate effect measure tells the distance of the observational effect value from the fixed moderate effect value. In multiple-target-situation decisions with a synthesized target, if we need the observational effect value to be like "the greater the better", "the more the better", etc., we can apply the concept of upper effect measure. If we need the observational effect value to be like "the smaller the better", "the fewer the better", etc., we can apply the concept of lower effect measure. If we need the observational effect value to be like "not too great, not too small", "not too many, not too few", etc., we can apply the concept of moderate effect measure.

Proposition 10.7.1. *The three effect measures* $r_{ij}^{(k)}$, $i = 1, 2, ..., n$, $j =$ *1,2, ...,* m, $k = 1, 2, ..., s$, *as given in Definition 10.7.1, satisfy the following.*

1. $r_{ij}^{(k)}$ *has no dimension;*
2. $r_{ij}^{(k)} \in [0, 1]$; *and*
3. *The more ideal the effect is, the greater* $r_{ij}^{(k)}$ *is.*

Definition 10.7.2. *The following matrix*

$$R^{(k)} = \left[r_{ij}^{(k)} \right]_{n \times m} = \begin{bmatrix} r_{11}^{(k)} & r_{12}^{(k)} & \cdots & r_{1m}^{(k)} \\ r_{21}^{(k)} & r_{22}^{(k)} & \cdots & r_{2m}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1}^{(k)} & r_{n2}^{(k)} & \cdots & r_{nm}^{(k)} \end{bmatrix}$$

is called the $matrix\ of\ uniform\ effect\ measures$ *of the situation set* S *with objective* k.

Definition 10.7.3. *Suppose that* $s_{ij} \in S$; *then*

$$r_{ij} = \left(r_{ij}^{(1)}, r_{ij}^{(2)}, ..., r_{ij}^{(s)}\right)$$

is called the vector of uniform effect measures of the situation s_{ij}.

Definition 10.7.4. *Assume that* η_k, $k = 1, 2, ..., s$, *is the decision weight of objective* k, *satisfying* $\sum_{k=1}^{s} \eta_k = 1$. *Then,*

$$\sum_{k=1}^{s} r_{ij}^{(k)} \cdot \eta_k$$

is called the synthetic effect measure of the situation s_{ij}, *which is still denoted as* r_{ij}. *That is,*

$$r_{ij} = \sum_{k=1}^{s} r_{ij}^{(k)} \cdot \eta_k.$$

Definition 10.7.5. *The following*

$$R = [r_{ij}]_{n \times m} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}$$

is called the matrix of synthetic effect measures.

Definition 10.7.6. *1. If*

$$\max_{1 \le j \le m} \{r_{ij}\} = r_{ij_0},$$

then b_{j_0} *is called the optimum countermeasure of the event* a_i.

2. If

$$\max_{1 \le j \le m} \{r_{ij}\} = r_{i_0 j},$$

then a_{i_0} *is called the optimum event corresponding to the countermeasure* b_j. *And*

3. If

$$\max_{1 \le i \le n} \max_{1 \le j \le m} \{r_{ij}\} = r_{i_0 j_0},$$

then $s_{i_0 j_0}$ *is called the optimum situation.*

Each multiple-target-situation decision with a synthesized target can be performed according to the following steps.

Step 1: Based on the set of events

$$A = \{a_1, a_2, ..., a_n\}$$

and the set

$$B = \{b_1, b_2, ..., b_m\}$$

of countermeasures, the following situation set is constructed

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}.$$

Step 2: Determine the objectives $k = 1, 2, ..., s$.

Step 3: For each objective $k = 1, 2, ..., s$, find the relevant observational effect matrix

$$U^{(k)} = \left[u_{ij}^{(k)}\right]_{n\times m}.$$

Step 4: Compute the matrix of uniform effect measures with objective k

$$R^{(k)} = \left[r_{ij}^{(k)}\right]_{n\times m},$$

for $k = 1, 2, ..., s$.

Step 5: Determine decision weights for all objectives: $\eta_1, \eta_2, ..., \eta_s$.

Step 6: From

$$r_{ij} = \sum_{k=1}^{s} r_{ij}^{(k)} \cdot \eta_k,$$

compute the matrix of synthetic effect measures

$$R = [r_{ij}]_{n\times m}.$$

Step 7: Determine the optimum situation $s_{i_0 j_0}$.

Example 10.7.1. In a certain county, there are three economic districts. Other than that each of them needs to develop farming, they have decided to develop individual but cooperative businesses in different areas of "forestry," "livestock husbandry," and "industry and services," with their own emphasis of development in order to achieve optimum economical effects. Here, they have three objectives for their decision making:

Objective 1: average per capita income;

Objective 2: occupied capital of each ¥100 income; and

Objective 3: occupied labor force of each ¥100 income.

We use $a_i, i = 1, 2, 3$, to represent the economic districts with a_1 = district 1, a_2 = district 2, and a_3 = district 3; and $b_j, j = 1, 2, 3$, for different kinds of businesses with b_1 = forestry, b_2 = livestock husbandry, and b_3 = industry and services. Then, we have obtained the set of events

$$A = \{a_1, a_2, a_3\},$$

the set of countermeasures

$$B = \{b_1, b_2, b_3\},$$

and the set of situations

$$S = \{s_{ij} = (a_i, b_j) | a_i \in A, b_j \in B\}.$$

Suppose that $u_{ij}^{(k)}$ is the observational effect value of situation s_{ij} with objective k, $i = 1, 2, 3$, $j = 1, 2, 3$, $k = 1, 2, 3$.

For objective 1, the observational effect matrix is

$$U^{(1)} = \left[u_{ij}^{(1)}\right]_{3\times3} = \begin{bmatrix} 0.55 & 22.4 & 3.9 \\ 0.9 & 4.4 & 14 \\ 1.14 & 5.3 & 4.9 \end{bmatrix},$$

where the unit is ¥100.

For objective 2, the observational effect matrix is

$$U^{(2)} = \left[u_{ij}^{(2)}\right]_{3\times3} = \begin{bmatrix} 0.8 & 3 & 3.5 \\ 0.6 & 2 & 4 \\ 0.1 & 0.9 & 5 \end{bmatrix}$$

with ¥100 being its unit.

For objective 3, the observational effect matrix is

$$U^{(3)} = \left[u_{ij}^{(3)}\right]_{3\times3} = \begin{bmatrix} 0.3 & 1.8 & 1 \\ 0.7 & 1 & 1.4 \\ 0.9 & 1.4 & 0.8 \end{bmatrix},$$

whose unit is person.

Because objective 1 is the average per capita income, it satisfies that the greater the better. So, we employ the upper effect measure. Because

$$u_{i_0j_0}^{(1)} = \max_{1\le i\le 3} \max_{1\le j\le 3} \left\{u_{ij}^{(1)}\right\} = 22.4,$$

from

$$r_{ij}^{(1)} = \frac{u_{ij}^{(1)}}{u_{i_0j_0}^{(1)}} = \frac{u_{ij}^{(1)}}{22.4},$$

it follows that we can obtain the matrix of uniform effect measures of objective 1 as follows.

$$R^{(1)} = \left[r_{ij}^{(1)}\right]_{3\times 3} = \begin{bmatrix} 0.025 & 1 & 0.174 \\ 0.040 & 0.196 & 0.625 \\ 0.051 & 0.237 & 0.219 \end{bmatrix}.$$

Because objective 2 stands for the occupied capital of each ¥100 income, it is hoped to satisfy that the less the better. So, we employ the lower effect measure. Because

$$u_{i_0j_0}^{(2)} = \min_{1\le i\le 3}\min_{1\le j\le 3}\left\{u_{ij}^{(2)}\right\} = 0.1,$$

from

$$r_{ij}^{(2)} = \frac{u_{i_0j_0}^{(2)}}{u_{ij}^{(2)}} = \frac{0.1}{u_{ij}^{(2)}},$$

it follows that we can obtain the matrix of uniform effect measures of objective 2 as follows.

$$R^{(2)} = \left[r_{ij}^{(2)}\right]_{3\times 3} = \begin{bmatrix} 0.125 & 0.033 & 0.029 \\ 0.167 & 0.050 & 0.025 \\ 1 & 0.111 & 0.020 \end{bmatrix}.$$

Objective 3 represents the occupied labor force of each ¥100 income. Based on the relevant Chinese policies and practical conditions of the county of our study, requiring appropriately arranging people to work, we employ the moderate effect measure. Take the moderate value as $u_{i_0j_0}^{(3)} = 1$. From

$$r_{ij}^{(3)} = \frac{u_{i_0j_0}^{(3)}}{u_{i_0j_0}^{(3)} + |u_{ij}^{(3)} - u_{i_0j_0}^{(3)}|} = \frac{1}{1 + |u_{ij}^{(3)} - 1|}$$

it follows that we can obtain the matrix of uniform effect measures of objective 3 as follows.

$$R^{(3)} = \left[r_{ij}^{(3)}\right]_{3\times 3} = \begin{bmatrix} 0.588 & 0.556 & 1 \\ 0.769 & 1 & 0.714 \\ 0.909 & 0.714 & 0.833 \end{bmatrix}.$$

Let the decision weights of objectives 1, 2, and 3 be, respectively,

$$\eta_1 = 0.5, \eta_2 = 0.35, \eta_3 = 0.15$$

From

$$r_{ij} = \sum_{k=1}^{3} r_{ij}^{(k)} \cdot \eta_k,$$

we can obtain the matrix of synthetic effect measures as follows:

$$R = [r_{ij}]_{3\times 3} = \begin{bmatrix} 0.144 & 0.595 & 0.247 \\ 0.194 & 0.266 & 0.428 \\ 0.512 & 0.264 & 0.241 \end{bmatrix}.$$

For $i = 1$,

$$\max_{1\le j\le 3} \{r_{1j}\} = 0.595 = r_{12}$$

the corresponding situation is

$$s_{12} = (a_1, b_2) = (\text{district 1, livestock husbandry}).$$

That is, district 1 should put its emphasis on the development of livestock.
For $i = 2$,

$$\max_{1\le j\le 3} \{r_{2j}\} = 0.428 = r_{23}$$

the corresponding situation is

$$s_{23} = (a_2, b_3) = (\text{district 2, industry and services}).$$

That is, district 2 should direct its attention to the development of industry and service-related businesses.
For $i = 3$,

$$\max_{1\le j\le 3} \{r_{3j}\} = 0.512 = r_{31}$$

the corresponding situation is

$$s_{31} = (a_3, b_1) = (\text{district 3, forestry}).$$

That is, district 3 should put its eyes on the development of forestry.
Furthermore, from

$$\max_{1\le i\le 3} \max_{1\le j\le 3} \{r_{ij}\} = 0.595 = r_{12}$$

it follows that

$$s_{12} = (a_1, b_2) = (\text{district 1, livestock husbandry})$$

is the optimum situation. That is, among the three business emphases, the county as a whole, consisting of all three economic districts, should first support the development of livestock husbandry in district 1.

10.8 Grey Stratified Decisions

We conclude this chapter by looking at a theory of stratified decision making, where each optimum decision is made by synthesizing decisions of several layers of decision makers.

Definition 10.8.1. *A group of decision makers with close and similar decision intentions is called a decision layer*.

Generally, the group of all parties involved in a decision-making process is divided into the layer of mass, the layer of experts, and the layer of administrators.These layers are denoted as L_1, L_2, and L_3, respectively.

Definition 10.8.2. *1. Each group of people with relatively small range of responsibility, scattered information, and a great number of individuals, is called the layer of mass.*

2. Each medium-sized group of people with responsibility for technical issues and relatively more decision information is called a layer of experts. And

3. The group of a few people with legislative responsibility and relatively uniform information on decision criteria is called the layer of administrators.

Definition 10.8.3. *The process of decision making, in which the optimum decision is achieved through synthesizing the decision intentions of different layers, is called a grey stratified decision making. According to Definition 10.8.2, the group of all parties involved in the decision making is divided into three layers L_1, L_2, and L_3. This kind of decision making is called a three-layered grey stratified decision making.*

In a three-layered grey stratified decision making, grey statistics is often used to reflect the decision intention of the layer of mass. So, L_1 is also called the grey statistics layer. Because experts tend to look at problems from the angle of future development, their decision intention is often reflected by using GM(1, 1) predicted values or GM(1, 1) development coefficients. So, L_2 is also called the grey development layer. The layer of administrators needs to balance the current situations and to consider all the legislative restrictions. So, for this layer, grey cluster is the method for its decision making. Therefore, L_3 is also called the grey cluster layer.

Each three-layered grey stratified decision making can be performed by going through the following steps.

Step 1: Order all situations in the set of situations

$$S = \{s_{ij} = (a_i, b_j)|a_i \in A, b_j \in B\}$$

in the dictionary order. If

$$A = \{a_1, a_2, ..., a_n\}$$

and

$$B = \{b_1, b_2, ..., b_m\}$$

then

$$S = \{s_q | q = 1, 2, ..., nm\}.$$

Step 2: Based the method as described in Section 10.5, compute the statistical decision coefficient σ_q^k for situation s_q to be in grey class $k, q = 1, 2, ..., nm$, $k = 1, 2, ..., s$, so that we have the vector of statistical decision coefficients

$$\sigma_q = \left(\sigma_q^1, \sigma_q^2, ..., \sigma_q^s\right),$$

$q = 1, 2, ..., nm$.

Step 3: Uniformly treat the observational effect time sequences of situation s_q with different objectives (for more details, see Section 10.7). Calculate the time sequence of synthetic effect measures

$$r_q = (r_q(1), r_q(2), ..., r_q(n_0))$$

and the development coefficient $-a_q$ of the relevant GM(1, 1) model using least squares estimates, q = 1, 2, ..., nm.

Step 4: Based on the method as described in Section 10.6 (here situations s_q should be seen as decision objects and the objectives in Step 3 as decision criteria), and a similar division of grey classes as in Step 2, compute grey cluster decision coefficients δ_q^k for situation s_q to be in grey class k so that the vector of grey cluster decision coefficients

$$\delta_q = \left(\delta_q^1, \delta_q^2, ..., \delta_q^s\right),$$

$q = 1, 2, ..., nm$, is obtained.

Step 5: Calculate the united decision vector of L_1 and L_2. Let

$$\begin{aligned}\omega &= \left(\omega^1, \omega^2, ..., \omega^s\right) \\ &= \begin{bmatrix} -a_1 \\ -a_2 \\ \cdots \\ -a_{nm} \end{bmatrix}^T \cdot \begin{bmatrix} \sigma_1^1 & \sigma_1^2 & \cdots & \sigma_1^s \\ \sigma_2^1 & \sigma_2^2 & \cdots & \sigma_2^s \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{nm}^1 & \sigma_{nm}^2 & \cdots & \sigma_{nm}^s \end{bmatrix}.\end{aligned}$$

Step 6: Compute the united decision of L_1, L_2, and L_3. For $q = 1, 2, ..., nm$, calculate the degree $\gamma(\omega, \delta_q)$ of grey incidence of δ_q and ω. Assume that

$$\max\{\gamma(\omega, \delta_q)\} = \gamma(\omega, \delta_{q^*}).$$

Then s_{q^*} is the optimum situation of the three-layered grey stratified decision making.

Example 10.8.1. Let us look at a three-layered grey stratified decision making done for a proposed technical innovation of a certain industrial company.

Step 1: The situation set S is the same as that in Example 10.4.1. That is,

$$S = \{s_{11}, s_{12}, s_{13}\}.$$

When ordered in a dictionary order, we have

$$S = \{s_1, s_2, s_3\},$$

where $n = 1, m = 3$, so, $nm = 3$.

Step 2: Assume that the grey statistics layer consists of five workshops, whose decision weights, based on the numbers of people involved, are given, respectively:

$$\eta_1 = 0.25, \eta_2 = 0.25, \eta_3 = 0.2, \eta_4 = 0.2, \eta_5 = 0.1.$$

And, the matrix of quantified evaluations of each workshop about the situation s_q, $q = 1, 2, 3$, is given by

$$C = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{bmatrix} = \begin{bmatrix} 80 & 60 & 40 \\ 60 & 50 & 50 \\ 75 & 70 & 60 \\ 90 & 80 & 80 \\ 50 & 70 & 60 \end{bmatrix}.$$

The whitenization weight functions for the four evaluation grey classes, best, implementable, basically implementable, and not implementable, are given, respectively (for more details, see Figure 10.1):

$$f^1(x) = \begin{cases} 0, & x < 60 \\ \dfrac{x-60}{30}, & 60 \le x \le 90 \\ 1, & x > 90; \end{cases}$$

$$f^2(x) = \begin{cases} 0, & x < 50 \\ \dfrac{x-50}{20}, & 50 \le x \le 70 \\ \dfrac{90-x}{20}, & 70 < x \le 90 \\ 0, & x > 90; \end{cases}$$

$$f^3(x) = \begin{cases} 0, & x < 35 \\ \dfrac{x-35}{20}, & 35 \le x \le 55 \\ \dfrac{75-x}{20}, & 55 < x \le 75 \\ 0, & x > 75; \end{cases}$$

and

$$f^4(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \le x \le 20 \\ \dfrac{40-x}{20}, & 20 < x \le 40 \\ 0, & x > 40. \end{cases}$$

From

$$\sigma_q^k = \frac{\sum_{i=1}^{5} f^k(x_{iq}) \cdot \eta_i}{\sum_{k=1}^{4}\sum_{i=1}^{5} f^k(x_{iq}) \cdot \eta_i},$$

$q = 1, 2, 3,\ k = 1, 2, 3, 4$, it follows that

$$\sigma_1 = \left(\sigma_1^1, \sigma_1^2, \sigma_1^3, \sigma_1^4\right) = (0.42, 0.35, 0.23, 0)\,,$$

$$\sigma_2 = \left(\sigma_2^1, \sigma_2^2, \sigma_2^3, \sigma_2^4\right) = (0.19, 0.44, 0.37, 0)\,,$$

and

$$\sigma_3 = \left(\sigma_3^1, \sigma_3^2, \sigma_3^3, \sigma_3^4\right) = (0.15, 0.29, 0.56, 0)\,.$$

Step 3: Assume that we have four objectives: production, profit, innovation investment, and production cost, and that the time sequences of synthetic effect measures of situation s_q, $q = 1, 2, 3$, with respect to the previous four objectives are, respectively:

$$r_1 = (r_1(i))_{i=1}^4 = (0.32, 0.44, 0.58, 0.70),$$

$$r_2 = (r_2(i))_{i=1}^4 = (0.23, 0.39, 0.69, 0.83),$$

and

$$r_3 = (r_3(i))_{i=1}^4 = (0.12, 0.14, 0.81, 0.96).$$

The least squares estimates of the GM(1, 1) parameter sequences of r_q, $q = 1, 2, 3$, are given by

$$\widehat{a}_1 = \begin{bmatrix} a_1 & b_1 \end{bmatrix}^T = \begin{bmatrix} -0.23 & 0.3215 \end{bmatrix}^T,$$

$$\widehat{a}_2 = \begin{bmatrix} a_2 & b_2 \end{bmatrix}^T = \begin{bmatrix} -0.32 & 0.2987 \end{bmatrix}^T,$$

and

$$\widehat{a}_3 = \begin{bmatrix} a_3 & b_3 \end{bmatrix}^T = \begin{bmatrix} -0.55 & 0.1974 \end{bmatrix}^T.$$

with the relevant development coefficients being

$$-a_1 = 0.23, -a_2 = 0.32, -a_3 = 0.55.$$

Step 4: Determine the decision objects of the situations s_1, s_2, and s_3 with the objectives listed in Step 3: production, profit, innovation investment, and production cost, as decision criteria here.

Assume that the matrix of quantified evaluation values of the situations s_1, s_2, and s_3 with respect to the previous four objectives is

$$D = [y_{qj}]_{3\times 4} = \begin{bmatrix} 420 & 51 & 26 & 0.67 \\ 874 & 93 & 30 & 0.61 \\ 1035 & 124 & 38 & 0.58 \end{bmatrix}.$$

Now, the whitenization weight function $f_j^k(\cdot)$, $j = 1, 2, 3, 4$, $k = 1, 2, 3, 4$, for j criterion to be in grey class k is shown in Figures 10.3 $\sim$ 10.6.

From these four figures, it is not hard to see the piecewise definitions of the whitenization weight functions $f_j^k(\cdot)$, $j = 1, 2, 3, 4$, $k = 1, 2, 3, 4$, for criterion j to be in grey class k. Assume again that the decision weights of the four criteria: production, profit, innovation investment, and production cost, are

$$\zeta_1 = 0.3, \zeta_2 = 0.3, \zeta_3 = 0.2, \zeta_4 = 0.2.$$

From

$$\delta_q^k = \sum_{j=1}^4 f_j^k(y_{qj}) \cdot \zeta_j,$$

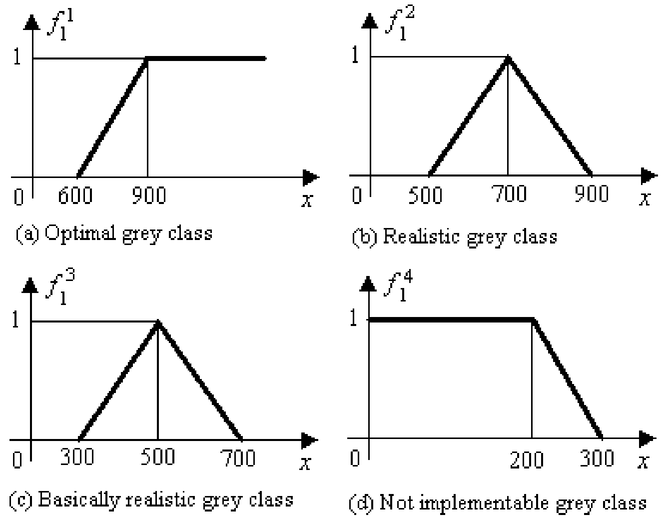

FIGURE 10.3. The whitenization weight functions with respect to the criterion: production

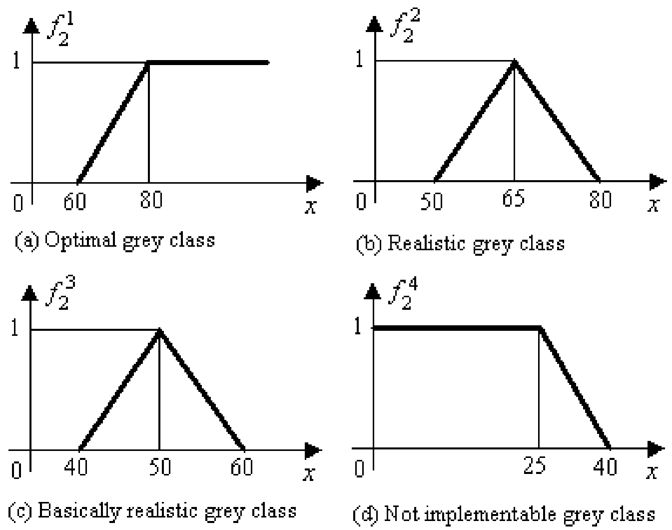

FIGURE 10.4. The whitenization weight functions with respect to the criterion: profit

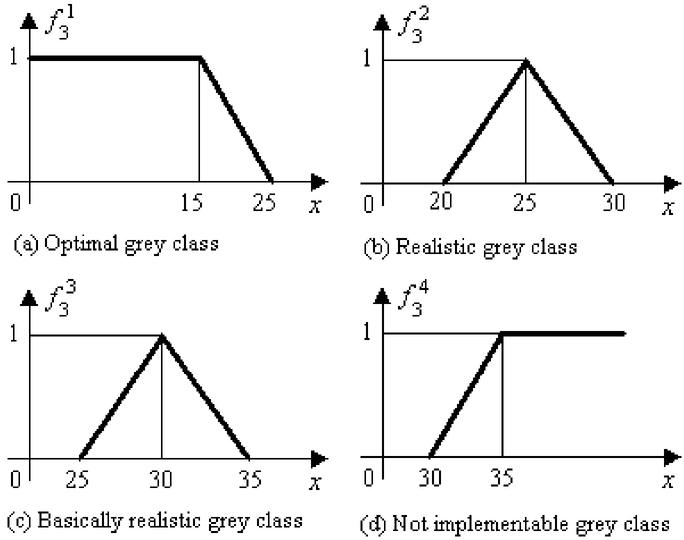

FIGURE 10.5. The whitenization weight functions with respect to the criterion: innovation investment

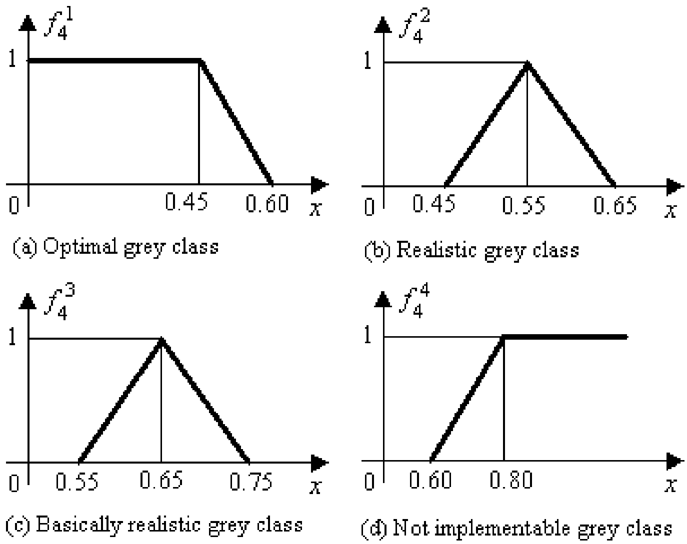

FIGURE 10.6. The whitenization weight functions with respect to the criterion: production cost

$q = 1, 2, 3$, $k = 1, 2, 3, 4$, it follows that the grey cluster decision coefficients are

$$\delta_1 = \left(\delta_1^1, \delta_1^2, \delta_1^3, \delta_1^4\right) \quad = (0, 0.18, 0.65, 0.07)\,,$$

$$\delta_2 = \left(\delta_2^1, \delta_2^2, \delta_2^3, \delta_2^4\right) \quad = (0.57, 0.12, 0.32, 0.01)\,,$$

and

$$\delta_3 = \left(\delta_3^1, \delta_3^2, \delta_3^3, \delta_3^4\right) \quad = (0.63, 0.14, 0.06, 0.2)\,.$$

Step 5: Compute the united decision vector of the layers L_1 and L_2 of the mass and experts. From

$$\begin{bmatrix} -a_1 \\ -a_2 \\ -a_3 \end{bmatrix}^T \cdot \begin{bmatrix} \sigma_1^1 & \sigma_1^2 & \sigma_1^3 & \sigma_1^4 \\ \sigma_2^1 & \sigma_2^2 & \sigma_2^3 & \sigma_2^4 \\ \sigma_3^1 & \sigma_3^2 & \sigma_3^3 & \sigma_3^4 \end{bmatrix}$$

$$= \begin{bmatrix} 0.23 \\ 0.32 \\ 0.55 \end{bmatrix}^T \cdot \begin{bmatrix} 0.42 & 0.35 & 0.23 & 0 \\ 0.19 & 0.44 & 0.37 & 0 \\ 0.15 & 0.29 & 0.56 & 0 \end{bmatrix}$$

$$= \quad (0.2399, 0.3808, 0.4793, 0)\,,$$

it follows that

$$\omega \quad = \left(\omega^1, \omega^2, \omega^3, \omega^4\right) = (0.2399, 0.3808, 0.4793, 0)\,.$$

Step 6: Determine the united decision of the three layers L_1, L_2, and L_3. Calculate the degree of grey incidence between δ_q and ω, $q = 1, 2, 3$. We obtain that

$$\gamma(\omega, \delta_1) = 0.59672,$$

$$\gamma(\omega, \delta_2) = 0.61725,$$

and

$$\gamma(\omega, \delta_3) = 0.43475.$$

From

$$\max\left\{\gamma(\omega, \delta_q)\right\} = 0.61725 = \gamma(\omega, \delta_2),$$

it follows that s_2 is the optimum situation of our three-layered grey stratified decision making. That is, an innovation done by stages is recommended. This is a conclusion obtained after synthesizing the opinions of the three layers L_1, L_2, and L_3. If it is seen from the angle of a certain layer, then s_2 is not an optimum situation. For example, from the decision vector of the layer of mass, it can be seen that s_2 is an implementable plan. From the decision coefficient of the layer of experts, it can be seen that the development coefficient of s_2 is located in the middle, which implies that it is not the optimum plan. And in the eyes of the layer of administrators, the plan of innovation by stages is optimum.

11
Grey Programming

11.1 Introduction

The so-called programming essentially belongs to the category of decision making. It mainly studies under certain constraints how to guarantee the objective of achieving the possible optimum. In summary, problems studied in programming are mainly the following kinds.

1. Problems of production plans: Under the condition of limited resources, determine the products and relevant quantities so that the output values and profits would be the maximum.

2. Problems of management of scientific research: Determine how to allocate the limited amount of available funds among various research disciplines and projects and who should be responsible for the research of what projects so that the productivity and profits of the research could be the highest.

3. Problems of military commands: Determine how to arrange the limited size of armed forces to effectively defeat enemies.

4. Problems of agricultural division into districts: Make decisions on how to choose different economic patterns, based on different conditions of soil, weather, and resources, to determine the sizes of areas of planting of various crops so that the total effect would be maximized.

5. Problems of industrial distribution and urban planning: Determine how to distribute industries and how to develop cities so that the overall economy would benefit.

6. Problems of transportation: Determine how to transport goods and materials from areas of supply to areas of demands in the network of allocation of goods and materials so that the transportation cost would be the least while all demands are met.

7. Problems of stocks: Determine the kinds and quantities of products, and time periods of storage so that under the condition of limited available spaces, the stocking profits would be maximum.

8. Problems of compounding materials: Determine how much each material is needed, under the condition that technological processes, qualities, etc., are pre-fixed, so that the cost would be minimum.

9. Problems of material-cutting: Determine how to cut materials into usable pieces so that the rate of usage of the materials would be maximum or that the number of complete sets is maximum.

10. Other problems: For example, how to allocate investment in advertising and determine methods of advertising so that an optimal effect can be reached. Under certain requirements, how to determine the number of people working so that the minimum number of employees are hired, etc.

In these problems, if the constraint condition and objective function is linear, the problems are called linear programming problems. When the objective function or the constraint condition is a nonlinear function, the corresponding problem is called a nonlinear programming problem. If we are solving a problem with a possible answer as "Do it" or "Don't do it," or in the constraint, there appears a phrase like "Either, or", we denote "Do it" or "Either this" as 1 and "Don't do it" or "Or that" as 0. In this way, the problem we are solving involves two variables 0 and 1 only, which is consequently called a 0-1 programming problem. Here, linear programming is one of the most important branches in operations research, which developed early and matured fast with a wide range of practical applications. However, the normal linear programming, nonlinear programming, and 0-1 programming all have the following problems.

1. They are all "static" programming, which cannot be used to reflect the situation of change when the constraints are changing with time;

2. When grey numbers appear in either the programming model or the constraint, applications will become difficult; and

3. In theory, each convex function defined on a convex set has a solution. However, in practical applications, due to technical reasons, the process of finding the solution cannot be finished.

By using the idea and modeling method of grey systems theory, these problems with programmings as listed above can be resolved to a certain degree. In this chapter, we mainly study grey linear programming, grey 0-1 programming, and grey nonlinear programming.

11.2 Linear Programming Models with Grey Parameters

In this section, we first look at linear programming models.

Definition 11.2.1. *Assume that* a_{ij}, b_i, *and* c_j, i *=1,2, ...,* m, $j = 1, 2$, ..., n, *are all constants, and* x_j, $j = 1, 2$, ..., n, *are unknown quantities. Then*

$$\max(\min) S = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

$$s.t. \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq (=, \geq) b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq (=, \geq) b_2 \\ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdot\cdot \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq (=, \geq) b_m \\ x_1 \geq 0, \ x_2 \geq 0, \ \cdot\cdot\cdot, \ x_n \geq 0 \end{cases}$$

is called a general mathematical model of linear programming problems, where

$$S = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

is called an objective function, and

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq (=, \geq) b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq (=, \geq) b_2 \\ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ \cdot\cdot \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq (=, \geq) b_m \\ x_1 \geq 0, \ x_2 \geq 0, \ \cdot\cdot\cdot, \ x_n \geq 0 \end{cases}$$

the constraint conditions.

Definition 11.2.2. *The following*

$$\max S = CX$$

$$s.t. \begin{cases} AX = b \\ X \geq 0 \end{cases}$$

is called the standardized type of linear programming problems, where

$$C = [c_1, c_2, ..., c_n],$$

$$X = [x_1, x_2, ..., x_n]^T,$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

$$b = [b_1, b_2, ..., b_m]^T$$

and $b_i \geq 0$, $i = 1, 2, ..., m$.

Definition 11.2.3. *Assume that*

$$X = [x_1, x_2, ..., x_n]^T,$$

$$C = [c_1(\otimes), c_2(\otimes), ..., c_n(\otimes)],$$

$$b = [b_1(\otimes), b_2(\otimes), ..., b_m(\otimes)]^T,$$

$$A(\otimes) = \begin{bmatrix} a_{11}(\otimes) & a_{12}(\otimes) \cdots & a_{1n}(\otimes) \\ a_{21}(\otimes) & a_{22}(\otimes) \cdots & a_{2n}(\otimes) \\ \cdots & \cdots & \cdots \\ a_{m1}(\otimes) & a_{m2}(\otimes) \cdots & a_{mn}(\otimes) \end{bmatrix},$$

where

$$c_j(\otimes) \in [\underline{c}_j, \bar{c}_j], \ \underline{c}_j \geq 0,$$

$j = 1, 2, ..., n,$

$$b_i(\otimes) \in [\underline{b}_i, \bar{b}_i], \underline{b}_i \geq 0,$$

$i = 1, 2, ..., m,$ and

$$a_{ij}(\otimes) \in [\underline{a}_{ij}, \bar{a}_{ij}], \underline{a}_{ij} \geq 0,$$

$i = 1, 2, ..., m, j = 1, 2, ..., n.$

Then

$$\max S = C(\otimes)X$$

$$s.t. \begin{cases} A(\otimes)X \leq b(\otimes) \\ X \geq 0 \end{cases}$$

is called a problem of linear programming with grey parameters $(LPGP)$, *and* $C(\otimes)$ *a grey price vector,* $A(\otimes)$ *a grey consumption matrix,* $b(\otimes)$ *a grey constraints vector for resource, and* X *the decision vector of the* $LPGP$.

As a matter of fact, X is a grey vector as well.

Definition 11.2.4. *Suppose that* α_j, β_i, $\gamma_{ij} \in [\,0,1]$, $i = 1, 2, ..., m$, $j = 1, 2, ..., n$, *and let the white values of grey parameters be, respectively, as follows*

$$\widetilde{c}_j(\otimes) = \alpha_j \overline{c}_j + (1-\alpha_j)\underline{c}_j,$$

$j = 1, 2, ..., n$,

$$\widetilde{b}_i(\otimes) = \beta_i \overline{b}_i + (1-\beta_i)\underline{b}_i,$$

$i = 1, 2, ..., m$, and

$$\widetilde{a}_{ij}(\otimes) = \gamma_{ij}\overline{a}_{ij} + (1-\gamma_{ij})\underline{a}_{ij},$$

$i = 1, 2, ..., m$, $j = 1, 2, ..., n$, *where* $\widetilde{C}(\otimes), \widetilde{b}(\otimes)$ *and* $\widetilde{A}(\otimes)$ *are, respectively, the whitenization vector of price, constraints for resources, and the whitenization matrix of consumption. Then*

$$\max S = \widetilde{C}(\otimes)X$$

$$s.t. \begin{cases} \widetilde{A}(\otimes)X \leq \widetilde{b}(\otimes) \\ X \geq 0 \end{cases}$$

is called a positioned programming of the $LPGP$*; and* α_j $(j = 1, 2, ..., n)$ *the positioned coefficients of price vector,* β_i $(i = 1, 2, ..., m)$ *the positioned coefficients of constraint vector for resources, and* γ_{ij} $(i = 1, 2, ..., m, j = 1, 2, ..., n)$ *the positioned coefficients of consumption.*

In Definition 11.2.4, α_j is a reflection of price fluctuation of the jth product. It can be determined by using a market analysis. Less α_j reflects a lower expected price of the jth product, and larger α_j reflects a higher expected price of the jth product.

The coefficient β_i is a reflection of market supplies of the ith resource. Less β_i expresses short supply of the ith resource, and larger β_i expresses sufficient supply of the ith resource.

Similarly, less γ_{ij} expresses lower consumption of the ith resource to produce an unit of the jth product, and larger γ_{ij} expresses higher consumption of the ith resource to produce the same unit of the jth product.

Proposition 11.2.1. *The optimal value max$S$ of the positioned programming of a LPGP is a function with $n+m+mn$ variables of α_j, β_i and γ_{ij} ($i = 1, 2, ..., m$, $j = 1, 2, ..., n$).*

Therefore from Proposition 11.2.1, the optimal value max$S$ of the positioned programming can be marked as follows,

$$\max S = f\left((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m,\ j = 1, 2, ..., n\right).$$

Similarly, the positioned programming can be marked as follows,

$$LP\left((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m,\ j = 1, 2, ..., n\right).$$

For the sake of convenience, we first make the following suppositions.

1. Rank$(\widetilde{A}(\otimes)) = m < n$. *Here* Rank$(\widetilde{A}(\otimes))$ *refers to the rank of matrix* $\widetilde{A}(\otimes)$, *and suppose* Rank$(\widetilde{A}(\otimes)) = m$.

2. The set composed of the feasible solution of $LP((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m, j = 1, 2, ..., n)$ is non-empty.

3. The set

$$\left\{X|\widetilde{A}(\otimes)X \leq \widetilde{b}(\otimes),\ X \geq 0\right\}$$

composed of real vectors is bounded. At the same time, the positioned programming $LP((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m, j = 1, 2, ..., n)$ can be rewritten into the following form,

$$\max S = \left[\widetilde{C}_B(\otimes), \widetilde{C}_N(\otimes)\right]\begin{bmatrix} X_B \\ X_N \end{bmatrix}$$

$$s.t. \begin{cases} \left[\widetilde{B}(\otimes), \widetilde{N}(\otimes)\right]\begin{bmatrix} X_B \\ X_N \end{bmatrix} \leq \widetilde{b}(\otimes) \\ \\ X_B \geq 0, X_N \geq 0. \end{cases}$$

That is, the first m columns of the whitenization matrix of consumer $\widetilde{A}(\otimes)$ are the basis matrix $\widetilde{B}(\otimes)$; the last $n-m$ columns are the non-basis matrix $\widetilde{N}(\otimes)$. The basis vectors and non-basis vectors corresponding to $\widetilde{B}(\otimes)$ and $\widetilde{N}(\otimes)$ can be written, respectively, as X_B and X_N. The whitenization vectors of price corresponding to X_B and X_N can be written, respectively, as $\widetilde{C}_B(\otimes)$ and $\widetilde{C}_N(\otimes)$. From supposition 3, and noticing the fact that $X_N = 0$, it is clear that

$$X = [X_B, X_N]^T = \left[\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes), 0\right]^T$$

and

$$S = \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes),$$

and the test vector is

$$r = \widetilde{C}(\otimes) - \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{A}(\otimes).$$

Proposition 11.2.2. *Suppose that the positioned programming in* 3 *satisfies the above suppositions 1,2,and 3, and*

$$X = [x_1,\ x_2,\ ...,\ x_n]^T$$

is the basic solution of the positioned programming in 3. Then,

$$\{x_j|\ j = 1,\ 2,\ ..., n\}$$

is bounded.

Proposition 11.2.3. *There is at least one basic feasible solution of the positioned programming LP* $((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1,\ 2,\ ...,\ m,\ j = 1,\ 2,\ ...,\ n)$*, which satisfies the suppositions* $1, 2$*, and* 3.

11.3 Grey Linear Programming of Prediction Type

In this section, we focus on linear programming problems such that they can be potentially used to make predictions.

Definition 11.3.1. *For the grey linear programming problem in Definition 11.2.3, first whitenize* $C(\otimes)$*, and* $A(\otimes)$*. Assume that*

$$\widetilde{C} = [\widetilde{c}_1, \widetilde{c}_2, ..., \widetilde{c}_n]$$

and

$$\widetilde{A} = \begin{bmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} & \cdots & \widetilde{a}_{1n} \\ \widetilde{a}_{21} & \widetilde{a}_{22} & \cdots & \widetilde{a}_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \widetilde{a}_{m1} & \widetilde{a}_{m2} & \cdots & \widetilde{a}_{mn} \end{bmatrix}.$$

Based on the historical data of $b_i(\otimes)$*,* $i = 1, 2,\ ...,\ m$*,*

$$b_i(\otimes) = (b_i(1), b_i(2), ..., b_i(s)),$$

establish a GM(1, 1) model and solve for its predicted value $\widehat{b}_i(s+k)$ *at the time moment* $s+k$*, for* $i = 1, 2,\ ..., m$*. Denote*

$$\widehat{b} = \left[\widehat{b}_1(s+k), \widehat{b}_2(s+k), ..., \widehat{b}_m(s+k)\right].$$

Then

$$\max S = \widetilde{C}X$$

$$s.t. \begin{cases} \widetilde{A}X = \widehat{b} \\ X \geq 0 \end{cases}$$

is called a linear programming problem of grey prediction type.

For a linear programming problem of grey prediction type, we can solve it according to the method of solving a general linear programming problem.

Example 11.3.1. A manufacturing company produces two products: Product A and Product B. Each piece of Product A requires $2.5 \sim 3.5$ work days, $3 \sim 5$ kilowatt-hours of electricity, and $7 \sim 11$ tons of coals. And each piece of Product B requires $8 \sim 12$ work days, $3.5 \sim 6.5$ kilowatt-hours of electricity, and $3 \sim 5$ tons of coals. The profit from each piece of Product A is \$600 $\sim$ 800, and the profit from each piece of Product B is \$900 $\sim$ 1500. This company has 300 movable laborers, 360 tons of daily consumable coals, and the daily electricity supplies are given as in Table 11.1.

Table 11.1. Electricity supply during 1993 to 1996

Year	1993	1994	1995	1996
Daily electricity supply (in kilowhour)	168	174	180	190

For the years of 1997 and 1998, how should the company arrange its daily production of Products A and B to maximize its profits?

Solution: Assume that the production of Products A and B are x_1 and x_2, respectively. Then, we have the following grey linear programming problem.

$$\max S = c_1(\otimes)x_1 + c_2(\otimes)x_2$$

$$s.t. \begin{cases} \otimes_{11}x_1 + \otimes_{12}x_2 \leq b_1(\otimes) \\ \otimes_{21}x_1 + \otimes_{22}x_2 \leq b_2 = 360 \\ \otimes_{31}x_1 + \otimes_{32}x_2 \leq b_3 = 300 \\ x_1 \geq 0,\ x_2 \geq 0, \end{cases}$$

where

$$c_1(\otimes) \in [600, 800], c_2(\otimes) \in [900, 1500],$$

$$\otimes_{11} \in [3, 5], \otimes_{12} \in [3.5, 6.5],$$

$$\otimes_{21} \in [7, 11,], \otimes_{22} \in [3, 5],$$

and

$$\otimes_{31} \in [2.5, 3.5], \otimes_{32} \in [8, 12].$$

Mean whitenizing all the grey elements in the objective function and constraints gives us that

$$\widetilde{C} = [\widetilde{c}_1, \widetilde{c}_1] = [700, 1200]$$

and

$$\widetilde{A} = \begin{bmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} \\ \widetilde{a}_{21} & \widetilde{a}_{22} \\ \widetilde{a}_{31} & \widetilde{a}_{32} \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 9 & 4 \\ 3 & 10 \end{bmatrix}.$$

From Table 11.1, we obtain a sequence of restrained variables $b_1(\otimes)$ as follows,

$$b_1(\otimes) \quad = (b_1(i))_{i=1}^4 = (168, 174, 180, 190)\,.$$

Its $GM(1,1)$ time response sequence is

$$\begin{cases} \widehat{b}_1^{(1)}(k+1) & = 3829.125 \cdot e^{0.0442k} - 3661.125 \\ \widehat{b}_1(k+1) & = \widehat{b}_1^{(1)}(k+1) - \widehat{b}_1^{(1)}(k). \end{cases}$$

From this end, it follows that

$$\widehat{b}_1(5) \approx 198, \quad (1997),$$

$$\widehat{b}_1(6) \approx 207, \quad (1998).$$

Therefore, the programming model for the year of 1997 is

$$\max S = 700x_1 + 1200x_2$$

$$s.t. \begin{cases} 4x_1 + 5x_2 \le 198 \\ 9x_1 + 4x_2 \le 360 \\ 3x_1 + 10x_2 \le 300 \\ x_1 \ge 0,\ x_2 \ge 0. \end{cases}$$

This is a linear programming problem in two variables, which can be solved by using graphs. The feasible region of this problem is given in Figure 11.1.

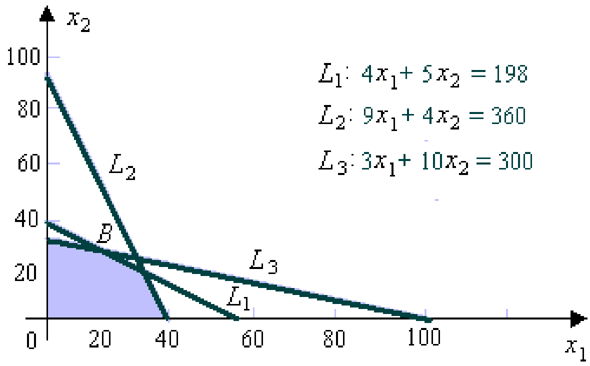

FIGURE 11.1. The feasible region of the linear programming model for 1997

The point B is the last point intersecting all lines with slope $-\frac{7}{12}$ parallel to the objective function $S = 700x_1 + 1200x_2$. So, point B is the optimal point. Solving the system

$$\begin{cases} 4x_1 + 5x_2 = 198 \\ 3x_1 + 10x_2 = 300 \end{cases}$$

gives the optimal solution

$$x_1 = 19.2, x_2 = 24.24.$$

That is, in 1997, in order to maximize profit for the company, the daily production of Products A and B should be 19.2 pieces and 24.24 pieces, respectively. The maximum daily profit would be

$$\max S = 700 \cdot 19.2 + 1200 \cdot 24.24 = 42528$$

with an annual profit \$15,522,720. This problem can also be solved by introducing slack variables to convert to the standard form and then applying the simplex algorithm.

The programming model for the year of 1998 is

$$\max S = 700x_1 + 1200x_2$$

$$s.t. \begin{cases} 4x_1 + 5x_2 \leq 207 \\ 9x_1 + 4x_2 \leq 360 \\ 3x_1 + 10x_2 \leq 300 \\ x_1 \geq 0,\ x_2 \geq 0. \end{cases}$$

Similarly, we can find the optimal solution as follows:

$$x_1 = 22.8, x_2 = 23.16$$

with the maximum daily profit

$$\max S = 700 \cdot 22.8 + 1200 \cdot 23.16 = 43752 \ (\$)$$

and maximum annual profit \$15,969,480.

Comparing the programming model of production for the year 1998 with that of the year 1997, the daily electricity supply is increased by 9 kilowatt-hours with profit increased by an amount of \$1,224, and there is a surplus supply of coal. For the year 1997, the daily usage of coal is

$$9 \cdot 19.2 + 4 \cdot 24.24 = 269.76 \ (\text{ton}),$$

which is 90.24 tons less than the scheduled supply. For the year of 1998, the daily consumption of coal is

$$9 \cdot 22.8 + 4 \cdot 23.16 = 297.84 \ (\text{ton}),$$

which is 62.16 tons less than the scheduled supply. If there is no plan to devote more manpower and electricity, the scheduled supply of coal can be reduced appropriately.

11.4 Several Theorems on Positioned Solutions of LPGP

An *LPGP* is also called grey drifting linear programming. In reality, a problem of *LPGP* is a set composed of some ordinary problems of linear programming.

In the following proof, we suppose that the whitenization vectors and the whitenization matrix, given in the following

$$\max S = \left[\widetilde{C}_B(\otimes), \widetilde{C}_N(\otimes)\right] \begin{bmatrix} X_B \\ X_N \end{bmatrix}$$

$$s.t. \begin{cases} \left[\widetilde{B}(\otimes), \widetilde{N}(\otimes)\right] \begin{bmatrix} X_B \\ X_N \end{bmatrix} \leq \widetilde{b}(\otimes) \\ X_B \geq 0, X_N \geq 0 \end{cases}$$

still keep the property of non-negativity.

Theorem 11.4.1. *For a positioned programming of a LPGP, when the positioned coefficients of the price vector satisfy*

$$\alpha_j \leq \alpha_j^{'} \ ,$$

$j = 1, 2, ..., n$, *we have*

$$\begin{aligned} \max S &= f\left((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m;\ j = 1, 2, ..., n\right) \\ &\leq f\left((\alpha_j^{'}, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m;\ j = 1, 2, ..., n\right) = \max S^{'}. \end{aligned}$$

Proof: Because $\alpha_j \leq \alpha_j^{'}$, we have

$$\widetilde{C}(\otimes) \leq \widetilde{C}^{'}(\otimes).$$

Suppose that

$$\widetilde{C}^{'}(\otimes) = \widetilde{C}(\otimes) + \Delta\widetilde{C}(\otimes), \text{ and } \Delta\widetilde{C}(\otimes) \geq 0.$$

There are now the following two cases. Without loss of generality,we assume that $\widetilde{B}(\otimes)$ is the optimal basis of $LP((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2,\ ..., n)$.

1.

$$\widetilde{C}^{'}(\otimes) - \widetilde{C}_B^{'}(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{A}(\otimes) \leq 0.$$

Here, the optimal basis $\widetilde{B}(\otimes)$ of the corresponding positioned programming doesn't change, nor does the optimum solution

$$X = \left[\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes), 0\right]^T.$$

Obviously,

$$\begin{aligned} \max S^{'} &= \widetilde{C}_B^{'}(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \\ &= \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \\ &\quad +\Delta\widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \geq \max S \end{aligned}$$

2.

$$\widetilde{C}^{'}(\otimes) - \widetilde{C}_B^{'}(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{A}(\otimes) > 0.$$

Suppose that the test number $r_k^{'}(\otimes) > 0$, and that $\widetilde{B}(\otimes)$ is not the optimal basis of the positioned programming $LP((\alpha_j^{'}, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m,$ $j = 1, 2,\ ...,\ n)$. Moreover, suppose that we use the simplex method to work out its optimal basis $\widetilde{B}_1(\otimes)$ and its optimal solution

$$\left[\widetilde{B}_1^{-1}(\otimes)\widetilde{b}(\otimes), 0\right]^T.$$

Notice that

$$\left[\widetilde{B}_1^{-1}(\otimes)\widetilde{b}(\otimes), 0\right]^T$$

is the feasible basic solution of the positioned programming $LP((\alpha_j^{'}, \beta_i, \gamma_{ij})|$ $i = 1, 2, ..., m, j = 1, 2, ..., n)$; it is readily to have

$$\begin{aligned} \max S^{'} &= \widetilde{C}_{B_1}^{'}(\otimes)\widetilde{B}_1^{-1}(\otimes)\widetilde{b}(\otimes) \geq \widetilde{C}_B^{'}(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \\ &= [\widetilde{C}_B(\otimes) + \Delta\widetilde{C}_B(\otimes)]\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \\ &= \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) + \Delta\widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \geq \max S. \ \square \end{aligned}$$

Theorem 11.4.2. *For a positioned programming of a LPGP, when the positioned coefficients of restriction vectors for resource satisfy the following*

$$\beta_i \leq \beta_i^{'},$$

$i = 1, 2, ..., m$, *we have*

$$\begin{aligned} \max S &= f\left((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m;\ j = 1, 2, ..., n\right) \\ &\leq f\left((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2, ..., m;\ j = 1, 2, ..., n\right) = \max S^{'}. \end{aligned}$$

Proof: From

$$\beta_i \leq \beta_i^{'},$$

$i = 1, 2, ..., m$, we know that $\widetilde{b}(\otimes) \leq \widetilde{b}^{'}(\otimes)$.

Suppose that

$$\widetilde{b}^{'}(\otimes) = \widetilde{b}(\otimes) + \Delta\widetilde{b}(\otimes),\ \ \Delta\widetilde{b}(\otimes) \geq 0,$$

then we have

$$\widetilde{B}^{-1}(\otimes)\widetilde{b}^{'}(\otimes) = \widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) + \widetilde{B}^{-1}(\otimes)\Delta\widetilde{b}(\otimes).$$

Here, $\widetilde{B}(\otimes)$ is the optimal basis of $LP((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m, j = 1, 2, ..., n)$.

1. Assume

$$\widetilde{B}^{-1}(\otimes)\Delta\widetilde{b}(\otimes) \geq 0.$$

Then

$$\begin{aligned} &\widetilde{B}^{-1}(\otimes)\widetilde{b}^{'}(\otimes) \\ &= \widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) + \widetilde{B}^{-1}(\otimes)\Delta\widetilde{b}(\otimes) \geq 0. \end{aligned}$$

Hence, $\widetilde{B}(\otimes)$ is still the optimal basis of the positioned programming $LP((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2, ..., m, j = 1, 2, ..., n)$. So

$$\begin{aligned} \max S^{'} &= \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}^{'}(\otimes) \\ &= \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) + \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\Delta\widetilde{b}(\otimes) \\ &= \max S + \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\Delta\widetilde{b}(\otimes) \geq \max S \end{aligned}$$

2. Assume

$$\widetilde{B}^{-1}(\otimes)\Delta\widetilde{b}(\otimes) < 0.$$

Suppose that $\exists k$, $\Delta x_k < 0$. Now we discuss the situation in two cases as follows.

(a) $x_k^{'} = x_k + \Delta x_k \geq 0$. Here, $\widetilde{B}(\otimes)$ is still the optimal basis of the positioned programming $LP((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2,\ ..., n)$. So, the optimal solution of $LP((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2,\ ..., n)$ is a feasible basic solution of $LP((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2,\ ..., n)$. Therefore, we have

$$\max S = \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \leq \max S^{'}.$$

(b) $x_k^{'} = x_k + \Delta x_k < 0$. Now

$$[\widetilde{B}^{-1}(\otimes)\widetilde{b}^{'}(\otimes), 0]^T$$

is not the feasible basic solution of $LP((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2,\ ...,\ m,\ j = 1, 2,\ ...,\ n)$. But $\widetilde{B}(\otimes)$ is a regular basis. By using the dual simplex method, we can obtain the optimal solution

$$X^{'} = [\widetilde{B}_1^{-1}(\otimes)\widetilde{b}^{'}(\otimes), 0]^T$$

and the optimal basis $\widetilde{B}_1(\otimes)$ of $LP((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2,\ ..., n)$. And, noticing that the optimal solution of $LP((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2,\ ...,\ n)$ is a feasible basic solution of $LP((\alpha_j, \beta_i^{'}, \gamma_{ij})|\ i = 1, 2,\ ..., m,\ j = 1, 2, ..., n)$, we have

$$\max S = \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \leq \max S^{'}. \ \square$$

Theorem 11.4.3. *For a positioned programming* $LP(\ (\alpha_j, \beta_i, \gamma_{ij})|\ i = 1,\ 2,\ ..., m,\ j = 1, 2,\ ..., n)$ *of a LPGP, when the positioned coefficients of consumption satisfy the following*

$$\gamma_{ij} \geq \gamma_{ij}^{'},$$

$i = 1, 2, ..., m,\ j = 1, 2, ..., n,$ *we have*

$$\begin{aligned} \max S &= f\left((\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m;\ j = 1, 2, ..., n\right) \\ &\leq f\left((\alpha_j, \beta_i, \gamma_{ij}^{'})|\ i = 1, 2, ..., m;\ j = 1, 2, ..., n\right) = \max S^{'}. \end{aligned}$$

Proof: From that

$$\gamma_{ij} \geq \gamma_{ij}^{'},$$

$i = 1, 2, ..., m, j = 1, 2, ..., n$, it follows that

$$\widetilde{A}(\otimes) \geq \widetilde{A}'(\otimes) \geq 0.$$

Assume the kth column satisfies that

$$\widetilde{P}_k(\otimes) \geq \widetilde{P}'_k(\otimes).$$

1. $\widetilde{P}_k(\otimes)$ is not a basis vector.

When $\widetilde{P}_k(\otimes)$ is changed to $\widetilde{P'}_k(\otimes)$, the basis $\widetilde{B}(\otimes)$ does not change. However, the test number

$$r'_k = \widetilde{C}_k(\otimes) - \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{P'}_k(\otimes)$$

may have been changed.

(a) If $r'_k \leq 0$, then the optimal solution of $LP(\ (\alpha_j, \beta_i, \gamma_{ij})|\ i = 1, 2, ..., m,$ $j = 1, 2, ..., n)$ is still the optimal solution of $LP((\alpha_j, \beta_i, \gamma'_{ij})|\ i = 1, 2, ..., m,$ $j = 1, 2, ..., n)$. And, the optimal value doesn't change. So,

$$\max S = \max S'.$$

(b) If $r'_k > 0$, then x'_k, which corresponds to $\widetilde{P}'_k(\otimes)$, will become a basis variable. We can obtain the optimal solution

$$X' = [\widetilde{B}_1^{-1}(\otimes)\widetilde{b}(\otimes), 0]^T$$

of $LP((\alpha_j, \beta_i, \gamma'_{ij})|\ i = 1, 2, ..., m,\ j = 1, 2, ..., n)$ by using the simplex algorithm. Noticing that

$$[\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes), 0]^T$$

is the feasible basic solution of $LP((\alpha_j, \beta_i, \gamma'_{ij})|\ i = 1, 2, ..., m,\ j = 1, 2, ..., n)$, we have

$$\begin{aligned} \max S &= \widetilde{C}_B(\otimes)\widetilde{B}^{-1}(\otimes)\widetilde{b}(\otimes) \\ &\leq \widetilde{C}_{B_1}(\otimes)\widetilde{B}_1^{-1}(\otimes)\widetilde{b}(\otimes) = \max S'. \ \square \end{aligned}$$

According to Theorems 11.4.1, 2 and 3, we know that the optimal value of a positioned programming is an increasing function about the positioned coefficients α_j ($j = 1, 2, ..., n$) of the price vector and the positioned coefficients β_i ($i = 1, 2, ..., m$) of the constraint vector, and a decreasing function about the positioned coefficients γ_{ij} ($i = 1, 2, ..., m, j = 1, 2, ..., n$) of consumption.

Definition 11.4.1. *Assume that* $\forall i = 1, 2, \cdots, m$ *and* $j = 1, 2, ..., n$

$$\alpha_j = \alpha, \beta_i = \beta, \gamma_{ij} = \gamma.$$

Then, the corresponding positioned programming is called a (α, β, γ)*-positioned programming. It is written as* $LP(\alpha,\beta,\gamma)$*. Its optimal value is denoted* $\max S(\alpha, \beta, \gamma)$*, called the* (α,β,γ)*-positioned optimal value.*

Theorem 11.4.4. *For a positioned programming* $LP(\alpha,\beta,\gamma)$*, we have*

1.When $\alpha=\alpha_0$ *and* $\beta=\beta_0$ *are fixed, if* $\gamma_1 \leq \gamma_2$*, then*

$$\max S(\alpha_0,\beta_0,\gamma_1) \geq \max S(\alpha_0,\beta_0,\gamma_2).$$

2. When $\beta=\beta_0$ *and* $\gamma=\gamma_0$ *are fixed, if* $\alpha_1 \leq \alpha_2$*, then*

$$\max S(\alpha_1,\beta_0,\gamma_0) \leq \max S(\alpha_2,\beta_0,\gamma_0).$$

3. When $\alpha=\alpha_0$ *and* $\gamma=\gamma_0$ *are fixed, if* $\beta_1 \leq \beta_2$*, then*

$$\max S(\alpha_0,\beta_1,\gamma_0) \leq \max S(\alpha_0,\beta_2,\gamma_0).$$

Here, α reflects the general price level of n kinds of products; β reflects the general supplying state of m kinds of resources; and γ is a collective reflection of the level of manufacturing technique, the quantity of the labor force, and managerial level applied in production.

11.5 Satisfactory Solutions of Grey Linear Programming

In this section, we study grey linear programming problems such that not necessarily the optimal solutions but satisfactory solutions can be practically reached.

Definition 11.5.1. *When* $\alpha=\beta=1$ *and* $\gamma=0$*, the corresponding positioned programming* $LP(1,1,0)$ *is called an ideal model of the LPGP. Its optimal value is written as* $\max \overline{S}$*.*

The ideal model stands for an ideal condition such that the highest prices of its products, the most sufficient resource supply, the most developed manufacturing technique, and the quality of labor force and managerial level are all at their optimal states. The aim is to demonstrate the feasibility of a new project. In fact, only a few firms can potentially come up to the ideal state.

Definition 11.5.2. *When* $\alpha=\beta=0$ *and* $\gamma=1$*, the corresponding positioned programming* $LP(0,0,1)$ *is called a critical model of the LPGP. Its optimal value is written as* $\max\underline{S}$*.*

The critical model stands for a condition such that the lowest prices, the shortest resource supply, the less-developed manufacturing technique, and the lowest quantity of labor force and managerial level are employed. With such a condition in place, the firm is at the edge of bankruptcy. The only choice for the firm to take is to change its products, to improve its production techniques, to find alternative resources, and to reeducate its management and workers all around.

Definition 11.5.3. *When $\alpha = \beta = \gamma = \theta$, the corresponding positioned programming is called a θ-positioned programming. It is written as $LP(\theta)$. Similarly, its optimal value is written as $\max S(\theta)$, which is called the θ-positioned optimal value.*

Especially when $\theta = 0.5$, the corresponding θ- positioned programming $LP(0.5)$ is called the mean whitenization programming. Generally, the mean whitenization programming is the most typical one for $LPGP$.

Theorem 11.5.1. $\forall\ \alpha, \beta, \gamma, \theta \in [0, 1]$, *we have*

$$1.\quad \max \underline{S} \leq \max S(\alpha, \beta, \gamma) \leq \max \overline{S},$$

and

$$2.\quad \max \underline{S} \leq \max S(\theta) \leq \max \overline{S}.$$

Proof: We prove 1. only. The second statement is left to the reader to prove.

Because

$$0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, 0 \leq \gamma \leq 1,$$

from Theorem 11.4.4, it follows that

$$\begin{aligned}\max \underline{S} &\leq \max S(0, 0, 1) \leq \max S(\alpha, 0, 1) \\ &\leq \max S(\alpha, \beta, 1) \leq \max S(\alpha, \beta, \gamma).\end{aligned}$$

Similarly, we can prove that $\max \overline{S} \geq \max S(\alpha, \beta, \gamma)$. □

Definition 11.5.4. *For fixed* $\alpha, \beta, \gamma \in [0, 1]$,

$$\mu(\alpha, \beta, \gamma) = \frac{1}{2}\left(1 - \frac{\max \underline{S}}{\max S(\alpha, \beta, \gamma)}\right) + \frac{1}{2}\frac{\max S(\alpha, \beta, \gamma)}{\max \overline{S}}$$

is called the pleased degree of the positioned programming $LP(\alpha, \beta, \gamma)$.

The pleased degree of $LP(\alpha, \beta, \gamma)$ reflects the relationship among the positioned optimal value $\max S(\alpha, \beta, \gamma)$, the optimal value $\max \underline{S}$ of its critical model, and the optimal value $\max \overline{S}$ of the ideal model. The nearer $\max S(\alpha, \beta, \gamma)$ approaches $\max \overline{S}$, the bigger $\mu(\alpha, \beta, \gamma)$ is; the nearer $\max S(\alpha, \beta, \gamma)$ approaches $\max \underline{S}$, the smaller $\mu(\alpha, \beta, \gamma)$ is.

Similarly, we can define the concept of pleased degree of $\mu(\theta)$ for θ-positioned programming $LP(\theta)$.

Proposition 11.5.1. $\forall\ \alpha, \beta, \gamma \in [0, 1]$, *we have that*

$$0 \leq \mu(\alpha, \beta, \gamma) \leq 1.$$

Definition 11.5.5. *Given a grey target $D = [\mu_0, 1]$, if $\mu(\alpha, \beta, \gamma) \in D$, then the corresponding optimal solution is called the pleased solution of the LPGP.*

Example 11.5.1. For the grey linear programming problem as shown in Example 11.3.1, assume that the daily electricity supply $b_1(\otimes) \in [150, 235]$, daily coal supply $b_2(\otimes) \in [280, 360]$, and movable labor $b_3(\otimes) \in [270, 330]$. Try to find

1. The ideal optimal value $\max \overline{S}$;
2. The critical optimal value $\max \underline{S}$;
3. The θ-positioned optimal value, when $\theta = 0.6$;
4. The (α, β, γ)-positioned optimal value with $\alpha = 0.7$, $\beta = 0.9$, $\gamma = 0.5$; and
5. Study the pleased degrees of these optimal values

of the grey linear programming of the drifting type

$$\max S = C(\otimes)X$$

$$s.t. \begin{cases} A(\otimes)X \leq b(\otimes) \\ X \geq 0. \end{cases}$$

Solution: 1. Find the ideal optimal value $\max \overline{S}$.
Take $\alpha = 1$, $\beta = 1$, and $\gamma = 0$. Then

$$\overline{C} = [\overline{c}_1, \overline{c}_2] = [800, 1500],$$

$$\overline{b} = \left[\overline{b}_1, \overline{b}_2, \overline{b}_3\right]^T = [235, 360, 330]^T,$$

and

$$\underline{A} = \begin{bmatrix} \underline{a}_{11} & \underline{a}_{12} \\ \underline{a}_{21} & \underline{a}_{22} \\ \underline{a}_{31} & \underline{a}_{32} \end{bmatrix} = \begin{bmatrix} 3 & 3.5 \\ 7 & 3 \\ 2.5 & 8 \end{bmatrix}.$$

So, we obtain the ideal model as follows

$$\max S = 800x_1 + 1500x_2$$

$$s.t. \begin{cases} 3x_1 + 3.5x_2 \leq 235 \\ 7x_1 + 3x_2 \leq 360 \\ 2.5x_1 + 8x_2 \leq 330 \\ x_1 \geq 0, \ x_2 \geq 0, \end{cases}$$

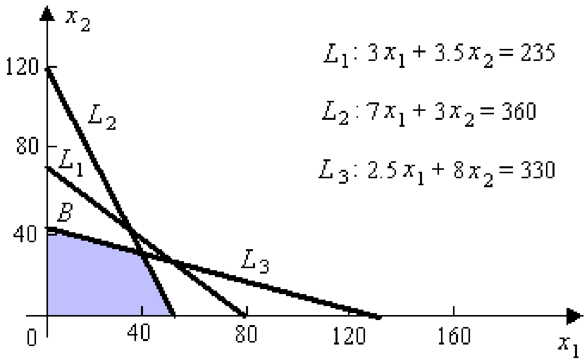

FIGURE 11.2. The feasible region of the ideal model

whose feasible region is shown in Figure 11.2.

All the lines parallel to the objective function $S = 800x_1 + 1500x_2$ intersect the last point at the right upper corner of the feasible region at point B, which is the optimal point. From

$$\begin{cases} 2.5x_1 + 8x_2 = 330 \\ x_1 = 0 \end{cases}$$

we obtain the optimal solution

$$x_1 = 0, x_2 = 41.25$$

and the ideal optimal value

$$\max \overline{S} = 800 \cdot 0 + 1500 \cdot 41.25 = 61875.$$

2. Find the critical optimal value $\max \underline{S}$.

Take $\alpha = 0$, $\beta = 0$, and $\gamma = 1$. Then, we have

$$\underline{C} = [\underline{c}_1, \underline{c}_2] = [600, 900],$$

$$\underline{b} = [\underline{b}_1, \underline{b}_2, \underline{b}_3]^T = [150, 280, 270]^T,$$

and

$$\overline{A} = \begin{bmatrix} \overline{a}_{11} & \overline{a}_{12} \\ \overline{a}_{21} & \overline{a}_{22} \\ \overline{a}_{31} & \overline{a}_{32} \end{bmatrix} = \begin{bmatrix} 5 & 6.5 \\ 11 & 5 \\ 3.5 & 12 \end{bmatrix}.$$

So, we have the critical model

$$\max S = 600x_1 + 900x_2$$

$$s.t. \begin{cases} 5x_1 + 6.5x_2 \le 150 \\ 11x_1 + 5x_2 \le 280 \\ 3.5x_1 + 12x_2 \le 270 \\ x_1 \ge 0,\ x_2 \ge 0. \end{cases}$$

Similarly, we know that the optimal solution is the solution of the system

$$\begin{cases} 3.5x_1 + 12x_2 = 270 \\ 5x_1 + 6.5x_2 = 150. \end{cases}$$

That is,

$$x_1 = 1.2,\ x_2 = 22.15.$$

So, the critical optimal value is

$$\max \underline{S} = 600 \cdot 1.2 + 900 \cdot 22.15 = 20655.$$

3. Find the θ-positioned optimal value $\max S(0.6)$, when $\theta = 0.6$. Take $\alpha = \beta = \gamma = \theta = 0.6$. Then

$$\widetilde{c}_1(\otimes) = 0.6\overline{c}_1 + 0.4\underline{c}_1 = 0.6 \cdot 800 + 0.4 \cdot 600 = 720,$$

$$\widetilde{c}_2(\otimes) = 0.6\overline{c}_2 + 0.4\underline{c}_2 = 0.6 \cdot 1500 + 0.4 \cdot 900 = 1260,$$

so,

$$\widetilde{c}(\otimes) = [\widetilde{c}_1(\otimes), \widetilde{c}_2(\otimes)] = [720, 1260].$$

Similarly, we have

$$\widetilde{b}(\otimes) = \left[\widetilde{b}_1(\otimes), \widetilde{b}_2(\otimes), \widetilde{b}_3(\otimes)\right]^T = [201, 328, 306]^T$$

and

$$\widetilde{A}(\otimes) = \begin{bmatrix} \widetilde{\otimes}_{11} & \widetilde{\otimes}_{12} \\ \widetilde{\otimes}_{21} & \widetilde{\otimes}_{22} \\ \widetilde{\otimes}_{31} & \widetilde{\otimes}_{32} \end{bmatrix} = \begin{bmatrix} 4.2 & 5.3 \\ 9.4 & 4.2 \\ 3.1 & 10.4 \end{bmatrix}.$$

Therefore, we obtain the θ- positioned model as follows

$$\max S = 720x_1 + 1260x_2$$

$$s.t.\begin{cases} 4.2x_1 + 5.3x_2 \le 201 \\ 9.4x_1 + 4.2x_2 \le 328 \\ 3.1x_1 + 10.4x_2 \le 306 \\ x_1 \ge 0,\ x_2 \ge 0, \end{cases}$$

whose optimal solution is the solution of the system

$$\begin{cases} 3.1x_1 + 10.4x_2 = 306 \\ 4.2x_1 + 5.3x_2 = 201. \end{cases}$$

That is, the optimal solution is

$$x_1 = 17.19,\ x_2 = 24.30$$

with the corresponding 0.6-positioned optimal value

$$\max S(0.6) = 720 \cdot 17.19 + 1260 \cdot 24.3 = 42994.8.$$

4. Find the (α, β, γ)-positioned optimal value $\max S(0.7, 0.9, 0.5)$, when $\alpha = 0.7$, $\beta = 0.9$, and $\gamma = 0.5$.

From $\alpha = 0.7$, it follows that

$$\widetilde{c}_1(\otimes) = 0.7\overline{c}_1 + 0.3\underline{c}_1 = 0.7 \cdot 800 + 0.3 \cdot 600 = 740,$$

$$\widetilde{c}_2(\otimes) = 0.7\overline{c}_2 + 0.3\underline{c}_2 = 0.7 \cdot 1500 + 0.3 \cdot 900 = 1320,$$

so,

$$\widetilde{c}(\otimes) = [\widetilde{c}_1(\otimes), \widetilde{c}_2(\otimes)] = [740, 1320]\,.$$

Similarly, we have

$$\widetilde{b}(\otimes) = \left[\widetilde{b}_1(\otimes), \widetilde{b}_2(\otimes), \widetilde{b}_3(\otimes)\right]^T = [226.5, 352, 324]^T$$

and

$$\widetilde{A}(\otimes) = \begin{bmatrix} \widetilde{\otimes}_{11} & \widetilde{\otimes}_{12} \\ \widetilde{\otimes}_{21} & \widetilde{\otimes}_{22} \\ \widetilde{\otimes}_{31} & \widetilde{\otimes}_{32} \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 9 & 4 \\ 3 & 10 \end{bmatrix}.$$

Therefore, we obtain the (α, β, γ)-positioned model as follows.

$$\max S = 740x_1 + 1320x_2$$

$$s.t. \begin{cases} 4x_1 + 5x_2 \leq 226.5 \\ 9x_1 + 4x_2 \leq 352 \\ 3x_1 + 10x_2 \leq 324 \\ x_1 \geq 0, \ x_2 \geq 0. \end{cases}$$

Studying its feasible region and objective function shows that the optimal solution is the solution of the system

$$\begin{cases} 4x_1 + 5x_2 = 226.5 \\ 3x_1 + 10x_2 = 324. \end{cases}$$

That is, the optimal solution is

$$x_1 = 25.8, \ x_2 = 24.66,$$

and the corresponding (α, β, γ)-positioned optimal value is

$$\begin{aligned} &\max S(0.7, 0.9, 0.5) \\ &= 740 \cdot 25.8 + 1320 \cdot 24.66 = 51643.2. \end{aligned}$$

5. Take $\mu_0 = 0.5$. Then for the 0.6-positioned model, we have

$$\begin{aligned} \mu &= \frac{1}{2} \cdot \left(1 - \frac{\max \underline{S}}{\max S(0.6)}\right) + \frac{1}{2} \cdot \frac{\max S(0.6)}{\max \overline{S}} \\ &= \frac{1}{2} \cdot \left(1 - \frac{20655}{42994.8}\right) + \frac{1}{2} \cdot \frac{42994.8}{61875} \\ &\approx 0.62 > \mu_0. \end{aligned}$$

So, the optimal solution corresponding to $\max S(0.6)$

$$x_1 = 17.19, x_2 = 24.30$$

is the pleased solution of the grey linear programming of the drifting type.

For the (α, β, γ)-positioned model with $\alpha = 0,7$, $\beta = 0.9$, and $\gamma = 0.5$, we have

$$\begin{aligned}
\mu &= \frac{1}{2} \cdot \left(1 - \frac{\max \underline{S}}{\max S(\alpha, \beta, \gamma)}\right) + \frac{1}{2} \cdot \frac{\max S(\alpha, \beta, \gamma)}{\max \overline{S}} \\
&= \frac{1}{2} \cdot \left(1 - \frac{20655}{51643.2}\right) + \frac{1}{2} \cdot \frac{51643.2}{61875} \\
&\approx 0.72 > \mu_0.
\end{aligned}$$

So, the optimal solution corresponding to $\max S(0.7, 0.9, 0.5)$

$$x_1 = 25.8, x_2 = 24.66$$

is the pleased solution of the grey linear programming of the drifting type.

11.6 Quasi-Optimal Solutions of Grey Linear Programming

In the process of solving a linear programming problem, very often the researcher meets such a situation that the optimal solution cannot be found. In this case, the researcher may consider using other methods to seek an approximate optimal solution to the problem. In this section, we mainly study the method of alternative optimization with decision variables. The whole process of the method of alternative optimization is described by the following steps.

Step 1: Determine the positioned programming problem:

$$\max S = \widetilde{C}(\otimes)X$$

$$s.t. \begin{cases} \widetilde{A}(\otimes)X \le \widetilde{b}(\otimes) \\ X \ge 0 \end{cases}$$

of the following grey linear programming problem

$$\max S = C(\otimes)X$$

$$s.t. \begin{cases} A(\otimes)X \le b(\otimes) \\ X \ge 0. \end{cases}$$

Step 2: Solve the positioned programming problem using a regular linear programming method until the calculation cannot be continued. Assume the last feasible solution is

$$X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, ..., x_m^{(0)}\right).$$

Step 3: Solve for $x_1^{(1)}$ with the starting point $X^{(0)}$, for fixed $x_2^{(0)}, x_3^{(0)}$, ..., $x_m^{(0)}$. Assume that

$$X^{(1)} = \left(x_1^{(1)}, x_2^{(0)}, ..., x_m^{(0)}\right)$$

is the optimal solution with $x_2^{(0)}, x_3^{(0)}, ..., x_m^{(0)}$ fixed. Second, optimize x_2 with starting point $X^{(1)}$. Assume that

$$X^{(2)} = \left(x_1^{(1)}, x_2^{(1)}, x_3^{(0)}, ..., x_m^{(0)}\right)$$

is the optimal solution with $x_1^{(1)}, x_3^{(0)}, x_4^{(0)}, ..., x_m^{(0)}$ fixed. Third, optimize x_3 with the starting point $X^{(2)}$. This process continues until we have found

$$X^{(m)} = \left(x_1^{(1)}, x_2^{(1)}, ..., x_m^{(1)}\right).$$

Step 4: With the starting point $X^{(m)}$, we repeat the exploration as described in Step 3, and obtain that

$$X^{(2m)} = \left(x_1^{(2)}, x_2^{(2)}, ..., x_m^{(2)}\right),$$

$$X^{(3m)} = \left(x_1^{(3)}, x_2^{(3)}, ..., x_m^{(3)}\right),$$

....

$$X^{(km)} = \left(x_1^{(k)}, x_2^{(k)}, ..., x_m^{(k)}\right),$$

....

until $X^{(km)} = X^{[(k-1)m]}$ or $X^{(km)}$ is sufficiently close to $X^{[(k-1)m]}$ and the corresponding objective function values are sufficiently close to each other.

Definition 11.6.1. *The final solution out of the method of alternative optimization*

$$X^{(km)} = \left(x_1^{(k)}, x_2^{(k)}, ..., x_m^{(k)}\right)$$

is called the quasi-optimal solution of the grey linear programming, and the corresponding objective function value is called the quasi-optimal value.

Example 11.6.1. Assume that the positioned programming of a certain grey linear programming is given as follows.

$$\max S = \frac{5}{2}x_1 + \tfrac{1}{2}x_2 + 2x_3 + 2x_4$$

$$s.t. \begin{cases} 4x_1 + x_2 + 2x_3 \le 100 \\ x_1 + x_3 + 2x_4 \le 80 \\ \tfrac{1}{2}x_1 + x_3 + 2x_4 \le 60 \\ 0 \le x_j \le 15, \text{ j=1,2,3,4.} \end{cases}$$

Find its quasi-optimal solution.

Solution: Assume that the least feasible solution, obtained by using the simplex algorithm, is

$$X^{(0)} = \left(x_i^{(0)}\right)_{i=1}^4 = (12, 10, 10, 5)\,,$$

and the corresponding objective function value is

$$\begin{aligned} S^{(0)} &= \frac{5}{2}x_1^{(0)} + \frac{1}{2}x_2^{(0)} + 2x_3^{(0)} + 2x_4^{(0)} \\ &= \frac{5}{2}\cdot 12 + \frac{1}{2}\cdot 10 + 2\cdot 10 + 2\cdot 5 = 65. \end{aligned}$$

Now, we start the first round of alternative optimization with the point $X^{(0)}$.

First, we optimize x_1 with fixed $x_2^{(0)}, x_3^{(0)}, x_4^{(0)}$.

Substituting $x_2^{(0)} = 10$, $x_3^{(0)} = 10$, and $x_4^{(0)} = 5$ into the constraint inequalities gives

$$\begin{cases} 4x_1 + x_2^{(0)} + 2x_3^{(0)} \le 100 \\ x_1 + x_3^{(0)} + 2x_4^{(0)} \le 80 \\ \frac{1}{2}x_1 + x_3^{(0)} + 2x_4^{(0)} \le 60 \\ 0 \le x_1 \le 15, \end{cases}$$

that is,

$$\begin{cases} 4x_1 \le 100 - 10 - 20 = 70 \\ x_1 \le 80 - 10 - 10 = 60 \\ \dfrac{1}{2}x_1 \le 60 - 10 - 10 = 40 \\ 0 \le x_1 \le 15, \end{cases}$$

it follows that

$$\begin{cases} 4x_1 \le 17.5 \\ x_1 \le 60 \\ x_1 \le 80 \\ 0 \le x_1 \le 15. \end{cases}$$

So, $x_1^{(1)} = 15$ is optimal. Now, use

$$X^{(1)} = \left(x_1^{(1)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}\right) = (15, 10, 10, 5)\,,$$

as the starting point to optimize x_2 with $x_1^{(1)}, x_3^{(0)}, x_4^{(0)}$ fixed. Substituting $x_1^{(1)} = 15$, $x_3^{(0)} = 10$, and $x_4^{(0)} = 5$ into the constraint inequalities gives that

$$\begin{cases} x_2 \le 100 - 4x_1^{(1)} - 2x_3^{(0)} = 100 - 60 - 20 = 30 \\ 0 \le x_2 \le 15. \end{cases}$$

That is,

$$\begin{cases} x_2 \le 20 \\ 0 \le x_2 \le 15. \end{cases}$$

So, $x_2^{(1)} = 15$ is optimal. Now, use

$$X^{(2)} = \left(x_1^{(1)}, x_2^{(1)}, x_3^{(0)}, x_4^{(0)}\right) = (15, 15, 10, 5)$$

as the starting point to optimize x_3 with $x_1^{(1)}, x_2^{(1)}, x_4^{(0)}$ fixed. Substituting $x_1^{(1)} = 15$, $x_2^{(1)} = 15$, and $x_4^{(0)} = 5$ into the constraint inequalities gives that

$$\begin{cases} 2x_3 \leq 100 - 4x_1^{(1)} - x_2^{(1)} = 100 - 60 - 15 = 25 \\ x_3 \leq 80 - x_1^{(1)} - 2x_4^{(0)} 80 - 15 - 10 = 55 \\ x_3 \leq 60 - \frac{1}{2}x_1^{(1)} - 2x_4^{(0)} 60 - 7.5 - 10 = 42.5 \\ 0 \leq x_3 \leq 15. \end{cases}$$

That is,

$$\begin{cases} x_3 \leq 12.5 \\ x_3 \leq 55 \\ x_3 \leq 42.5 \\ 0 \leq x_3 \leq 15. \end{cases}$$

So, $x_3^{(1)} = 12.5$ is optimal. Now, use

$$X^{(3)} = \left(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(0)}\right) = (15, 15, 12.5, 5)$$

as the starting point to optimize x_4 with $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}$ fixed. Substituting $x_1^{(1)} = 15$, $x_2^{(1)} = 15$, and $x_3^{(1)} = 12.5$ into the constraint inequalities gives that

$$\begin{cases} 2x_4 \leq 80 - x_1^{(1)} - x_3^{(1)} = 80 - 15 - 12.5 = 52.5 \\ 2x_4 \leq 60 - \frac{1}{2}x_1^{(1)} - x_3^{(1)} = 60 - 7.5 - 12.5 = 40 \\ 0 \leq x_4 \leq 15. \end{cases}$$

That is,

$$\begin{cases} x_4 \leq 26.25 \\ x_4 \leq 20 \\ x_4 \leq 15. \end{cases}$$

So, $x_4^{(1)} = 15$ is optimal. Hence, we have obtained the result of the first round alternative optimization as follows,

$$X^{(4)} = \left(x_i^{(1)}\right)_{i=1}^{4} = (15, 15, 12.5, 15),$$

the corresponding objective function value is

$$\begin{aligned} S^{(4)} &= \frac{5}{2}x_1^{(1)} + \frac{1}{2}x_2^{(1)} + 2x_3^{(1)} + 2x_4^{(1)} \\ &= \frac{5}{2} \cdot 15 + \frac{1}{2} \cdot 15 + 2 \cdot 12.5 + 2 \cdot 15 = 100. \end{aligned}$$

Because $S^{(4)} > S^{(0)}$, $X^{(4)}$ is more optimal than $X^{(0)}$. Now, we start the second round alternative optimization with the starting point $X^{(4)}$. Because three of the four decision variables have reached their upper limits and the other variable x_3, which has not reached its upper limit, is also optimal when the other variables are fixed, we have the result of the second round alternative optimization

$$X^{(8)} = X^{(4)}.$$

So,

$$X^{(4)} = (15, 15, 12.5, 15)$$

is the quasi-optimal solution, and $S^{(4)} = 100$ is the quasi-optimal value.

If the quasi-optimal solution is still not satisfactory, then further optimization can be done based on the magnitude of the coefficients of the decision variables in the objective function. For example, in Example 11.6.1, because

$$\widetilde{C}(\otimes) = (\widetilde{c}_i(\otimes))_{i=1}^4 = \left(\frac{5}{2}, \frac{1}{2}, 2, 2\right),$$

where

$$\widetilde{c}_2(\otimes) = \frac{1}{2} < \widetilde{c}_3(\otimes) = 2,$$

so, we conclude that

$$x_2^{(1)} = 15 > x_3^{(1)} = 12.5$$

is not very reasonable. If it is allowed by the constraint conditions, we can increase the objective function value through decreasing x_2 and increasing x_3. Now, substituting $x_1^{(1)} = 15$ and $x_4^{(1)} = 15$ into the constraint inequalities gives that

$$\begin{cases} x_2 + 2x_3 \le 100 - 4x_1^{(1)} = 100 - 60 = 40 \\ 0 \le x_2 \le 15 \\ 0 \le x_3 \le 15. \end{cases}$$

Taking $x_3^{(2)} = 15$ gives

$$\begin{cases} x_2 \le 40 - 30 = 10 \\ 0 \le x_2 \le 15. \end{cases}$$

So, $x_2^{(2)} = 10$ is optimal. Hence, it follows that

$$X^{(5)} = \left(x_1^{(1)}, x_2^{(2)}, x_3^{(2)}, x_4^{(1)}\right) = (15, 10, 15, 15)\,.$$

The corresponding objective function value satisfies

$$S^{(5)} = 102.5 > S^{(4)} = 100;$$

that is, $X^{(5)}$ is more optimal than $X^{(4)}$.

11.7 Grey 0-1 Programming

Most typical in 0-1 *programming* are the so-called *assignment problems*. In this section, we mainly discuss how to solve assignment problems of grey prediction type.

Definition 11.7.1. *Let n tasks be assigned to m people. Assume that each person will finish only one task. When $n = m$, this kind of assignment is called a balanced assignment problem.*

Definition 11.7.2. *In a balanced assignment problem, let*

$$x_{ij} = \begin{cases} 1, & \text{if the } i\text{th task is assigned to the } j\text{th person} \\ 0, & \text{if the } i\text{th task is not assigned to the } j\text{th person.} \end{cases}$$

Assume that c_{ij} is the expense for the jth person to accomplish the ith task, for i, $j = 1, 2, ..., n$. Then,

$$\min S = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$s.t. \begin{cases} \sum_{j=1}^{n} x_{ij} = 1, & i = 1, 2, ..., n \\ \sum_{i=1}^{n} x_{ij} = 1, & j = 1, 2, ..., n \\ x_{ij} = 0 \text{ or } 1, & i,\ j = 1, 2, ..., n \end{cases}$$

is called a mathematical model of the assignment problem.

Here, the constraints

$$\sum_{j=1}^{n} x_{ij} = 1,$$

$i = 1, 2, ..., n$, represent that each task is assigned only to one person, and the constraints

$$\sum_{i=1}^{n} x_{ij} = 1,$$

$j = 1, 2, ..., n$, stand for that each person only accepts one task.

Definition 11.7.3. *The square matrix*

$$C = [c_{ij}]_{n \times n}$$

is called an $efficiency$ matrix.

Theorem 11.7.1. *If a constant is added to all the entries of a row or a column of the efficiency matrix C, then the optimal assignment, obtained from the new efficiency matrix is the same as that obtained from C.*

Proof: Assume that e_i and t_j are constants, and

$$d_{ij} = c_{ij} + e_i + t_j,$$

i, $j = 1, 2, ..., n$. Then, the new objective function is

$$\begin{aligned} S' &= \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}x_{ij} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} (c_{ij} + e_i + t_j)x_{ij} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{n} e_i x_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{n} t_j x_{ij} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{i=1}^{n} e_i \sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} t_j \sum_{i=1}^{n} x_{ij} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{i=1}^{n} e_i + \sum_{j=1}^{n} t_j. \end{aligned}$$

Because $\sum_{i=1}^{n} e_i$ and $\sum_{j=1}^{n} t_j$ are constants, S' and S take the minimum value at the same time. □

Definition 11.7.4. *When the entries in the efficiency matrix are some grey predicted values or grey development coefficients of efficiency sequences, the corresponding 0-1 programming is called grey 0-1 programming.*

When the values c_{ij} in the original problem are values for benefits, and the objective function is

$$\max S = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$$

we can take

$$c_{i_0 j_0} = \max_{1 \le i \le n} \max_{1 \le j \le n} \{c_{ij}\},$$

and let

$$c'_{ij} = c_{i_0 j_0} - c_{ij},$$

$i,\ j = 1, 2, ..., n$. Then, the objective function is converted to

$$\min S = \sum_{i=1}^{n} \sum_{j=1}^{n} c'_{ij} x_{ij}.$$

The solution process of grey 0-1 programming is described as follows.

Step 1: Collect the benefit time sequence

$$u_{ij}^{(0)} = \left(u_{ij}^{(0)}(1), u_{ij}^{(0)}(2), ..., u_{ij}^{(0)}(h)\right),$$

$i,\ j = 1, 2, ..., n$.

Step 2: Establish the GM(1, 1) time response series

$$\begin{cases} \widehat{u}_{ij}^{(1)}(k+1) & = \omega_{ij} \cdot e^{-a_{ij} k} - \varpi_{ij} \\ \widehat{u}_{ij}^{(0)}(k+1) & = \widehat{u}_{ij}^{(1)}(k+1) - \widehat{u}_{ij}^{(1)}(k), \end{cases}$$

for

$$u_{ij}^{(0)} = \left(u_{ij}^{(0)}(1), u_{ij}^{(0)}(2), ..., u_{ij}^{(0)}(h)\right)$$

$i,\ j = 1, 2, ..., n$.

Step 3: Write out the benefit matrix $C = [c_{ij}]_{n \times n}$. Here, we can define

$$c_{ij} = \widehat{u}_{ij}^{(0)}(h+s) \text{ or } c_{ij} = -a_{ij},$$

$i,\ j = 1, 2, ..., n$.

Step 4: Compute $c_{i_0 j_0} = \max_{1 \le i \le n} \max_{1 \le j \le n} \{c_{ij}\}$.

Step 5: Let

$$c'_{ij} = c_{i_0 j_0} - c_{ij},$$

$i,\ j = 1, 2, ..., n$. We obtain the following grey 0-1 programming model,

$$\min S = \sum_{i=1}^{n} \sum_{j=1}^{n} c'_{ij} x_{ij}$$

$$s.t. \begin{cases} \sum_{j=1}^{n} x_{ij} = 1, & i = 1, 2, ..., n \\ \sum_{i=1}^{n} x_{ij} = 1, & j = 1, 2, ..., n \\ x_{ij} = 0 \text{ or } 1, & i,\ j = 1, 2, ..., n. \end{cases}$$

Step 6: Convert the efficiency matrix

$$C' = \left[c'_{ij}\right]_{n\times n}.$$

Subtract the minimum entry of each row and each column of the efficiency matrix C' so that each row and each column contain at least one zero entry. If the number of zeros located at different rows and different columns equals the rank n of the efficiency matrix, stop the conversion. Otherwise, repeat the previous conversion until the number of zeros, located at different rows and different columns, equals the rank n of the efficiency matrix.

Step 7: Add "( )" to the n zero entries located at different row and different columns. These zero entries are called *independent zero*. And let

$$x_{ij} = \begin{cases} 1, & \text{if there exists an independent zero at } (i\ j) \text{ location} \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$X = \{x_{ij} \mid i,\ j = 1, 2, ..., n\}$$

is the optimal solution for which we are looking.

Example 11.7.1. There are three economic districts. All the districts need to develop farming, but each district needs to have a development emphasis in one of the three directions: industry and service, livestock husbandry, and forestry so that the overall benefits of the three districts will be maximized. Try to do a grey 0-1 programming.

First, we find the solution using predicted values.

Step 1: According to the statistical records available, we obtain the benefit time series for the ith district to develop in the jth business direction:

$$u_{ij}^{(0)} = \left(u_{ij}^{(0)}(1), u_{ij}^{(0)}(2), u_{ij}^{(0)}(3), u_{ij}^{(0)}(4)\right),$$

$i = 1, 2, 3$; $j = 1, 2, 3$, where $u_{ij}^{(0)}(1), u_{ij}^{(0)}(2), u_{ij}^{(0)}(3), u_{ij}^{(0)}(4)$ are, respectively, the benefit values for the years 1993, 1994, 1995, and 1996. Specifically, we have

$$u_{11}^{(0)} = \left(u_{11}^{(0)}(i)\right)_{i=1}^{4} = (4, 4.2, 4.4, 5),$$

$$u_{12}^{(0)} = \left(u_{12}^{(0)}(i)\right)_{i=1}^{4} = (5, 6, 8, 12),$$

$$u_{13}^{(0)} = \left(u_{13}^{(0)}(i)\right)_{i=1}^{4} = (0.9, 0.8, 2, 3);$$

$$u_{21}^{(0)} = \left(u_{21}^{(0)}(i)\right)_{i=1}^{4} = (3, 4, 4.2, 6)\,,$$

$$u_{22}^{(0)} = \left(u_{22}^{(0)}(i)\right)_{i=1}^{4} = (7, 9, 10, 12)\,,$$

$$u_{23}^{(0)} = \left(u_{23}^{(0)}(i)\right)_{i=1}^{4} = (0.1, 0.4, 0.5, 0.7)\,;$$

$$u_{31}^{(0)} = \left(u_{31}^{(0)}(i)\right)_{i=1}^{4} = (1, 2, 2.5, 3)\,,$$

$$u_{32}^{(0)} = \left(u_{32}^{(0)}(i)\right)_{i=1}^{4} = (3, 4.5, 5, 5.8)\,,$$

and

$$u_{33}^{(0)} = \left(u_{33}^{(0)}(i)\right)_{i=1}^{4} = (2, 3, 4, 6)\,.$$

Step 2: For $i = 1, 2, 3$, and $j = 1, 2, 3$, find the GM(1, 1) time response sequence

$$\begin{cases} \widehat{u}_{ij}^{(1)}(k+1) = \omega_{ij} \cdot e^{-a_{ij}k} - \varpi_{ij} \\ \widehat{u}_{ij}^{(0)}(k+1) = \widehat{u}_{ij}^{(1)}(k+1) - \widehat{u}_{ij}^{(1)}(k); \end{cases}$$

that is,

$$\widehat{u}_{ij}^{(0)}(k+1) = \omega_{ij} \cdot (1 - e^{a_{ij}}) \cdot e^{-a_{ij}k}.$$

It follows that

$$\widehat{u}_{11}^{(0)}(k+1) = 3.7736 \cdot e^{0.0894k}, \widehat{u}_{12}^{(0)}(k+1) = 3.9936 \cdot e^{0.3561k},$$

$$\widehat{u}_{13}^{(0)}(k+1) = 0.5797 \cdot e^{0.5448k}; \widehat{u}_{21}^{(0)}(k+1) = 2.9559 \cdot e^{0.2228k},$$

$$\widehat{u}_{22}^{(0)}(k+1) = 7.62 \cdot e^{0.1472k}, \widehat{u}_{23}^{(0)}(k+1) = 0.288 \cdot e^{0.2883k};$$

$$\widehat{u}_{31}^{(0)}(k+1) = 1.6489 \cdot e^{0.1993k}, \widehat{u}_{32}^{(0)}(k+1) = 3.9142 \cdot e^{0.1285k},$$

and

$$\widehat{u}_{33}^{(0)}(k+1) = 1.9983 \cdot e^{0.3561k}.$$

Step 3: Based on the predicted values for the year 1997,

$$\widehat{u}_{11}^{(0)}(5) = 5.3958, \widehat{u}_{12}^{(0)}(5) = 12.5949, \widehat{u}_{13}^{(0)}(5) = 5.1241;$$

$$\widehat{u}_{21}^{(0)}(5) = 7.2066, \widehat{u}_{22}^{(0)}(5) = 13.7299, \widehat{u}_{23}^{(0)}(5) = 0.9125;$$

$$\widehat{u}_{31}^{(0)}(5) = 3.6594, \widehat{u}_{32}^{(0)}(5) = 6.5444, \widehat{u}_{33}^{(0)}(5) = 8.3037,$$

take

$$c_{ij} = \widehat{u}_{ij}^{(0)}(5),$$

$i, j = 1, 2, 3$. So, we obtain the benefit matrix

$$C = [c_{ij}]_{3\times 3} = \begin{bmatrix} 5.3958 & 16.5949 & 5.1241 \\ 7.2066 & 13.7299 & 0.9125 \\ 3.6594 & 6.5444 & 8.3037 \end{bmatrix}.$$

Step 4:

$$\max_{1\le i\le 3} \max_{1\le j\le 3} \{c_{ij}\} = 16.5949 = c_{12}.$$

Step 5: Let

$$c_{ij}^{(0)} = 16.5949 - c_{ij},$$

$i, j = 1, 2, 3$. We obtain the efficiency matrix

$$C^{(0)} = \left[c_{ij}^{(0)}\right]_{3\times 3} = \begin{bmatrix} 11.1991 & 0 & 11.4708 \\ 9.3883 & 2.865 & 15.6824 \\ 12.9355 & 10.0505 & 8.2912 \end{bmatrix}.$$

Step 6: Convert the efficiency matrix by subtracting the minimum entries from the entries of their own columns. We have that

$$\min_{1\le i\le 3} \{c_{i1}^{(0)}\} = c_{21}^{(0)} = 9.3883,$$

$$\min_{1\le i\le 3} \{c_{i2}^{(0)}\} = c_{12}^{(0)} = 0,$$

$$\min_{1\le i\le 3} \{c_{i3}^{(0)}\} = c_{33}^{(0)} = 8.2912,$$

$$C^{(1)} = \begin{bmatrix} c_{11}^{(0)} - c_{21}^{(0)} & c_{12}^{(0)} - c_{12}^{(0)} & c_{13}^{(0)} - c_{33}^{(0)} \\ c_{21}^{(0)} - c_{21}^{(0)} & c_{22}^{(0)} - c_{12}^{(0)} & c_{23}^{(0)} - c_{33}^{(0)} \\ c_{31}^{(0)} - c_{21}^{(0)} & c_{32}^{(0)} - c_{12}^{(0)} & c_{33}^{(0)} - c_{33}^{(0)} \end{bmatrix}$$

$$= \begin{bmatrix} 1.8108 & (0) & 3.1796 \\ (0) & 2.865 & 7.3912 \\ 3.5472 & 10.0505 & (0) \end{bmatrix},$$

where $C^{(1)}$ already contains three zero entries located at different rows and different columns.

Step 7: Corresponding to the independent zeros, let

$$x_{12} = 1, x_{21} = 1, x_{33} = 1,$$

and $x_{ij} = 0$ for all other i and j. So, we obtain the optimal solution

$$\begin{aligned} X &= (x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}, x_{31}, x_{32}, x_{33}) \\ &= (0, 1, 0, 1, 0, 0, 0, 0, 1)\,. \end{aligned}$$

That is, district 1 should put its development emphasis on livestock husbandry, district 2 on industry and services, and district 3 on forestry so that the overall benefits would be maximized.

In the following, we find the solution based on the development coefficients.

From the GM(1, 1) response sequence, we have the development coefficients

$$-a_{11} = 0.0894, -a_{12} = 0.3561, -a_{13} = 0.5448;$$

$$-a_{21} = 0.2228, -a_{22} = 0.1472, -a_{23} = 0.2883;$$

$$-a_{31} = 0.1993, -a_{32} = 0.1285, -a_{33} = 0.3561.$$

Take

$$c_{ij} = -a_{ij},$$

i, $j = 1, 2, 3$; we obtain the benefit matrix

$$C = [c_{ij}]_{3\times 3} = \begin{bmatrix} 0.0894 & 0.3561 & 0.5448 \\ 0.2228 & 0.1472 & 0.2883 \\ 0.1993 & 0.1285 & 0.3561 \end{bmatrix},$$

and

$$\max_{1\le i\le 3} \max_{1\le j\le 3} \{c_{ij}\} = 0.5448 = c_{13.}$$

Let

$$c_{ij}^{(0)} = c_{13} - c_{ij} = 0.5448 - c_{ij};$$

we obtain the efficiency matrix

$$C^{(0)} = \left[c_{ij}^{(0)}\right]_{3\times 3} = \begin{bmatrix} 0.4554 & 0.1887 & 0 \\ 0.3220 & 0.3976 & 0.2565 \\ 0.3455 & 0.4163 & 0.1887 \end{bmatrix}.$$

Now, in $C^{(0)}$, all the entries in each column subtract their minimum entry. From

$$\min_{1\le i\le 3}\{c_{i1}^{(0)}\} = 0.3220 = c_{21}^{(0)},$$

$$\min_{1\le i\le 3}\{c_{i2}^{(0)}\} = 0.1887 = c_{12}^{(0)},$$

and

$$\min_{1\le i\le 3}\{c_{i3}^{(0)}\} = 0 = c_{13}^{(0)},$$

let

$$c_{i1}^{(1)} = c_{i1}^{(0)} - c_{21}^{(0)} = c_{i1}^{(0)} - 0.3220,$$

$$c_{i2}^{(1)} = c_{i2}^{(0)} - c_{12}^{(0)} = c_{i2}^{(0)} - 0.1887,$$

and

$$c_{i3}^{(1)} = c_{i3}^{(0)} - c_{13}^{(0)} = c_{i3}^{(0)} - 0 = c_{i3}^{(0)},$$

$i = 1, 2, 3$. So it follows that

$$C^{(1)} = \left[c_{ij}^{(1)}\right]_{3\times 3} = \begin{bmatrix} 0.1334 & 0 & 0 \\ 0 & 0.2089 & 0.2565 \\ 0.0235 & 0.2276 & 0.1887 \end{bmatrix}.$$

In $C^{(1)}$, all the entries of each row subtract their respective row minimum entry. We have

$$C^{(2)} = \left[c_{ij}^{(2)}\right]_{3\times 3} = \begin{bmatrix} 0.1334 & 0 & 0 \\ 0 & 0.2089 & 0.2565 \\ 0 & 0.2041 & 0.1652 \end{bmatrix}.$$

Now, there do not appear three zero entries located at different rows and different columns in $C^{(2)}$, and the minimum entries of all the rows and columns are zero. In this case, we can proceed according to the following steps.

Step 1: Start with the row or the column in $C^{(2)}$ with the least number of zero entries to select a zero entry as independent zero, and use "()" to embrace it;

Step 2: Draw a "$\surd$" on the right side of each row in $C^{(2)}$ that does not contain an independent zero;

Step 3: Draw a "$\surd$" underneath each column of the zero entries in a row with a check "$\surd$";

Step 4: Draw a "$\surd$" on the right side of each row with independent zeros, located in a column with a check "$\surd$";

Step 5: Draw a horizontal line over the rows without a check "$\surd$", and a vertical line over the columns with a check "$\surd$";

Step 6: Find the minimum entry among all the entries that have not been lined out;

Step 7: Subtract this minimum entry from all the entries that have not been lined out, add this minimum entry to all the entries that have been lined twice, and keep the rest of the entries unchanged. Then, we obtain a matrix $C^{(3)}$;

Step 8: If the number of zero entries, located at different rows and different columns, equals the rank of $C^{(3)}$, stop the computation. Otherwise start from Step 1.

As for the matrix in our example,

$$C^{(2)} = \begin{bmatrix} 0.1334 & (0) & 0 \\ (0) & 0.2089 & 0.2565 \\ 0 & 0.2041 & 0.1652 \end{bmatrix} \begin{matrix} \\ \surd \\ \surd \end{matrix}$$
$$\begin{matrix} \surd & & \end{matrix}$$

we take $c_{12}^{(2)}$ and $c_{21}^{(2)}$ as independent zeros, and put a check "$\surd$" on the right side of the third row. Because there is a zero entry at the first column of the third row, we put a check "$\surd$" underneath the first column. Because there is an independent zero at the location (2,1), we put a check "$\surd$" on the right side of the second row. Draw a horizontal line across the first row that does not have a check "$\surd$" and a vertical line across the first column that has a check "$\surd$". Now, the remaining entries, which have not been lined out, are $c_{22}^{(2)}$, $c_{23}^{(2)}$, $c_{32}^{(2)}$, and $c_{33}^{(2)}$. Because

$$\min\{c_{22}^{(2)}, c_{23}^{(2)}, c_{32}^{(2)}, c_{33}^{(2)}\} = c_{33}^{(2)} = 0.1652,$$

we subtract $c_{33}^{(2)}$ from each of the remaining entries $c_{22}^{(2)}$, $c_{23}^{(2)}$, $c_{32}^{(2)}$, and $c_{33}^{(2)}$, add $c_{33}^{(2)}$ to the entry $c_{11}^{(2)}$ which was lined across twice, and keep the other entries unchanged. We obtain that

$$C^{(3)} = \left[c_{ij}^{(3)}\right]_{3\times 3} = \begin{bmatrix} c_{11}^{(2)} + c_{33}^{(2)} & c_{12}^{(2)} & c_{13}^{(2)} \\ c_{21}^{(2)} & c_{22}^{(2)} - c_{33}^{(2)} & c_{23}^{(2)} - c_{33}^{(2)} \\ c_{31}^{(2)} & c_{32}^{(2)} - c_{33}^{(2)} & c_{33}^{(2)} - c_{33}^{(2)} \end{bmatrix}$$

$$= \begin{bmatrix} 0.2986 & (0) & 0 \\ (0) & 0.0437 & 0.0913 \\ 0 & 0.0389 & (0) \end{bmatrix}.$$

There already are three zero entries in $C^{(3)}$ located at different rows and different columns. The corresponding optimal solution is

$$\begin{aligned} X &= (x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}, x_{31}, x_{32}, x_{33}) \\ &= (0, 1, 0, 1, 0, 0, 0, 0, 1), \end{aligned}$$

which agrees with the solution of prediction type.

11.8 Grey Nonlinear Programming Without Constraints

Now, in this section, we turn our attention to nonlinear programming problems.

Definition 11.8.1. *Assume that*

$$X = (x_1, x_2, ..., x_n)$$

is a decision vector, and $\otimes$ *a set of grey parameters. Then,*

$$\max(\min) S = f(X, \otimes)$$

is called a grey nonlinear programming *problem without constraints, where* $f(X, \otimes)$ *is a grey price or consumption functional.*

Definition 11.8.2. *Whitenizing all grey elements in* $f(X, \otimes)$ *results in a programming problem called a whitenized programming of*

$$\max(\min) S = f(X, \otimes)$$

which is denoted

$$\max(\min) S = f(X).$$

That is, for grey nonlinear programming problems, one can first whitenize it and then solve for the solution.

Definition 11.8.3. *Assume that $f(X)$ is a differentiable function, where $X = (x_1, x_2, ..., x_n)$. Then, the solution of the gradient vector*

$$\operatorname{grad} f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n} \right) = 0$$

is called a stationary point of $f(X)$.

Theorem 11.8.1. *Assume that $f(X)$ is second-order differentiable, and its Hesse matrix is*

$$H(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

If X^0 is a fixed point of $f(X)$, then

1. *When $H(X^0)$ is a positive definite matrix, X^0 is a minimum point.*
2. *When $H(X^0)$ is a negative definite matrix, X^0 is a maximum point.*
3. *When $H(X^0)$ is semi-positive definite, if there exists a neighborhood $U(X^0, \delta)$ of X^0 such that for any $X \in U(X^0, \delta)$, $H(X^0)$ is semi-positive definite, then X^0 is a minimum point.*
4. *When $H(X^0)$ is semi-negative definite, if there exists a neighborhood $U(X^0, \delta)$ of X^0 such that for any $X \in U(X^0, \delta)$, $H(X^0)$ is semi-negative, then X^0 is a maximum point.*
5. *When $H(X^0)$ is a non-definite matrix, X^0 is not an extreme point of the functional $f(X)$.*

Example 11.8.1. Solve the grey nonlinear programming:

$$\begin{aligned} \max S &= f(X, \otimes) \\ &= \otimes_1 x_1 + \otimes_2 x_3 + \otimes_3 x_2 x_3 - x_1^2 - x_2^2 + \otimes_4 x_3^2, \end{aligned}$$

where

$$\otimes_1 \in [0, 2], \otimes_2 \in [1.5, 2.5],$$

$$\otimes_3 \in [0.5, 1.5], \text{ and } \otimes_4 \in [-2, 0].$$

Solution: Step 1: Mean whitenizing the grey elements $\otimes_i, i = 1, 2, 3, 4$, gives that

$$\tilde{\otimes}_1 = 1, \tilde{\otimes}_2 = 2, \tilde{\otimes}_3 = 1, \tilde{\otimes}_4 = -1.$$

So, the whitenized programming problem is

$$\max S = f(X) = x_1 + 2x_3 + x_2x_3 - x_1^2 - x_2^2 - x_3^2.$$

Step 2: Find the gradient vector

$$\begin{aligned} \mathrm{grad} f(X) &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right) \\ &= (1 - 2x_1, x_3 - 2x_2, 2 + x_2 - 2x_3). \end{aligned}$$

Step 3: Let grad $f(X) = 0$. From

$$\begin{cases} 1 - 2x_1 = 0 \\ x_3 - 2x_2 = 0 \\ 2 + x_2 - 2x_3 = 0 \end{cases}$$

we solve and obtain a stationary point

$$X^0 = (x_1^0, x_2^0, x_3^0) = \left(\frac{1}{2}, \frac{2}{3}, \frac{4}{3} \right).$$

Step 4: Determine the Hesse matrix $H(X)$. From

$$\frac{\partial^2 f}{\partial x_1^2} = -2, \ \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0, \ \frac{\partial^2 f}{\partial x_1 \partial x_3} = 0,$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = 0, \ \frac{\partial^2 f}{\partial x_2^2} = -2, \ \frac{\partial^2 f}{\partial x_2 \partial x_3} = 1,$$

and

$$\frac{\partial^2 f}{\partial x_3 \partial x_1} = 0, \ \frac{\partial^2 f}{\partial x_3 \partial x_2} = 1, \ \frac{\partial^2 f}{\partial x_3^2} = -2,$$

it follows that

$$H(X) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}.$$

Step 5: Substituting the stationary point

$$X^0 = (x_1^0, x_2^0, x_3^0) = \left(\frac{1}{2}, \frac{2}{3}, \frac{4}{3} \right)$$

into $H(X)$ gives

$$H(X^0) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}.$$

Step 6: Compute various leading principal minors (LPM) of $H(X^0)$. We have

$$\begin{aligned} &\text{the 1st-order LPM} && = -2 < 0, \\ &\text{the 2nd-order LPM} && = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0, \\ &\text{the 3rd-order LPM} && = \left|H(X^0)\right| = -6 < 0. \end{aligned}$$

Because the odd-order leading principal minor is < 0, and the even-order leading principal minors are > 0, $H(X^0)$ is a negative definite matrix and X^0 a maximum point.

Step 7: Compute the maximum value.

$$\max S = \frac{1}{2} + 2 \cdot \frac{4}{3} + \frac{2}{3} \cdot \frac{4}{3} - \left(\frac{1}{2}\right)^2 - \left(\frac{2}{3}\right)^2 - \left(\frac{4}{3}\right)^2 = \frac{19}{12}.$$

11.9 Grey Nonlinear Programming with Constraints

In this section, we discuss problems of grey nonlinear programming with constraints.

Definition 11.9.1. *Assume that*

$$X = (x_1, x_2, ..., x_n)$$

is a decision vector and $\otimes^{(1)}, \otimes^{(j)}, \otimes^{(i)}$, $j = 1, 2, ..., m$, $i = 1, 2, ..., s$, *are sets of grey parameters. Then,*

$$\min S = f(X, \otimes^{(1)})$$

$$s.t. \begin{cases} g_j(X, \otimes^{(j)}) \geq 0, \ j \in J = \{1, 2, ..., m\} \\ h_i(X, \otimes^{(i)}) = 0, \ i \in I = \{1, 2, ..., s\} \end{cases}$$

is called a problem of nonlinear programming with constraints, where $f(X, \otimes^{(1)})$ *is the grey consumption functional, and* $g_j(X, \otimes^{(j)})$ *and* $h_i(X, \otimes^{(i)})$ *are, respectively, grey constraint functionals.*

The whitenized programming problem

$$\min S = f(X)$$

$$s.t. \begin{cases} g_j(X) \geq 0, \ j \in J = \{1, 2, ..., m\} \\ h_i(X) = 0, \ i \in I = \{1, 2, ..., s\} \end{cases}$$

of the grey nonlinear programming problem can be solved by following the steps below.

Step 1: If $I \neq \emptyset$, go to Step 2 directly. When $I = \emptyset$, that is, when there does not exist any equation constraint, solve the programming problem as one without constraints

$$\min S = f(X).$$

Assume that $X^{(0)}$ is the optimal solution. Let $k = 0$, go to Step 3.

Step 2: Solve the sub-programming problem (Lagrange multipliers can be applied):

$$\min S = f(X)$$

$$s.t. \{h_i(X) = 0, \ i \in I = \{1, 2, \cdots, s\}$$

with equation constraints. Assume that $X^{(0)}$ is the optimal solution. Let $k = 0$.

Step 3: Substitute $X^{(k)}$ into the inequality constraints and compute the index set of inequalities that are not satisfied:

$$J_k = \left\{ p | g_p(X^{(k)}) < 0, \ p \in J = \{1, 2, ..., m\} \right\}.$$

When $J_k = \emptyset$, $X^{(k)}$ is the optimal solution of the original programming problem. So, stop the calculation. If $J_k \neq \emptyset$, go to the next step.

Step 4: Select an arbitrary element p from J_k, introduce non-negative slack variable y_p^2, and change the corresponding inequality constraints into equation constraints. Now, solve the following programming problem with augmented equation constraints,

$$\min S = f(X)$$

$$s.t. \begin{cases} g_p(X) - y_p^2 = 0, \ p \in J_k \\ h_i(X) = 0, \ i \in I = \{1, 2, ..., s\}. \end{cases}$$

Assume that $X^{(k+1)}(y_p^2)$ is the optimal solution.

Step 5: Solve the following programming problem without constraints,

$$\min S = f\left[X^{(k+1)}(y_p^2)\right].$$

Assume that $X^{(k+1)}$ is the optimal solution. Replace $k+1$ with k and go back to Step 3.

Example 11.9.1. Solve the following grey nonlinear programming problem,

$$\min S = f(X, \otimes^{(1)}) = \otimes_1 x_1 + \otimes_2 x_2^2$$

$$s.t. \begin{cases} g_1(X, \otimes^{(2)}) = \otimes_3 - (x_1 - 4)^2 + \otimes_4 x_2^2 \geq 0 \\ h_1(X, \otimes^{(3)}) = (x_1 - 3)^2 + (x_2 - 2)^2 + \otimes_5 = 0, \end{cases}$$

where

$$\otimes_1 \in [0.8, 1.2], \otimes_2 \in [-0.15, 0.15], \otimes_3 \in [12, 17],$$

$$\otimes_4 \in [-1.4, -0.9], \text{ and } \otimes_5 \in [9, 14].$$

Solution: Step 1: Mean whitenizing the grey elements in the objective functional, and 0.8-positioning whitenizing the grey elements in the constraint functional give that

$$\widetilde{\otimes}_1 = 1, \widetilde{\otimes}_2 = 0, \widetilde{\otimes}_3 = 16,$$

$$\widetilde{\otimes}_4 = -1, \widetilde{\otimes}_5 = 13.$$

Therefore, we have the following whitenized programming problem,

$$\min S = f(X) = x_1$$

$$s.t. \begin{cases} g_1(X) = 16 - (x_1 - 4)^2 - x_2^2 \geq 0 \\ h_1(X) = (x_1 - 3)^2 + (x_2 - 2)^2 - 13 = 0. \end{cases}$$

Step 2: Solve the following sub-programming problem with an equation constraint

$$\min S = x_1$$

$$s.t. \left\{ (x_1 - 3)^2 + (x_2 - 2)^2 - 13 = 0, \right.$$

and obtain the optimal solution

$$X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}\right) = \left(3 - \sqrt{13}, 2\right).$$

Step 3: Substituting

$$X^{(0)} = \left(3 - \sqrt{13}, 2\right)$$

into the inequality constraint gives that

$$g_1(X^{(0)}) = -2(1 + \sqrt{13}) < 0,$$

which shows that the inequality constraint is not satisfied.

Step 4: Introduce non-negative slack variable y^2, and solve the following sub-programming problem with augmented equation constraints:

$$\min S = x_1$$

$$s.t. \begin{cases} (x_1 - 3)^2 + (x_2 - 2)^2 - 13 = 0 \\ (x_1 - 4)^2 + x_2^2 - 16 - y^2 = 0. \end{cases}$$

and obtain the following optimal solution

$$\begin{aligned} X^{(1)}(y^2) &= \left(x_1^{(1)}, x_2^{(1)}\right) \\ &= \left(\frac{2}{5}(8 \pm \sqrt{64 - 5y^2}), \frac{1}{5}(8 \pm \sqrt{64 - 5y^2})\right). \end{aligned}$$

Step 5: Solving the following sub-programming problem without constraints,

$$\begin{aligned} \min S &= f\left(X^{(1)}(y^2)\right) = x_1^{(1)} \\ &= \frac{2}{5}(8 \pm \sqrt{64 - 5y^2}), \end{aligned}$$

gives us $y = 0$. So, the corresponding optimal solution is

$$X^{(2)} = \left(x_1^{(2)}, x_2^{(2)}\right) = (0, 0).$$

Step 6: Substitute $X^{(2)}$ into the inequality constraint. It is found that the inequality is satisfied and $J_2 = \emptyset$, so stop the calculation. Now,

$$X^{(2)} = (0, 0)$$

is the optimal solution of the original whitenized programming with the optimal value

$$\min S = f(X^{(2)}) = 0.$$

If we solve the grey nonlinear programming problem directly without first whitenizing the problem, then in general the optimal solution contains grey parameters. We can study the changes of the optimal solution and the optimal value based on the fields where the grey parameters take values.

Example 11.9.2. Solve the following grey nonlinear programming problem,

$$\min S = f(X, \otimes^{(1)}) = x_1^2 + \otimes_1 x_2$$

$$s.t. \begin{cases} g_1(X) = 1 - (x_1 + x_2^2) \geq 0 \\ g_2(X, \otimes^{(2)}) = \otimes_2 - x_1 - x_2 \geq 0 \\ h_1(X, \otimes^{(3)}) = x_1^2 + x_2^2 - \otimes_3^2 = 0, \end{cases}$$

where

$$\otimes_1 \in [0.9, 1.2], \otimes_2 \in [0.8, 1.4], \text{ and } \otimes_3 \in [2.5, 3.5].$$

Solution: Step 1: Solve the following sub-programming problem with an equation constraint,

$$\min S = x_1^2 + \otimes_1 x_2$$

$$s.t.\ x_1^2 + x_2^2 - \otimes_3^2 = 0.$$

We obtain the optimal solution

$$X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}\right) = (0, -\otimes_3)\,.$$

Step 2: Substituting

$$X^{(0)} = (0, -\otimes_3)$$

into the inequality constraints gives that

$$g_1(X^{(0)}) = 1 - \otimes_3^2.$$

Because

$$\tilde{\otimes}_3 \geq 2.5, \tilde{\otimes}_3^2 \geq 6.25,$$

we have

$$g_1(X^{(0)}) \leq 1 - 6.25 = -5.25 < 0.$$

That is,

$$g_1(X^{(0)}) \geq 0$$

is not satisfied. Similar discussion tells us that

$$g_2(X^{(0)}, \otimes^{(2)}) \geq 0$$

is satisfied. Therefore, $J_0 = \{1\}$.

Step 3: Introduce non-negative slack variable y_1^2, and solve the following sub-programming problem with augmented equation constraints.

$$\min S = x_1^2 + \otimes_1 x_2$$

$$s.t. \begin{cases} g_1(X) - y_1^2 = 1 - (x_1 + x_2^2) - y_1^2 = 0 \\ h_1(X, \otimes^{(3)}) = x_1^2 + x_2^2 - \otimes_3^2 = 0, \end{cases}$$

and obtain the following optimal solution

$$\begin{aligned} X^{(1)}(y_1^2) &= = \left(x_1^{(1)}, x_2^{(1)}\right) \\ &= = \left(\frac{1}{2}[1 - \sqrt{4y_1^2 + 4 \otimes_3^2 - 3}],\right. \\ &\quad \left. - \frac{\sqrt{2}}{2}\sqrt{1 - 2y_1^2 + \sqrt{4y_1^2 + 4 \otimes_3^2 - 3}}\right). \end{aligned}$$

Step 4: Solving the following sub-programming problem without constraints,

$$\begin{aligned} \min S &= f(X^{(1)}(y_1^2)) = [x_1^{(1)}]^2 + \otimes_1 x_2^{(1)} \\ &= [\frac{1}{2}[1 - \sqrt{4y_1^2 + 4 \otimes_3^2 - 3}]]^2 \\ &\quad - \otimes_1 \cdot \frac{\sqrt{2}}{2}\sqrt{1 - 2y_1^2 + \sqrt{4y_1^2 + 4 \otimes_3^2 - 3}}, \end{aligned}$$

gives that $y_1^2 = 0$, and the corresponding optimal solution is

$$\begin{aligned} X^{(2)} &= \left(x_1^{(2)}, x_2^{(2)}\right) \\ &= \left(\frac{1}{2}[1 - \sqrt{4 \otimes_3^2 - 3}], -\frac{\sqrt{2}}{2}\sqrt{1 + \sqrt{4 \otimes_3^2 - 3}}\right). \end{aligned}$$

Step 5: Substituting $X^{(2)}$ into the inequality constraints provides

$$\begin{aligned} g_1(X^{(2)}, \otimes^{(2)}) &= 1 - \frac{1}{2}[1 - \sqrt{4 \otimes_3^2 - 3}] \\ &\quad -\frac{1}{2}[1 + \sqrt{4 \otimes_3^2 - 3}] = 0 \end{aligned}$$

and

$$\begin{aligned} g_2(X^{(2)}, \otimes^{(2)}) &= \otimes_2 - \frac{1}{2}[1 - \sqrt{4 \otimes_3^2 - 3}] \\ &\quad +\frac{\sqrt{2}}{2}\sqrt{1 + \sqrt{4 \otimes_3^2 - 3}}. \end{aligned}$$

From the fields of $\otimes_2$ and $\otimes$, it follows that

$$g_2(X^{(2)}, \otimes^{(2)}) \geq 0.$$

That is, all the inequality constraints are satisfied. So, $J_2 = \emptyset$, and we stop the calculation. Now,

$$X^{(2)} = \left(\frac{1}{2}[1 - \sqrt{4\otimes_3^2 - 3}], -\frac{\sqrt{2}}{2}\sqrt{1 + \sqrt{4\otimes_3^2 - 3}}\right)$$

is the optimal solution of the original grey nonlinear programming, and the corresponding optimal value is

$$\begin{aligned}\min S &= f(X^{(2)}, \otimes^{(1)}) = [x_1^{(2)}]^2 + \otimes_1 x_2^{(2)} \\ &= \frac{1}{4}\left[1 - \sqrt{4\otimes_3^2 - 3}\right]^2 - \otimes_1 \cdot \frac{\sqrt{2}}{2}\sqrt{1 + \sqrt{4\otimes_3^2 - 3}}.\end{aligned}$$

When $\otimes_1$ takes the upper limit and $\otimes_3$ takes the lower limit, we can obtain the ideal optimal value

$$\min \overline{S} \approx 1.381.$$

12
Grey Input and Output

In each macroeconomic analysis, most of all the information contained in statistical data, typical investigations, laboratory reports, etc., are grey quantities. Therefore, what is listed in an input-output table inevitably contains various kinds of grey numbers. Besides, as a dynamic economic system, various parameters contained in each input-output problem must be in a state of constant change. So, these parameters are also grey numbers with possibly an upper and a lower limit. In this chapter, we learn how we can combine the methodology of grey systems theory with the study of input-output systems so that problems of grey input-output can be introduced and studied thoroughly. The main contents here in this chapter include basic concepts used in grey input and output, the basic theory, grey input-output optimal models, grey industrial incidence coefficients, grey dynamic input-output analysis, grey von Neumann model, etc.

12.1 Basic Concepts for Grey Input and Output

In this section, we first look at some of the basic concepts studied in grey input-output analysis.

Definition 12.1.1. *Assume that* $x_{ij}, i, j = 1, 2, ..., n$, *stands for the total value of the products of the* i*th department consumed by the* j*th depart-*

ment. Then

$$Q = [x_{ij}]_{n \times n} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & & \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix}$$

is called the flow matrix.

Definition 12.1.2. *Assume that $x_{ij}, i, j = 1, 2, ..., n$, is the same as above, and x_j the total output of the jth department. Then*

$$a_{ij} = \frac{x_{ij}}{x_j}, \ i, j = 1, 2, ..., n,$$

is called the direct consumption coefficient.

The concept of direct consumption coefficients a_{ij} reflects the amount of consumption of the products produced by the ith production department consumed by the jth production unit. It represents the degree of dependence of the jth department on the ith production unit. The larger the a_{ij} value, the more closely the jth department is dependent on the ith production unit.

Definition 12.1.3. *The matrix*

$$A = [a_{ij}]_{n \times n}$$

is called the direct consumption coefficient matrix.

Proposition 12.1.1. *Each entry a_{ij} in A satisfies $a_{ij} \geq 0$, $i, j = 1, 2, ..., n$.*

Proof. Because the total output of products of the jth department satisfies $x_j > 0$, the total cost spent by the jth department on the products of the ith production unit satisfies $x_{ij} \geq 0$. Therefore,

$$a_{ij} \geq 0, i, j = 1, 2, ..., n. \ \square$$

Proposition 12.1.2. *The sum of the entries in any row of A is less than 1. That is, $\sum_{i=1}^{n} a_{ij} < 1$, j =1, 2, ..., n.*

Proof. We prove this result by contradiction. Assume that there is a k such that $\sum_{i=1}^{n} a_{ik} \geq 1$. From $a_{ik} = a_{ik}/x_k$, it follows that $\sum_{i=1}^{n} x_{ik} \geq x_k$. That is, the total output x_k of the kth department is less than or equal to the total cost this department spent on products of all departments. Therefore, there is no way for the kth department to have its own production activities going on. Hence, it is impossible to have that $\sum_{i=1}^{n} a_{ij} \geq 1$. From the arbitrariness of k, it follows that $\sum_{i=1}^{n} a_{ij} < 1$ holds true for $j = 1, 2, ..., n$. $\square$

Because of difficulties of obtaining information, as a matter of fact, the total amount of consumption of the jth department of all products produced by the ith department is a grey number $x_{ij}(\otimes)$, $i, j = 1, 2, ..., n$.

Definition 12.1.4. *The following*

$$C = (E - A)^{-1} - E$$

is called a complete consumption coefficient matrix, where E stands for the identity (unit) matrix of the size of A.

Definition 12.1.5. *The following*

$$Q(\otimes) = [x_{ij}(\otimes)]_{n \times n}$$

is called a grey flow matrix.

When the amount of flow is a grey number, obviously the direct consumption coefficient

$$a_{ij}(\otimes) = \frac{x_{ij}(\otimes)}{x_j}, i, j = 1, 2, ..., n$$

is also a grey number.

Definition 12.1.6. *The following*

$$A(\otimes) = [a_{ij}(\otimes)]_{n \times n}$$

is called a grey direct consumption coefficient matrix.

Proposition 12.1.3. *Assume that $X = [x_1, x_2, ..., x_n]^T$ stands for the total output vector, $Y = [y_1, y_2, ..., y_n]^T$ the ultimate product vector, $S = [s_1, s_2, ..., s_n]^T$ the newly created value vector, $P = [p_1, p_2, ..., p_n]^T$ the price vector, and $A(\otimes) = [a_{ij}(\otimes)]_{n \times n}$ the grey direct consumption coefficient matrix. Then, the following hold true.*

1. $X = [E - A(\otimes)]^{-1} Y$; *and*
2. $P = \left[E - A^T(\otimes)\right]^{-1} S$.

Proposition 12.1.4. *The statements 1 and 2 in the previous proposition can be written in equation forms as follows.*

1. $\sum_{j=1}^{n} a_{ij}(\otimes) x_j + y_i = x_i, i = 1, 2, ..., n$. *And*
2. $\sum_{i=1}^{n} a_{ij}(\otimes) p_i + s_j = p_j, j = 1, 2, ..., n$.

Proposition 12.1.5. *The statements 1^o and 2^o in the previous proposition can be simplified as follows.*

1. $\sum_{j=1}^{n} x_{ij}(\otimes) + y_i = x_i, i = 1, 2, ..., n$. *And*
2. $\sum_{i=1}^{n} x_{ij}(\otimes) + s_j = p_j, j = 1, 2, ..., n$.

The first statement of Proposition 12.1.5 indicates that the total output of the ith department equal the sum of all the products of the ith department consumed by all departments and the ultimate products of the ith department. This set of equations is normally known as the distribution system. The second statement of Proposition 12.1.5 indicates that the total values created by the jth department equals the sum of the total consumption of all departments and the newly created value of the jth department. This set of equations is often known as the production system of equations.

Definition 12.1.7. *The following*

$$C(\otimes) = [E - A(\otimes)]^{-1} - E$$

is called a grey complete consumption coefficient matrix.

The grey input-output model, as described above, reflects the grey relationships between various production departments, and ultimate products and the total production, price, and consumption, and between newly created values. It is expected to be the foundation for studies on the structure of production systems and for analysis of development mechanisms of economic systems.

12.2 P–F Theorems of Grey Non-Negative Matrices

The so-called grey flow matrix and grey direct consumption coefficient matrix, as studied in the previous section, are all non-negative matrices. Therefore, the study on the spectral radii and characteristic values of grey non-negative matrices becomes the theoretical foundation for any attempt to solve grey input-output models. In this section, we provide a proof for Perron–Frobenius (P–F) theorem of grey non-negative matrices.

Definition 12.2.1. *Assume that a grey element* $\otimes \in [\underline{a}, \overline{a}], \underline{a} < \overline{a}$, *is given. If*

1. $\otimes$ *is a continuous grey number, then*

$$\widehat{a} = \frac{1}{2}(\underline{a} + \overline{a})$$

is called the mean-value whitenization (number) of the grey element $\otimes$;

2. $\otimes$ *is a discrete grey number and*

$$a_i \in [\underline{a}, \overline{a}], i = 1, 2, ..., n,$$

are the values the grey element $\otimes$ *takes, then*

$$\widehat{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$$

is called the mean-value whitenization (number) of the grey element $\otimes$. *(Note: If* $a_k(\otimes)$ *is a grey element such that* $a_k(\otimes) \in [\underline{a}_k, \overline{a}_k], \underline{a}_k < \overline{a}_k$, *let* $a_k = \widehat{a}_k$.*)*

If we write a grey number $\otimes$ as $\otimes = \widehat{a} + \delta$, then δ is known as the perturbation grey element of $\widehat{a}$.

Definition 12.2.2. *Assume that a grey matrix* $A(\otimes) = (\otimes_{ij}) \in G^{n\times n}$ *satisfies that each grey element* $\otimes_{ij} = \widehat{a}_{ij} + \delta_{ij} \in \left[\underline{a}_{ij}, \overline{a}_{ij}\right], \underline{a}_{ij} < \overline{a}_{ij}$, *where* $\widehat{a}_{ij}$ *stands for the mean-value whitenization of the grey element* $\otimes_{ij}$, *and* δ_{ij} *the perturbation grey element of* $\otimes_{ij}$ *on the basis of* $\widehat{a}_{ij}$. *Then, accordingly, for the matrix representation of*

$$A(\otimes) = \widehat{A} + A_\delta$$

the matrices

$$\widehat{A} = [\widehat{a}_{ij}]_{n\times n} \text{ and } A_\delta = [\delta_{ij}]_{n\times n}$$

are, respectively, called the mean-value matrix of the grey matrix $A(\otimes)$ *and perturbation grey matrix of* $A(\otimes)$ *on the basis of* $\widehat{A}$.

Definition 12.2.3. *For* $A(\otimes) \in G^{n\times n}$, *if the mean-value matrix* $\widehat{A} \geq 0$, *then* $A(\otimes)$ *is called a grey non-negative matrix.*

Definition 12.2.4. *For* $A(\otimes) \in G^{n\times n}$, *assume that* $\lambda_i(\otimes) = \widehat{\lambda}_i + \delta_i$, $i = 1, 2, ..., n$, *is a characteristic value of* $A(\otimes)$ *and* $\max\left\{\widehat{\lambda}_i\right\} = \widehat{\lambda}_k$. *Then* $\rho(A(\otimes)) = \widehat{\lambda}_k + \delta_k$ *is called the spectral radius of* $A(\otimes)$.

Obviously, in general, the spectral radius of a grey matrix is also a grey element.

Proposition 12.2.1. *If* $A(\otimes) \in G^{n\times n}$, *then* $\widehat{\rho}(A(\otimes)) = \rho\left(\widehat{A}\right)$. *That is, the mean-value whitenization of the spectral radius of the grey matrix* $A(\otimes)$ *equals the spectral radius of the mean-value matrix of* $A(\otimes)$.

Definition 12.2.5. *For* $A(\otimes) \in G^{n\times n}$, *if the mean-value matrix* $\widehat{A} = (\widehat{a}_{ij})$ *of* $A(\otimes)$ *satisfies the following conditions,*

1. $\widehat{a}_{ij} \leq 0$, $i \neq j$; *and*
2. $\widehat{A}^{-1}$*exists, and* $\widehat{A}^{-1} \geq 0$,

then $A(\otimes)$ *is called a grey M matrix.*

Proposition 12.2.2. *If* $A(\otimes) \in G^{n\times n}$ *satisfies that the mean-value matrix* $\widehat{A} \geq 0$, *then* $E - A(\otimes)$ *is a grey M matrix, if and only if* $\rho\left(\widehat{A}\right) < 1$.

Definition 12.2.6. *If* $A(\otimes) = \widehat{A} + A_\delta \in G^{n\times n}$ *satisfies that* $\det \widehat{A}(i_1, i_2, ..., i_k) > 0$, $k = 1, 2, ..., n$, *then* $A(\otimes)$ *is called a grey P matrix.*

Proposition 12.2.3. *If* $A(\otimes) \in G^{n\times n}$ *satisfies that* $\widehat{a}_{ij} \leq 0, i \neq j$, *then* $A(\otimes)$ *is a grey M matrix, if and only if* $A(\otimes)$ *is a grey P matrix.*

Lemma 12.2.1. *Assume that* $A(\otimes)$ *is a grey non-negative matrix such that its mean-value matrix* $\widehat{A}$ *is irreducible. Then* $A(\otimes)$ *has a grey characteristic value* $\lambda^*(A(\otimes)) = \lambda^*\left(\widehat{A}\right) + \delta$, *where* $\lambda^*\left(\widehat{A}\right) > 0$.

Proof. It suffices to prove that the mean-value matrix $\widehat{A}$ has a positive characteristic value. Consider the following set

$$S = \left\{ X \middle| \ X \geq 0, \sum_{i=1}^{n} X_i = 1 \right\}.$$

Then, S is a compact convex set in the n-dimensional Euclidean space. Define a mapping

$$f : S \to S, f(X) \rightarrowtail \frac{\widehat{A}X}{\left\| \widehat{A}X \right\|}.$$

From the assumption of non-negativeness of $A(\otimes)$, it follows that $\left\| \widehat{A}X \right\| > 0$. Therefore, $f : S \to S$ is a continuous function. From Brouwer's fixed point theorem it follows that there is at least one fixed point for f in the set S. That is, there exists $X^* \in S$ such that $f(X^*) = X^*$. From this fact, it follows that

$$\frac{\widehat{A}X^*}{\left\| \widehat{A}X \right\|} = X^*.$$

If we let $\lambda^* \left(\widehat{A} \right) = \left\| \widehat{A}X \right\|$, then we have

$$\widehat{A}X^* = \lambda^* \left(\widehat{A} \right) X^* \text{ and } \lambda^* \left(\widehat{A} \right) > 0. \ \square$$

Lemma 12.2.2. *Assume that $A(\otimes)$ is a grey non-negative matrix such that its mean-value matrix $\widehat{A}$ is irreducible. Then any kth order $(k < n)$ principal submatrix of $\rho(A(\otimes))E - A(\otimes)$ is a grey P matrix.*

Proof. First let us prove that any kth order principal submatrix of

$$\rho(A(\otimes))E - A(\otimes)$$

is a grey M matrix. Assume that $\widehat{A}_k$ is an arbitrary principal submatrix of $\widehat{A}$. Construct a matrix B as follows,

$$B = \begin{bmatrix} \widehat{A}_k & 0 \\ 0 & 0 \end{bmatrix}_{n \times n}.$$

Evidently, we have $\widehat{A} \geq B \geq 0$, $\rho(B) = \rho\left(\widehat{A}_k\right)$ and $\rho\left(\widehat{A}\right) > \rho(B)$. Therefore, the following principal submatrix of $\rho\left(\widehat{A}\right)E - \widehat{A}$ satisfies

$$\rho\left(\widehat{A}\right)E_k - \widehat{A}_k = \rho\left(\widehat{A}\right)\left[E_k - \frac{\widehat{A}_k}{\rho\left(\widehat{A}\right)}\right]$$

and

$$\rho\left(\frac{\widehat{A}_k}{\rho\left(\widehat{A}\right)}\right) = \frac{\rho\left(\widehat{A}_k\right)}{\rho\left(\widehat{A}\right)} < 1.$$

From Proposition 12.2.2, it follows that $\rho\left(\widehat{A}\right) E_k - \widehat{A}_k(\otimes)$ is a grey M matrix. Now, Proposition 12.2.3 implies that $\rho\left(\widehat{A}\right) E_k - \widehat{A}_k(\otimes)$ is a grey P matrix. □

Theorem 12.2.1. *(P–F Theorem 1) Assume that $A(\otimes)$ is a grey non-negative matrix such that its mean-value matrix $\widehat{A}$ is irreducible. Then, the following conclusions are true.*

1. $A(\otimes)$ has a grey characteristic value $\lambda^*(A(\otimes))$ whose mean-value whitenization $\lambda^* > 0$.

2. If $X^*(\otimes)$ is a grey characteristic vector corresponding to $\lambda^*(A(\otimes))$ and $X^*(\otimes) = \widehat{X}^* + X^*_\delta$ then its mean-value vector satisfies $\widehat{X}^* > 0$.

3. $\lambda^*(A(\otimes)) = \rho(A(\otimes))$.

4. $\lambda^*(A(\otimes))$ flows in the same direction as the entry $\otimes_{ij}$ of $A(\otimes)$. And

5. $\lambda^*(A(\otimes))$ is a characteristic value of $A(\otimes)$ of multiplicity 1.

Proof. 1. is a direct consequence of Lemma 12.2.1 and Proposition 12.2.1.

2. Let us treat the set S in Lemma 12.2.1 as the set of all mean-value vectors of the characteristic vectors $X(\otimes)$. From $\widehat{X}^* \in S$, it follows that $\widehat{X}^* \geq 0$.

Assume that $\widehat{X}^* > 0$ does not hold true. Without loss of generality, let us assume that $\widehat{X}^* = \left[\widehat{X}_1^*, \widehat{X}_2^*\right]^T$, where $\widehat{X}_1^* > 0$ and $\widehat{X}_2^* = 0$. From $\widehat{A}\widehat{X}^* = \lambda\left(\widehat{A}\right)\widehat{X}^*$, it can be seen that we can divide $\widehat{A}$ into the following

$$\begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ \widehat{A}_{21} & \widehat{A}_{22} \end{bmatrix} \begin{bmatrix} \widehat{X}_1^* \\ \widehat{X}_2^* \end{bmatrix} = \lambda\left(\widehat{A}\right) \begin{bmatrix} \widehat{X}_1^* \\ \widehat{X}_2^* \end{bmatrix}.$$

Therefore, it follows that $\widehat{A}_{21}\widehat{X}_1^* = 0$. However, $\widehat{A}_{21} \geq 0$ and $\widehat{X}_1^* > 0$. Hence, $\widehat{A}_{21} = 0$. However, this end contradicts the assumption that $\widehat{A}$ is irreducible. This end contradiction implies that we must have $\widehat{X}^* > 0$.

3. Assume that $\lambda(\otimes)$ is an arbitrary characteristic value of $A(\otimes)$ and $X(\otimes)$ the corresponding characteristic vector satisfying $\lambda(\otimes)X(\otimes) = A(\otimes)X(\otimes)$. When written in the form of components, we have

$$\lambda(\otimes)X_i(\otimes) = \sum_{j=1}^{n} \otimes_{ij}X_j(\otimes), i = 1, 2, ..., n.$$

Taking mean-value whitenization on both sides of this equation leads to

$$\widehat{\lambda}\widehat{X}_i = \sum_{j=1}^{n} \widehat{a}_{ij}\widehat{X}_j, i = 1, 2, ..., n.$$

By taking absolute values, we have

$$\left|\widehat{\lambda}\right| \left|\widehat{X}_i\right| \leq \sum_{j=1}^{n} \left|\widehat{a}_{ij}\right| \left|\widehat{X}_j\right| = \sum_{j=1}^{n} \widehat{a}_{ij} \left|\widehat{X}_j\right|, i = 1, 2, ..., n.$$

Denoting $\widehat{X}_a = \left(\left|\widehat{X}_1\right|, \left|\widehat{X}_2\right|, ..., \left|\widehat{X}_n\right|\right)^T \geq 0$ leads to

$$\left|\widehat{\lambda}\right| \widehat{X}_a \leq \widehat{A}\widehat{X}_a \tag{12.1}$$

From the assumptions that $A(\otimes)$ is non-negative and that $\widehat{A}$ is irreducible and by noticing that $\lambda^*\left(\widehat{A}^T\right) = \lambda^*\left(\widehat{A}\right)$, it follows that

$$\widehat{A}^T\widehat{X}^* = \lambda^*\left(\widehat{A}\right)\widehat{X}^*.$$

Applying the inner product by $\widehat{X}^{*T}$ to the two sides of equ. (12.1) provides that

$$\begin{aligned} \widehat{X}^{*T}\left|\widehat{\lambda}\right|\widehat{X}_a &= \left|\widehat{\lambda}\right|\left(\widehat{X}^{*T}\widehat{X}_a\right) \\ &\leq \widehat{X}^{*T}\widehat{A}\widehat{X}_a \\ &= \left(\widehat{A}^T\widehat{X}^*\right)^T\widehat{X}_a \\ &= \lambda^*\left(\widehat{A}\right)\left(\widehat{X}^{*T}\widehat{X}_a\right). \end{aligned}$$

From $\widehat{X}^{*T}\widehat{X}_a > 0$, it follows that $\widehat{\lambda} \leq \lambda^*\left(\widehat{A}\right)$. Because $\lambda(\otimes)$ is arbitrary, we have that $\lambda^*(A(\otimes)) = \lambda^*\left(\widehat{A}\right) + \delta$ is the spectral radius of $A(\otimes)$.

4. It suffices to show that the mean-value whitenization of $\lambda^*\left(\widehat{A}(\otimes)\right)$ is an increasing function of the mean-value whitenization number of $\otimes_{ij}$. Assume that $A(\otimes)$ and $A'(\otimes)$ are two grey non-negative matrices such that their mean-value matrices $\widehat{A}$ and $\widehat{A'}$ are irreducible and $\widehat{A} \leq \widehat{A'}$. From 1 and 2, it follows that the matrices $A(\otimes)$ and $A'(\otimes)$, respectively, have grey characteristic values $\lambda^*(A(\otimes))$ and $\lambda^*(A'(\otimes))$ such that their corresponding characteristic vectors $X^*(\otimes)$ and $X'^*(\otimes)$ satisfy

$$\lambda^*\left(\widehat{A}\right) > 0, \lambda^*\left(\widehat{A'}\right) > 0, \widehat{X}^* > 0, \widehat{X'}^* > 0.$$

Now, from $\widehat{A}\widehat{X}^* \leq \widehat{A'}\widehat{X'}^*$ and $\lambda^*\left(\widehat{A}\right)\widehat{X}^* = \widehat{A}\widehat{X}^*$, it follows that

$$\lambda^*\left(\widehat{A}\right)\widehat{X}^* \leq \widehat{A'}\widehat{X}^* \tag{12.2}$$

By applying the inner product to the two sides of equ. (12.2) using $\widehat{X}^{*T}$ and from the fact that $\lambda^*\left(\widehat{A'}^T\right) = \lambda^*\left(\widehat{A'}\right)$, it follows that

$$\begin{aligned}\lambda^*\left(\widehat{A}\right)\left(\widehat{X}^{*T}\widehat{X}^*\right) &\leq \widehat{X}^{*T}\widehat{A'}\widehat{X}^* \\ &= \left(\widehat{A'}\widehat{X}^*\right)^T\widehat{X}^* \\ &= \lambda^*\left(\widehat{A'}\right)\left(\widehat{X}^{*T}\widehat{X}^*\right).\end{aligned}$$

Therefore, $\widehat{X}^{*T}\widehat{X}^* > 0$. So, $\lambda^*\left(\widehat{A}\right) < \lambda^*\left(\widehat{A'}\right)$.

5. From 3 and Lemma 12.2.2, it follows that $\lambda^*\left(A\left(\otimes\right)\right)E_k - A_k$ is a grey P matrix. Therefore, $\det\left(\lambda^*\left(\widehat{A}\left(\otimes\right)\right)E_k - \widehat{A}_k\left(\otimes\right)\right) > 0$.

Constructing the polynomial $f\left(\lambda\right) = \det\left(\lambda E - \widehat{A}\right)$ and computing its derivative give us

$$f'\left(\lambda\right) = \sum \det\left(\lambda E_k - \widehat{A}_k\right)$$

and

$$f'\left(\widehat{\lambda}^*\right) = \sum \det\left(\widehat{\lambda}^* E_k - \widehat{A}_k\right) > 0.$$

This end shows that $\lambda^*\left(A\left(\otimes\right)\right)$ is a characteristic value of $A\left(\otimes\right)$ of multiplicity 1. □

Theorem 12.2.2. *(P–F Theorem 2) If* $A\left(\otimes\right)$ *is a grey non-negative matrix, then*

1. $A\left(\otimes\right)$ *has a grey characteristic value* $\lambda^*\left(A\left(\otimes\right)\right) = \widehat{\lambda}^* + \delta$ *whose mean-value whitenization satisfies* $\widehat{\lambda}^* \geq 0$.

2. If $X^*\left(\otimes\right)$ *is the characteristic vector corresponding to* $\lambda^*\left(A\left(\otimes\right)\right)$ *and* $X^*\left(\otimes\right) = \widehat{X}^* + X^*_\delta$, *then* $\widehat{X}^* \geq 0$.

3. $\lambda^*\left(A\left(\otimes\right)\right) = \rho\left(A\left(\otimes\right)\right)$.

4. When an element $\otimes_{ij}$ *of* $A\left(\otimes\right)$ *flows,* $\lambda^*\left(A\left(\otimes\right)\right)$ *either stays the same or flows in the same direction.*

Proof. If the mean-value matrix $\widehat{A}$ is irreducible, the results are obvious.

If $\widehat{A}$ is reducible, then $\widehat{A}$ can be written as follows:

$$\widehat{A} = \begin{bmatrix} \widehat{A}_{11} & & & & & \\ & \widehat{A}_{22} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \widehat{A}_{rr} \end{bmatrix}$$

where $\widehat{A}_{ii}$ is either a square 0 matrix or an irreducible square matrix, $i = 1, 2, ..., r$. From $\lambda^* \left(\widehat{A}\right) = \max_{1 \le i \le r} \left\{ \lambda^* \left(\widehat{A}_{ii}\right) \right\}$, statements 1, 3, and 4 follow. In the following, we prove statement 2. Let us construct the following square matrix,

$$\widehat{A}_\varepsilon = \left(\widehat{a}_{ij} + \varepsilon\right), \ \varepsilon > 0.$$

Then $\widehat{A}_\varepsilon > 0$ and irreducible. Therefore, there exists $\lambda^* \left(\widehat{A}_\varepsilon\right) > 0$, whose corresponding characteristic vector $\widehat{X}_\varepsilon^* \in S = \{X | \ X \ge 0, \ \sum_{i=1}^n X_i = 1\}$. So, we have

$$\lambda^* \left(\widehat{A}_\varepsilon\right) \widehat{X}_\varepsilon^* = \widehat{A}_\varepsilon \widehat{X}_\varepsilon^*. \tag{12.3}$$

Let us take a sequence $\{\varepsilon_i\}$ satisfying that $\varepsilon_i \to 0$ $(i \to \infty)$. Then the corresponding sequence $\left\{\widehat{X}_{\varepsilon_i}^*\right\} \subset S$. From the compactness of S, it follows that there exists a subsequence

$$\left\{\widehat{X}_{\varepsilon_{i_k}}^*\right\}, \widehat{X}_{\varepsilon_{i_k}}^* \to \widehat{X}^* \ (i_k \to \infty), \ \widehat{X}^* \in S.$$

Therefore, $\widehat{X}^* \ge 0$.

Without loss of generality, let us assume $\lim_{\varepsilon \to 0} \widehat{X}_\varepsilon^* = \widehat{X}^*$. By taking limits for equ. (12.3), we obtain

$$\lambda^* \left(\widehat{A}\right) \widehat{X}^* = \widehat{A} \widehat{X}^*$$

and $\widehat{X}^* \ge 0$. □

12.3 Responsibility and Influence Coefficients

12.3.1 Grey Responsibility and Influence Coefficients

In each economy, the development of various business sectors is interconnected. The interdependency and incidence relationship between businesses can be studied using input-output tables. In this sub-section, we study such concepts as grey responsibility coefficients, grey influence coefficients, etc., so that input-output tables can be applied in studies of business structures.

Definition 12.3.1. *In the grey input-output model*

$$X(\otimes) = (E - A(\otimes))^{-1} Y(\otimes),$$

$(E - A(\otimes))^{-1}$ *is called an inverse matrix coefficient.*

Definition 12.3.2. *Assume that* $C(\otimes) = [E - A(\otimes)]^{-1} - E = (c_{ij}(\otimes))$ *is a grey complete consumption coefficient matrix. Then,*

$$u_i(\otimes) = \frac{\sum_{j=1}^{n} c_{ij}(\otimes)}{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(\otimes)} \tag{12.4}$$

is called the grey responsibility coefficient of the i*th department.*

Definition 12.3.3. *Assume that* $C(\otimes) = (c_{ij}(\otimes))$ *is a grey complete consumption coefficient matrix. Then,*

$$v_i(\otimes) = \frac{\sum_{i=1}^{n} c_{ij}(\otimes)}{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(\otimes)} \tag{12.5}$$

is called the grey influence coefficient of the j*th department.*

The concept of grey responsibility coefficient $u_i(\otimes)$ of the ith department reflects the impact on the output of the ith department caused by production increases of all departments. And, the concept of grey influence coefficient $v_i(\otimes)$ of the jth department stands for the impact on the output of all departments caused by a demand increase of the jth department. With these concepts in place, we are now ready to study the incidence relationship between different departments.

Example 12.3.1. Shown below (Table 12.1) is a price type grey input-output table for a certain economic district. Here, $\otimes_{11} \in [8, 12]$, $\otimes_{12} \in [25, 35]$, $\otimes_{13} \in [27, 33]$, $\otimes_{14} \in [37, 43]$, and $\otimes_{15} \in [40, 60]$. Compute the grey direct consumption coefficients, grey complete consumption coefficients, grey inverse matrix coefficients, grey responsibility coefficients, and grey influence coefficients.

Table 12.1. The price type input-ouput table of a certain economic district

Input / Output	1 Dept A	2 Dept B	3 Dept C	4 Dept D	5 Dept E	Total	Ultimate Product	Total Product
1. dept A	$\otimes_{11}$	$\otimes_{12}$	$\otimes_{13}$	$\otimes_{14}$	$\otimes_{15}$	$\otimes_1$	350	$\otimes$
2. dept B	40	30	20	50	60	200	200	400
3. dept C	50	30	60	20	10	100	220	400
4. dept D	30	30	20	10	40	130	270	400
5. dept E	40	30	20	20	30	140	360	500
6. goods	100	150	150	160	190			
7. energy	60	50	40	30	40			
8. wage,taxes	90	50	60	70	80			
total input	500	400	400	400	500			

Solution. Using the concept of the grey input-output coefficient matrix, we can compute grey direct consumption coefficients. Here, we only list the mean-value whitenization matrix for the grey direct consumption coefficient matrix in Table 12.2 below.

Table 12.2. Mean-value whitenization matrix of grey direct consumption coefficient matrix

Input / Output	1 Dept A	2 Dept B	3 Dept C	4 Dept D	5 Dept E
1. Dept A	0.020	0.050	0.075	0.100	0.100
2. Dept B	0.080	0.075	0.050	0.125	0.120
3. Dept C	0.100	0.102	0.153	0.050	0.020
4. Dept D	0.060	0.075	0.053	0.025	0.080
5. Dept E	0.080	0.075	0.053	0.025	0.080
6. Goods	0.360	0.375	0.375	0.400	0.380
7. Energy	0.120	0.125	0.100	0.075	0.080
8. Wages, taxes	0.180	0.125	0.150	0.175	0.160

From $C(\otimes) = (E - A(\otimes))^{-1} - E$, we can compute the mean-value whitenization matrix of the grey complete consumption coefficient matrix (Table 12.3 below). Similarly, we can calculate the needed grey inverse matrix coefficients, grey responsibility coefficients, and grey influence coefficients (see Table 12.4 below). Now, based on the obtained responsibility and influence coefficients we can study the incidence relationship between different departments in economic growth.

Table 12.3. Mean-value whitenization matrix of the grey complete consumption coefficient matrix

Output \ Input	1 Dept A	2 Dept B	3 Dept C	4 Dept D	5 Dept E
1. Dept A	0.059	0.094	0.116	0.133	0.138
2. Dept B	0.128	0.134	0.098	0.172	0.175
3. Dept C	0.147	0.155	0.209	0.099	0.070
4. Dept D	0.092	0.117	0.085	0.059	0.117
5. Dept E	0.114	0.118	0.087	0.087	0.100
6. Goods	0.564	0.609	0.599	0.606	0.606
7. Energy	0.174	0.187	0.160	0.134	0.142
8. Wages, taxes	0.263	0.221	0.243	0.260	0.254

Table 12.4. Mean-value whitenizations of grey inverse matrix, grey responsibility, and influence matrices

Output \ Input	1 Dept A	2 Dept B	3 Dept C	4 Dept D	5 Dept E	Sum of Inverse Matrix Coefficients	Influence Coeffiecients
1. Dept A	1.059	0.094	0.116	0.133	0.133	1.540	0.9743
2. Dept B	0.128	1.134	0.098	0.172	0.175	1.707	1.0799
3. Dept C	0.147	0.155	1.209	0.099	0.070	1.680	1.0629
4. Dept D	0.092	0.117	0.085	1.059	0.117	1.470	0.9300
5. Dept E	0.144	0.118	0.087	0.087	1.100	1.506	0.9528
Sum of inverse Matrix coefficients	1.540	1.610	1.595	1.55	1.600		
Influence coefficients	0.9743	1.0237	1.0097	0.9806	1.0122		

From Table 12.4, it can be seen that both departments B and C have relatively high responsibility and influence coefficients, indicating that these two departments have played a dominant role in the economic development of the region of our study.

12.3.2 Other Related Coefficients

In grey input-output analysis, one also uses other related coefficients. In this sub-section, we only list their relevant definitions.

Definition 12.3.4. *When the* j*th department increases its unit ultimate products, it causes each department to increase its individual production value. This increased total sum is called a grey impact coefficient, written*

d_j and defined by

$$d_j(\otimes) = \sum_{i=1}^{n} c_{ij}(\otimes), \tag{12.6}$$

where $C(\otimes) = [c_{ij}(\otimes)] = (E - A(\otimes))^{-1} - E$. So, $d_j(\otimes)$ is the sum of the entries in the jth column of the complete consumption coefficient matrix. It reflects the leading role played by the jth department on all other departments.

Definition 12.3.5. *In order to achieve the goal that each department will increase its unit ultimate products, the required production value increase out of the ith department is called a constraint coefficient denoted $z_i(\otimes)$. That is, in symbols, we have*

$$z_i(\otimes) = \sum_{j=1}^{n} c_{ij}(\otimes), i = 1, 2, ..., n. \tag{12.7}$$

The constraint coefficient in fact equals the sum of the entries in the ith row of the complete consumption matrix. It reflects the constraint the ith department has over all other departments.

Definition 12.3.6. *When the jth department increases its unit ultimate products, it may cause each department to earn increased net production value. The sum of these increased production values is called a grey induced economic benefit coefficient, denoted $p_i(\otimes)$. In symbols,*

$$p_i(\otimes) = \sum_{i=1}^{n} c_{ij}(\otimes) v_i, j = 1, 2, ..., n. \tag{12.8}$$

where v_i stands for the net production value rate of the ith department.

The concept of grey induced economic benefit coefficients reflects the enticing role played by the ith department through its increased ultimate goods production.

12.4 Optimal Input-Output Models

In practice, input-output analysis is indeed an indispensable scientific method used in global systems design and total systems coordination. However, this analysis cannot guarantee that various proportional relationships as indicated by the resultant coordination and synthetic equilibrium of the analysis are the optimal design meeting the practical requirements of the physical situation. In this section, we employ grey systems theory to establish optimal grey input-output models that combine the input-output method, as studied earlier, and linear programming into an organic whole.

The modeling idea of optimal grey input-output models can be described as follows. Seek a set X such that the objective function $f(X)$ reaches its extreme value and the following conditions are satisfied,

$$\begin{cases} (E - A(\otimes))X \geq Y(\otimes) \\ A(\otimes)X \leq B(\otimes) \\ X \geq 0. \end{cases} \tag{12.9}$$

In equ. (12.9), $(E - A(\otimes))$ stands for Leontief grey matrix, $Y(\otimes)$ the grey demand vector, whose components in general are (either continuous or discrete) interval grey numbers, representing the upper and lower limits of the societal demands, $A(\otimes)$ the grey consumption coefficient matrix, $B(\otimes)$ a grey constraint vector, and X the decision vector, that is, $X = [x_1, x_2, ..., x_n]$.

Each grey input-output optimal model can be established by going through the following steps.

1. Use a time series to represent the values of the constraint condition.

Through GM(1,1) model predictions, one can obtain a time series representing development patterns existing in the values of the constraints. Then, program based on the predicted values. The optimal grey input-output programming obtained in this way can not only reflect a certain (static state) situation of the system under consideration, but also present the evolutionary development of the constraints. The outcome obtained from such an optimal model may be either a single answer or a set of time series values. Such a solution not only reflects the optimal relationship (structure), but also provides a way for the researcher to understand the development tendency of the optimal relationship.

2. Parameters in the constraints are grey.

In the constraint conditions in equ. (12.9), we have the following grey matrices and grey vectors.

$$E - A(\otimes) = \begin{bmatrix} 1-\otimes_{11} & -\otimes_{12} & \cdots & -\otimes_{1n} \\ -\otimes_{21} & 1-\otimes_{22} & \cdots & -\otimes_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -\otimes_{m1} & -\otimes_{m2} & \cdots & 1-\otimes_{mn} \end{bmatrix},$$

$$A(\otimes) = \begin{bmatrix} \otimes_{11} & \otimes_{12} & \cdots & \otimes_{1n} \\ \otimes_{21} & \otimes_{22} & \cdots & \otimes_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \otimes_{m1} & \otimes_{m2} & \cdots & \otimes_{mn} \end{bmatrix}, \otimes_{ij} \in \left[\underline{a}_{ij}, \overline{a}_{ij}\right],$$

$$Y(\otimes) = [\otimes_1, \otimes_2, ..., \otimes_m]^T, \otimes_i \in \left[\underline{y}_i, \overline{y}_i\right],$$

and

$$B(\otimes) = [\otimes_1, \otimes_2, ..., \otimes_m]^T, \otimes_i \in \left[\underline{b}_i, \overline{b}_i\right].$$

3. The objective function.

$$\begin{cases} \max f(X) = \max K(\otimes)^T \cdot X \\ K(\otimes) \in \left[\underline{k}_i, \overline{k}_i\right] \\ K(\otimes) = (k_1(\otimes), k_2(\otimes), ..., k_n(\otimes)). \end{cases} \tag{12.10}$$

4. The method of finding the solution.

After all grey elements appearing in matrices or vectors are whitenized, one can get on a computer to find the solution. First, she will find the optimal solution for the ideal model and the optimal solution for the boundary model. Then, based on pre-determined requirements, she will need to study several different positioned programming problems in order to obtain enough designs for her final choice. After a comprehensive comparison, she will be able to find the most satisfactory programming design.

Example 12.4.1. Let us see how an optimal model is established for the productional structure of a farm of a certain region.

Based on relevant analysis, we obtain the following constraints:

1. Constraints on the demand of the ultimate products:

$$(E - A) X \geq G.$$

2. Constraints on resources:

Fertilizers:

$$\begin{aligned} 0.24343x_1 \quad &+0.21586x_2 + 0.22458x_3 + 0.20266x_4 + 0.1015x_5 \\ &+0.2935x_6 + 0.00439x_7 + 0.03733x_8 \leq G_{15}. \end{aligned}$$

Electricity:

$$\begin{aligned} 0.01248x_1 \quad &+G_{24}x_2 + 0.01135x_3 + G_{25}x_9 + G_{26}x_{10} \\ &+0.00954x_{11} + G_{27}x_{12} \leq G_{16}. \end{aligned}$$

Chemicals:

$$\begin{aligned} 0.01707x_1 \quad &+0.00757x_2 + 0.00364x_3 + 0.00468x_4 \\ &+0.0035x_5 + 0.00736x_6 \leq G_{17}. \end{aligned}$$

Capitals:

$$\begin{aligned} 0.38969x_1 \quad &+0.22343x_2 + 0.33457x_3 + 0.23069x_4 \\ &+0.21243x_5 + 0.30513x_6 + 0.0044x_7 + 0.03734x_8 \\ &+0.09003x_9 + 0.6316x_{10} + 0.20419x_{11} + 0.07751x_{12} \\ &+0.02016x_{13} + 0.10133x_{14} \leq G_{18}. \end{aligned}$$

3. Constraints on goods production abilities.

$$x_7 \leq G_{19}, x_8 \leq G_{20}, x_{13} \leq G_{21}, x_{14} \leq G_{22}.$$

The following is our objective function:

$$\begin{aligned}\max f(X) \quad = \quad & G_{28}x_1 + 0.51542x_2 + 0.67809x_3 + 0.61977x_4 \\ & +0.76343x_5 + 0.49995x_6 + 0.93798x_7 + 0.91692x_8 \\ & +0.39947x_9 + 0.17217x_{10} + 0.51828x_{11} + 0.3051x_{12} \\ & +0.97974x_{13} + 0.86666x_{14},\end{aligned}$$

where E stands for the unit matrix, A the direct consumption coefficient matrix, G the constraint value grey vector, and all the components of X are given in the following.

- x_1 stands for food,
- x_x crops for cooking oil,
- x_3 teas,
- x_4 fruits,
- x_5 vegetables,
- x_6 miscellaneous,
- x_7 bamboo products,
- x_8 other forestry products,
- x_9 livestock,
- x_{10} domestic birds and poultry,
- x_{11} food processing,
- x_{12} construction materials,
- x_{13} industry, and
- x_{14} fishery.

The definition intervals for the grey parameter G are given as follows.

- constraint on food demand $G_1 \in (14,225,15,000)$;
- constraint on oil ingredients $G_2 \in (965,1,000)$;
- constraint on the demand of teas $G_3 \in (7,870,7,900)$;
- constraint on the demand of fruits $G_4 \in (634,640)$;
- constraint on the demand of vegetables $G_5 \in (8,936,9,000)$;
- constraint on the demand of miscellaneous $G_6 \in (215,220)$;

- constraint on the demand of bamboo products $G_7 \in (297, 300)$;
- constraint on other forestry products $G_8 \in (1,249, 15,000)$;
- constraint on the demand of livestock $G_9 \in (10,158, 10,160)$;
- constraint on the demand of domestic birds and poultry $G_{10} \in$ (715, 720);
- constraint on the demand of farm products processing $G_{11} \in$ (7,846, 7,851);
- constraint on the demand of construction materials $G_{12} \in$ (9,045, 9,052);
- constraint on the demand of industrial products $G_{13} \in (336, 340)$;
- constraint on the demand of fishery products $G_{14} \in (97, 100)$;
- constraint on the supply of fertilizers $G_{15} \in (9,850, 9,900)$;
- constraint on the supply of electricity $G_{16} \in (450, 455)$;
- constraint on the supply of farm chemicals $G_{17} \in (475, 481)$;
- constraint on the supply of capitals $G_{18} \in (22,000, 22,110)$;
- constraint on production capability of bamboo products $G_{19} \in$ (885, 889);
- constraint on production capability of other forestry products $G_{20} \in$ (1,380, 1,385);
- constraint on production capability of industrial products $G_{21} \in$ (2,880, 2,886);
- constraint on production capability of fishery products $G_{22} \in (330, 335)$;
- products consumed when producing a unit of tea $G_{23} \in (0.054, 0.056)$;
- electricity consumed when producing a unit of oil materials $G_{24} \in$ (0.00023, 0.00035);
- electricity consumed when producing a unit of livestock $G_{25} \in$ (0.011, 0.012);
- electricity consumed when producing a unit of domestic birds and poultry $G_{26} \in (0.0003, 0.0005)$;
- electricity consumed when producing a unit of construction materials $G_{27} \in (0.0062, 0.0082)$; and

- net production value from producing a unit of food $G_{28} \in (0.2134, 0.3731)$.

After doing all necessary computations on a computer, we obtained four optimal plans for the final selection. These four plans were then evaluated by the region's politicians, experts, and other relevant personnel, and the most satisfactory plan was eventually selected (see Table 12.5 below for more details). On the basis of this selected plan, we established an expanded sequence of interconnected optimal grey input-output models. Because each agricultural economic system is huge with many layers, one must establish a group of models connecting different layers in order to realistically reflect the whole systemwide optimization and to be effectively implementable.

Table 12.5. Comparison between computer results and the original programming output in ¥10,000

Variable	Satisfactory plan	Theoretical plan	Comparison
x_1	209.77	209.71	+06
x_2	9.61	9.61	0
x_3	78.81	78.81	0
x_4	6.40	6.40	0
x_5	150.39	90.56	+59.83
x_6	4.84	4.84	0
x_7	8.55	5.37	+3.18
x_8	13.80	13.20	+0.60
x_9	137.46	137.35	+0.11
x_{10}	7.40	7.40	0
x_{11}	78.46	78.46	0
x_{12}	90.45	90.45	0
x_{13}	28.80	28.21	+0.59
x_{14}	3.30	1.08	+2.22
Net Production Value	434.03	382.26	Total: 51.77

In Table 12.5, the theoretical plan is the one produced out of the original programming model and the comparison is between the satisfactory plan and the theoretical plan.

From the angle of practical applications, the grey input-output optimal model has the following special characteristics.

1. It has the ability to supplement many different aspects of modeling. The method of input-output can balance and harmonize the system under consideration. However, it cannot optimize the situation. On the other hand, the method of linear programming can perform optimization without being able to achieve comprehensive balance. By combining the two methods, we can use the strength of one to overcome the weakness of the other.

2. It is relatively easy to find the solution. Because many parameters have a range to take their individual values, it is possible to be flexible on computers and to derive a set of many choices before a satisfactory solution is reached.

3. It is convenient for grey target decision making. Agricultural economic systems are always very complicated and grey with many uncertain and missing parameters. It is extremely difficult for the researcher to obtain a definite optimal value hitting a specific "target" (objective). Through qualitative analysis and quantitative tests, it becomes a relatively easy task for him to determine a grey region within which the development of the agricultural economic system under consideration falls in the "grey target."

12.5 Dynamic Input-Output Models

In each analysis of economic systems, it is often necessary to study the economic structure of the future, especially possible changes in investment structures and impacts of economic policy changes on the economy. On the basis of such study, predictions about future economic development can be made so that decisions about optimal investment scales and investment structures can be derived. Due to such a need, in this section we introduce a method that can capture the dynamics of economic systems' movements. More specifically, in this section we study the grey dynamic input-output method.

12.5.1 Dynamic Comprehensively Balanced Model

The so-called grey dynamic input-output model is developed on the basis of the grey static input-output method. It introduces the concepts of grey capital coefficients or investment coefficient matrices so that investment becomes a part of the model. That is, in this new class of study, through development of a model both production and investment are synchronously and quantitatively calculated with an introduction of a dynamic concept such as moving time. With this new model in place, demand for investment and economic development, the present, and the future are connected together. And, the relationship between the accumulation of fixed properties and expanded production can be dynamically considered on the basis of a given time series. A grey dynamic input-output model reflects the process of a business expansion and production and is an ideal research tool for studying problems of investment.

More specifically, a grey dynamic input-output model can be written as follows,

$$X_t(\otimes) - A_t(\otimes) X_t(\otimes) - B_{t+1}(\otimes) \left[X_{t+1}(\otimes) - X_t(\otimes) \right] = C_t(\otimes). \quad (12.11)$$

Each so-called grey dynamic comprehensively balanced model is established on the grey dynamic input-output model with its emphasis on the study of such a comprehensive balance relationship that exists in the process of reproduction of an economy. It describes the dynamic process of production and distribution of products of all departments of an economy. The main balance equation used in a grey dynamic comprehensively balanced model is given below.

1. The balance equation for production and consumption of products is:

$$\begin{aligned} (E - A(\otimes) &+ B(\otimes))\, X_t(\otimes) - B(\otimes)\, X_{t+1}(\otimes) = C_t(\otimes) \\ &t = 0, 1, ..., T-1 \\ (E - A(\otimes)) &\; X_t(\otimes) = C_t(\otimes), \end{aligned} \tag{12.12}$$

where $A(\otimes)$ stands for a grey direct consumption coefficient matrix, $B(\otimes)$ a grey investment coefficient matrix, $C_t(\otimes)$ the ultimate net demand grey vector of the tth year, $X_t(\otimes)$ the grey production vector of each department of the tth year, and T the objective number of years.

2. The balance equation for investment demand and supply is:

$$\sum_{j=1}^{n} b_j(\otimes)\left[x_j(t+1) - x_j(t)\right] = k(t), \tag{12.13}$$

for $t = 0, 1, ..., T-1$, where $b_j(\otimes)$ stands for the investment coefficient of the jth department, $x_j(\otimes)$ the production value of the jth department in the tth year, $k(t)$ the supply of investment capital in the tth year, and n the number of departments.

3. The balance equation for the demand of labor and supply is:

$$\sum_{j=1}^{n} L_j(\otimes)\, x_j(t) = L(t), \tag{12.14}$$

for $t = 0, 1, ..., T$, where $L_j(\otimes)$ stands for demand (for labor) coefficient for the jth department to produce a unit of production value, and $L(t)$ the available labor supply of the tth year.

12.5.2 Optimal Dynamic Input-Output Model

The grey dynamic input-output optimal model is a generalization of the grey static input-output optimal model studied in Section 12.4 above. Its specific form looks as follows.

$$\begin{aligned} &\max SX(T) \\ &s.t. \begin{cases} (E - A(\otimes) + B(\otimes))\, X_t(\otimes) - B(\otimes)\, X_{t+i}(\otimes) \geq C_t(\otimes) \\ A(\otimes)\, X_t(\otimes) \leq B(\otimes) \\ X_t(\otimes) \geq 0. \end{cases} \end{aligned} \tag{12.15}$$

Example 12.5.1. Let us see here in this example how a grey dynamic input-output optimal model for a certain geographic region is established.

Solution.

The objective function: $\max SX(T)$.

The constraints:

1. Productional balance constraint:

$$(E - A(\otimes) + B(\otimes)) X_t(\otimes) - B(\otimes) X_{t+i}(\otimes) \geq C_t(\otimes),$$

for $t = 0, 1, 2, ..., T - 1$;

2. Production value constraint:

$$X_{t+1}(\otimes) - X_t(\otimes) \geq 0,$$

for $t = 0, 1, 2, ..., T - 1$;

3. Capital constraint:

$$E \cdot B(\otimes) [X_{t+1}(\otimes) - X_t(\otimes)] \leq K_t(\otimes),$$

for $t = 0, 1, 2, ..., T - 1$;

4. Constraint of water resource:

$$W(\otimes) \cdot X_t(\otimes) \leq Q_t(\otimes),$$

for $t = 0, 1, 2, ..., T$;

5. Constraint of energy:

$$G(\otimes) \cdot X_t(\otimes) \leq P_t(\otimes),$$

for $t = 0, 1, 2, ..., T$;

6. Constraint of labor force:

$$L(\otimes) \cdot X_t(\otimes) \leq R_t(\otimes),$$

for $t = 0, 1, 2, ..., T$, and

7. Constraint of the environment:

$$U(\otimes) \cdot X_t(\otimes) \leq H_t(\otimes),$$

for $t = 0, 1, 2, ..., T$, where all variables are defined as follows.

- S stands for the vector of net production value coefficients,
- T the length of time period to be planned,
- $A(\otimes)$ the grey direct consumption coefficient matrix,

- $B(\otimes)$ the grey investment coefficient matrix,
- $C_t(\otimes)$ the net demand grey vector at the end of the tth year,
- $X_t(\otimes)$ the production value vector of the tth year,
- $K_t(\otimes)$ the grey vector of capital supply of the tth year,
- $W(\otimes)$ the grey vector of water consumption per unit production value of all the departments,
- $Q_t(\otimes)$ the grey vector of available water supply of the tth year,
- $G(\otimes)$ the grey vector of energy consumption per unit production value of all the departments,
- $P_t(\otimes)$ the grey vector of energy supply of the tth year,
- $L(\otimes)$ the grey vector of labor demand per unit production value of all the departments,
- $R_t(\otimes)$ the grey vector of labor supply of the tth year,
- $U(\otimes)$ the grey vector of pollution coefficients per unit production value of all the departments,
- $H_t(\otimes)$ the grey vector of upper limits of total pollution of the tth year, and
- E stands for the n-dimensional identify matrix.

12.5.3 Von Neumann Model

The so-called grey von Neumann model is a combined model using grey systems theory and the von Neumann model. The grey von Neumann model can be employed to study the optimal speed and optimal economic structure for balanced economic growth of a nation, a region, or a city. It can also be employed to explore the relationship between the rate of accumulation and the speed of economic growth.

The specific form of this model is given in the following,

$$\max PB(\otimes)X_T(\otimes)$$
$$s.t.\begin{cases} (E-A(\otimes) & +B(\otimes))X_t(\otimes)-B(\otimes)X_{t+1}(\otimes) \\ & \geq \delta(\otimes)D(\otimes)V(\otimes)X_t(\otimes),t=0,1,...,T-1 \\ X_t(\otimes)\geq 0, & t=0,1,...,T, \end{cases} \tag{12.16}$$

where all variables are defined as follows.

- $A(\otimes)$ stands for the grey direct consumption coefficient matrix,

- $B(\otimes)$ the grey investment coefficient matrix,
- $D(\otimes)$ a grey column vector for the ultimate net demand structure,
- $X_t(\otimes)$ the grey vector of production values of all the departments of the tth year,
- P a row vector of weights,
- $V(\otimes)$ a grey row vector of net production value coefficients,
- $\delta(\otimes)$ the consumption rate and $(1-\delta(\otimes))$ the accumulation rate,
- T the time span of the planning, and
- $V(\otimes)\cdot X_t(\otimes)$ the grey GDP total.

12.6 Practical Applications

During a time period where currencies still represent value, price, as a measure of labor content in commercial products, plays a significant role in the growth of a market economy. The feasibility of a price system is relative in terms of an optimal productivity strength. Developing such an optimal price system corresponding to von Neumann optimal productivity strength that it can guarantee the whole of an economy sustains a high speed and balanced growth should be the goal of each reform of an existing price system.

In this section, we apply grey input-output theory to analyze the tendency of the price system reform of modern China. In our example here, we divide the Chinese economy into four comprehensive departments: agriculture, agricultural processing industry, other industries, and services. In the department of agriculture, we further have the following five subdivisions: food crop planting, other crop planting, forestry, livestock husbandry, and farm industry. Based on the input-output table of Henan Province, the People's Republic of China, of 1987, and relevant statistics, we predicted the relevant data values for the economy of the national scale. Considering the fact that agricultural production contains a lot of uncertain information, we derived the following grey input matrices for the four comprehensive departments and five sub-divisions of the agriculture department, respectively,

$$A(\otimes)=\begin{bmatrix} \otimes_{11} & \otimes_{12} & \otimes_{13} & \otimes_{14} \\ 24.80 & 67.33 & 59.04 & 45.53 \\ 90.12 & 31.98 & 365.22 & 66.18 \\ 41.52 & 49.71 & 135.35 & 136.68 \end{bmatrix},$$

where $\otimes_{11} \in [155.00, 164.78]$, $\otimes_{12} \in [54.09, 56.09]$, $\otimes_{13} \in [25.22, 27.86]$, and $\otimes_{14} \in [18.7, 21.1]$, and

$$a(\otimes) = \begin{bmatrix} 56.44 & 26.83 & 3.66 & 27.02 & 6.97 \\ \otimes_{21} & \otimes_{22} & \otimes_{23} & \otimes_{24} & \otimes_{25} \\ 3.56 & 2.55 & 2.42 & 0.99 & 0.66 \\ 10.54 & 7.27 & 1.08 & 4.38 & 1.60 \\ 5.16 & 3.44 & 0.51 & 1.75 & 1.24 \end{bmatrix},$$

where $\otimes_{21} \in [6.18, 7.98]$, $\otimes_{22} \in [5.97, 7.57]$, $\otimes_{23} \in [0.38, 0.76]$, $\otimes_{24} \in [1.47, 2.33]$, and $\otimes_{25} \in [0.98, 1.40]$.

And, we also derived the following grey output matrices for the four comprehensive departments and five subdivisions of the agriculture department, respectively,

$$B(\otimes) = \begin{bmatrix} \otimes_{11} & 0 & 0 & 0 \\ 0 & 257.65 & 0 & 0 \\ 0 & 0 & 638.32 & 0 \\ 0 & 0 & 0 & 343.01 \end{bmatrix},$$

where $\otimes_{11} \in [388.83, 428.83]$, and

$$b(\otimes) = \begin{bmatrix} 89.74 & 0 & 0 & 0 & 0 \\ 0 & \otimes_{22} & 0 & 0 & 0 \\ 0 & 0 & 12.33 & 0 & 0 \\ 0 & 0 & 0 & 42.90 & 0 \\ 0 & 0 & 0 & 0 & 17.20 \end{bmatrix},$$

where $\otimes_{22} \in [60.00, 72.16]$.

Assume that Z^*, α^*, and P^* are, respectively, the optimal strength, expansion rate, and optimal price, of the four comprehensive departments, and that S^*, β^*, and V^*, respectively, are the optimal strength, expansion rate, and optimal price of the five sub-divisions within the agriculture department. From the following von Neumann models

$$\begin{cases} Z^* A(\otimes)\alpha^* \le Z^* B(\otimes) \\ \alpha^* A(\otimes) P^* \ge B(\otimes) P^* \\ Z^* A(\otimes)\alpha^* P^* = Z^* B(\otimes) P^* \end{cases} \tag{12.17}$$

and

$$\begin{cases} S^* a(\otimes)\beta^* \le S^* b(\otimes) \\ \beta^* a(\otimes) V^* \ge b(\otimes) V^* \\ S^* a(\otimes)\beta^* V^* = S^* b(\otimes) V^* \end{cases} \tag{12.18}$$

after mean-value whitenizing the grey input and grey output matrices, it is not difficult to find the optimal prices corresponding to the optimal strengths Z^* and S^* and expansion rates α^* and β^*:

$$P^* = \left([p_i^*]_{i=1}^4\right)^T = [1, 1.6015, 1.7658, 2.3923]^T$$

and

$$V^* = \left([v_i^*]_{i=1}^5\right)^T = [1, 0.174, 0.575, 0.404, 0.496]^T .$$

These optimal prices, which may well promote the Chinese national economy to effectively develop along the von Neumann high speed track, should be employed as one of the important bases for making price-related policies. According to the reasonable price ratios, the prices of the products of the agricultural processing industry should be 60% higher than those of agricultural products, the prices of the products of the other industries should be 77% higher than those of agricultural products, and the prices of the services area should be 139% higher. However, the reality is that the prices of agricultural products are extremely low. Its consequence has been that all input into agricultural production has been dropping annually and the foundation of the national economy has been greatly weakened. In recent years since China adopted a open door policy and started an economic reform, the economic market price system has been adjusted systematically and prices on agricultural products and preliminary mineral products have increased steadily. However, the problem of unreasonable pricing has not been resolved from its root level. In 1988, the tax income per ¥100 investment IP (investment price) and full-time labor productivity LP (labor price) for Chinese agriculture, industry, and construction is listed in Table 12.6 below.

Table 12.6. IP and LP data for 1988

	Area 1	Area 2	Area 3	Area 4	Area 5
IP	1.84	27.00	24.67	17.03	5.20
LP	1,486	16,008	17,547	12,160	10,345

where areas 1 to 5 stand, respectively, for agriculture, light industry with agricultural products as its raw materials, light industry with nonfarm products as its raw materials, heavy industry, and construction.

That is, we have the following market prices in 1988,

$$IP = (1, 14.7717, 13.4076, 9.2554, 2.8261)$$

$$LP = (1, 10.7725, 11.8082, 8.1830, 6.9616) .$$

It can also be said that the capital price and labor price, contained in the products of light industry with agricultural products as its raw materials, are respectively 1,377.17% and 977.25% higher than those contained in agricultural products. The capital and labor prices, contained in the products of light industry with non-farm products as its raw materials, heavy industry, and construction, are 1,240.76%, 1,080.82%; 825.54%, 718.30%; 182.61%, 596.16% higher than those contained in agricultural products, respectively. Even when we do not consider the differences existing in labor qualities from different business departments, it is still not difficult for us to see the seriousness of the distorted price relationship.

Similarly, from the angle of von Neumann optimal price relationship among the sub-divisions within the agriculture department, it can be seen that the prices of the products of the sub-division of food crop planting should be higher than those of the products of all other sub-divisions. However, the fact is exactly the opposite. No matter whether we look at the benefits of capital input or the benefits of labor input, food crop sub-division is lower than other sub-divisions. Consequently, it implies that it is extremely difficult for farmers to get ahead financially if they focus on planting food crops.

Such double distortions, one appearing in the fact that agricultural products are priced lowest among all products of the four comprehensive economic departments, and the other that food products are priced lowest among agricultural subdivisions, have made maintaining and developing agricultural and food production extremely expensive in terms of capital, labor, and supporting materials input without much adequate return. They have caused low return on agricultural investment and high benefits from industrial capital, food crop planting a low-benefit business and other crop planting get-rich-quick operations. Consequently, all the economic regions, focusing on food crop plantings, have been slow in terms of economic growth. Such a phenomenon has seriously affected the motivation for farmers to produce more food and greatly diminished the attractiveness of food crop planting. Therefore, it is extremely important for the Chinese government to consider price reform of the national economic market as a desperate task on hand. We believe that the starting point of such a price reform should be increasing the prices of agricultural products. And within the department of agriculture, the government should first consider increasing the prices of food crops.

On the other hand, one may ask: if prices for agricultural products and foods are increased dramatically, will it cause the broad range price index to jump, bringing hardship to people's daily living and national stability and economic development? To this end, we also did some initial research. Our calculated results indicate that when the prices of agricultural products and food change in the amounts of ΔP_1 and ΔV_1, the prices of the products of the other departments and sub-divisions will, respectively, change in the amounts of

$$\Delta P = \Delta P_1 \left[0.3426, 0.0700, 0.0519\right]$$

and

$$\Delta V = \Delta V_1 \left[0.0442, 0.0160, 0.3414, 0.0949\right].$$

So, it can be seen that if the prices for agricultural products increase 5%, 10%, 20%, 50%, or even 100%, then

1. The prices of the products of agricultural processing industry, as affected by increased prices of agricultural products, will increase accordingly 1.7%, 3.43%, 6.85%, 17.13%, or 34.26%.

2. The prices of the department of other industries will accordingly increase 0.3%, 0.7%, 1.4%, 3.5%, or 7%. And

3. The prices of the department of services will increase accordingly 0.25%, 0.52%, 1.04%, 2.6%, or 5.19%.

Within the department of agriculture, if the prices of food crops increase 5%, 10%, 20%, 50%, or 100%, then

1. The prices of all other crops will increase accordingly 0.2%, 0.44%, 0.88%, 2.21%, or 4.42%.

2. The prices of forestry products will increase accordingly 0.08%, 0.16%, 0.32%, 0.8%, or 1.6%.

3. The prices of the products of the sub-division of livestock husbandry will increase accordingly 1.7%, 3.4%, 6.8%, 17.07%, or 34.14%. And,

4. The prices of the products of farm industry will increase accordingly 0.47%, 0.95%, 1.9%, 4.7%, or 9.49%.

Therefore, it can be seen that bringing up the prices of agricultural products, other than the agricultural processing industry and livestock husbandry, does not really affect the prices of other products. That is, even under the consideration of market price bounces, if the market economy gradually evolves into maturity, with adequate adjustment on the speed of reform and increased strength of macroscopic control, the phenomenon of market-wide landslide price increases can be completely avoided.

We have also studied the impact of changes in price of agricultural and food products on consumption expense and accumulation expense. Our results indicate that when the prices of agricultural products change in the amount of ΔP_1, the changes in the consumption expense and accumulation expense on agricultural products are, respectively, given as follows:

$$\Delta\alpha_1 = \frac{\Delta P_1}{4.5179 + 1.4941\Delta P_1} \text{ and } \Delta\beta_1 = \frac{\Delta P_1}{7.675 + 1.182\Delta P_1}.$$

If the prices of agricultural products are increased 5%, 10%, 20%, 50%, or 100%, the corresponding consumption expenses on agricultural products increase, respectively, 1.09%, 2.14%, 4.15%, 9.5%, or 16.63%. And, the corresponding accumulation expenses increase 0.06%, 1.28%, 2.53%, 6.05%, or 11.29%, respectively.

The proportional connections between consumption expenses and accumulation expenses and price changes on foods are given by

$$\Delta\alpha_1' = \frac{\Delta V_1}{7.5957 + 1.1848\Delta V_1} \text{ and } \Delta\beta_1' = \frac{\Delta V_1}{18.748 + 1.0599\Delta V_1}.$$

If the prices of foods go up 5%, 10%, 20%, 50%, or 100%, the consumption expenses on foods will increase, respectively, 0.65%, 1.3%, 2.55%, 6.1%, or 11.39%. And the corresponding accumulation expenses will increase, respectively, 0.27%, 0.53%, 1.05%, 2.59%, or 5.05%. Therefore, we can say that the negative impact, if any, of a price increase for foods and other agricultural products on people's daily lives and the national economy will be next to nothing when compared to the consequent economic benefits.

Adjusting the price relationship is an important link facing the current Chinese economic reform. It is a very complicated systems project. In order to harmonize the price relationship existing in the Chinese economic price system so that the national economy can enjoy a sustained and healthy development, we believe that any potentially successful plan for economic reform of modern China needs to have increasing prices for agricultural products as its starting point and basic direction.

13
Grey Control

13.1 Introduction

As a scientific concept, control means that the controlling equipment or party is imposing a special function or action on the controlled equipment or party. This special function or action is a purposeful and selected dynamic activity. In a control system, there exist at least three parts: controlling equipment, controlled equipment, and a communication tunnel. A control system, consisting of these three parts only, is called a *open loop control system*, which is shown in Figure 13.1. For open loop control sys-

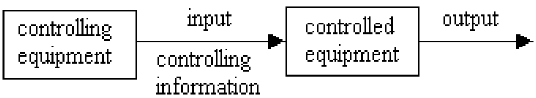

FIGURE 13.1. The flowchart of an open link control system

tems, it is relatively simple to do a needed control, because the output can be controlled directly by using the input. However, the weakness is that it is not very robust against disturbances. These control systems with *feedback loops* are called *closed loop control systems*, which are shown in Figure 13.2. In a closed loop control system, the control is realized through interactions of the input and the feedback of the output.

An outstanding strength of a closed loop control system is that it is very robust against disturbances with its output always around the pre-

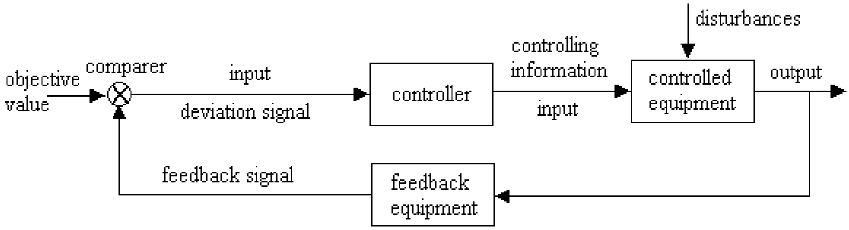

FIGURE 13.2. The flowchart for a closed link control system

defined objectives. That is, closed loop control systems possess some kind of stability.

So-called *grey control* means the control of essential grey systems, including general control systems with grey parameters and controls built on the analysis, modeling, prediction, and decision making of grey systems. When compared to traditional control theory, the methodology of grey control can more deeply reveal the characteristics of the problem under consideration, and is more beneficial for the realization of the objectives of control. Human recognition of the physical reality is a process of gradual accumulation and constant deepening of knowledge. We have to study new topics and new objects with the prerequisite of having received experience and knowledge from our forefathers. That is to say, what we are facing is often a grey system where we know to a certain degree about the structure and function of how it realizes an input-output relation. Controlling the output according to a predetermined objective is the control problem of essential grey systems. On the other hand, in order to establish a mathematical model for the system of interest, we must first recognize, distinguish, and measure the system, and study the structure of the system and determine the system's parameters through the records of input and output data. Due to the fact that there always exist errors in recognition and measuring, the structure and related parameters of the system contain a degree of greyness. Selecting input and output data values in different time intervals might also have an impact on the determination of the parameters. For example, based on the statistical data of total food production $\mathrm{x}^{(0)}(k)$ of Henan Province, the People's Republic of China, from 1949 to 1988, which is given in Table 13.1, the following GM(1, 1) model

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b$$

can be established so that different estimates of the parameters a and b can be obtained when different time intervals are considered.

Table 13.1. Food production of Henan from 1949 to 1988

Year	$x^{(0)}(k)$	Year	$x^{(0)}(k)$	Year	$x^{(0)}(k)$	Year	$x^{(0)}(k)$
1949	713.5	1959	974.5	1969	1321.5	1979	2134.5
1950	842.0	1960	887.0	1970	1555.5	1980	2148.5
1951	1052.5	1961	684.5	1971	1646.5	1981	2314.5
1952	1007.0	1962	903.0	1972	1627.0	1982	2217.1
1953	1091.0	1963	788.0	1973	1872.0	1983	2904.0
1954	1142.5	1964	950.0	1974	1861.5	1984	2893.5
1955	1250.0	1965	1166.0	1975	1941.5	1985	2710.5
1956	1211.0	1966	1227.5	1976	2122.0	1986	2545.7
1957	1180.0	1967	1382.0	1977	1947.5	1987	2948.4
1958	1264.5	1968	1330.5	1978	2097.5	1988	2663.0

For example, if we take the data values of the years 1949 $\sim$ 1952 to establish a GM(1,1) model, we obtain that

$$a = -0.0816, b = 795.0283.$$

If the data values of the years 1949 $\sim$ 1956 are used, we obtain that

$$a = -0.0498, b = 879.8842.$$

If the data values of the years 1949 $\sim$ 1960 are used, we obtain that

$$a = -0.0055, b = 1051.988.$$

If the data values of the years 1949 $\sim$ 1966 are used, we obtain that

$$a = -0.0044, b = 1083.621.$$

If the data values of the years 1949 $\sim$ 1969 are used, we obtain that

$$a = -0.0084, b = 991.1538.$$

If the data values of the years 1949 $\sim$ 1977 are used, we obtain that

$$a = -0.0319, b = 772.1509.$$

If the data values of the years 1949 $\sim$ 1982 are used, we obtain that

$$a = -0.0329, b = 759.8184.$$

If the data values of the years 1949 $\sim$ 1988 are used, we obtain that

$$a = -0.0348, b = 735.3096.$$

If we take respectively the time intervals 1971 $\sim$ 1978, 1971 $\sim$ 1983, 1971 $\sim$ 1988, 1979 $\sim$ 1983, 1979 $\sim$ 1988, and 1984 $\sim$ 1988 to establish models, we obtain, respectively,

$$a = -0.0331, b = 1655.061;$$

$$a = -0.0367, b = 1609.544;$$

$$a = -0.0319, b = 1673.111;$$

$$a = -0.0945, b = 1767.133;$$

$$a = -0.0272, b = 2231.559;$$

and

$$a = 0.0095, b = 2638.18.$$

Therefore, the parameters of the GM(1, 1) model are grey. That is, we have

$$\frac{dx^{(1)}}{dt} + a(\otimes)x^{(1)} = b(\otimes),$$

where

$$a(\otimes) \in [-0.0945, -0.0044]$$

and

$$b(\otimes) \in [772.1509, 2638.18].$$

When solving practical problems, systems with grey parameters can be seen everywhere.

The contents of grey control include grey linear systems, grey transfer functions, typical links, grey incidence control, control with abandonment, control of grey predictions, etc. In the following sections, we discuss these topics in detail.

13.2 Grey Linear Control Systems

In this section, let us first look at control problems studied in grey linear systems.

Definition 13.2.1. *Assume that*

$$U = [u_1, u_2, ..., u_s]^T$$

is a control vector,

$$X = [x_1, x_2, ..., x_n]^T$$

a state vector, and

$$Y = [y_1, y_2, ..., y_m]^T$$

is *the output vector. Then,*

$$\begin{cases} \dot{X} = A(\otimes)X + B(\otimes)U \\ Y = C(\otimes)X \end{cases}$$

is called the mathematical model of a grey linear control system, where $A(\otimes) \in G^{n \times n}$, $B(\otimes) \in G^{n \times s}$, *and* $C(\otimes) \in G^{m \times n}$. *Accordingly,* $A(\otimes)$ *is called a state grey matrix,* $B(\otimes)$ *a control grey matrix, and* $C(\otimes)$ *a grey output matrix.*

Sometimes, to emphasize changes of U, X, and Y with time, that is, the dynamic characteristics of the system, we also denote the control vector, state vector, and output vector as $U(t)$, $X(t)$, and $Y(t)$, respectively.

The first class of equations

$$\dot{X}(t) = A(\otimes)X(t) + B(\otimes)U(t)$$

in the mathematical model of grey linear control systems is called the *state equations*; the second class of equations

$$Y(t) = C(\otimes)X(t)$$

is the *output equations.*

Definition 13.2.3. *For a fixed time moment* t_0 *and predetermined accuracy requirement, if there exists* $t_1 \in [t_0, \infty)$ *such that based on the output* $Y(t)$, $t \in [t_0, t_1]$, *of the system, the system's state* $X(t)$ *can be measured with the desirable accuracy, then the system is said to be observable on the interval* $[t_0, t_1]$. *If for any* t_0, t_1, *the system is observable on the interval* $[t_0, t_1]$, *then the system is said to be observable.*

Definition 13.2.3. *For a given accuracy requirement and an objective vector*

$$J = [j_1, j_2, ..., j_m]^T ,$$

if the controlling equipment and the control vector $U(t)$ *can make the output* $Y(t)$ *reaches objective* J *with the desired accuracy through the control of the input, then the system is said to be controllable.*

Definition 13.2.4. *When some perturbations are applied to the initial value of the system,*

1. if the amplitude of the response (output) is bounded, the system is said to be stable;

2. if the response (output) can recover its initial state after a period of time, then the system is said to be asymptotically stable;

3. if the amplitude of the response becomes unbounded, then the system is said to be unstable.

In general, the concept of systems stability means that of asymptotic stability.

Theorem 13.2.1. *For the system*

$$\begin{cases} \dot{X}(t) = A(\otimes)X(t) + B(\otimes)U(t) \\ Y(t) = C(\otimes)X(t), \end{cases}$$

where

$$X(t) = [x_1(t), x_2(t), ..., x_n(t)]^T,$$
$$U(t) = [u_1(t), u_2(t), ..., u_s(t)]^T,$$
$$Y(t) = [y_1(t), y_2(t), ..., y_m(t)]^T,$$

$A(\otimes) \in G^{n\times n}$, $B(\otimes) \in G^{n\times s}$ *and* $C(\otimes) \in G^{m\times n}$. *Let*

$$D(\otimes) = \begin{bmatrix} C(\otimes) \\ C(\otimes)A(\otimes) \\ C(\otimes)A^2(\otimes) \\ \cdots \\ C(\otimes)A^{n-1}(\otimes) \end{bmatrix}$$

and

$$L(\otimes) = \begin{bmatrix} B(\otimes) & A(\otimes)B(\otimes) & A^2(\otimes)B(\otimes) & \cdots & A^{n-1}(\otimes)B(\otimes) \end{bmatrix}.$$

Then, we have the following.

1. *When* $\text{rank}(D(\otimes)) = n$, *the system is observable;*

2. *When* $\text{rank}(L(\otimes)) = n$, *the system is controllable; and*

3. *A sufficient and necessary condition for the system to be asymptotically stable is that the upper bounds of the grey elements of the real parts of the grey characteristic roots of the state grey matrix* $A(\otimes)$ *are all less than zero.*

Example 13.2.1. Assume that a grey linear system is given as follows,

$$\begin{cases} \dfrac{dx_1(t)}{dt} = \otimes_1 x_1(t) + \otimes_2 u_1(t) + \otimes_3 u_2(t) \\ \dfrac{dx_2(t)}{dt} = \otimes_4 x_1(t) \\ y(t) = \otimes_5 x_2(t), \end{cases}$$

where

$$\otimes_1 \in [2,4], \otimes_2 \in [0.8,1.2], \otimes_3 \in [1,3],$$

$$\otimes_4 \in [0.8,1.2], \otimes_5 \in [0.8,1.2].$$

Discuss the controllability, observability, and stability of this system.

Solution: Write the system in matrix form:

$$\begin{cases} \dot{X}(t) = A(\otimes)X(t) + B(\otimes)U(t) \\ Y(t) = C(\otimes)X(t), \end{cases}$$

where

$$X(t) = [x_1(t), x_2(t)]^T, \ U(t) = [u_1(t), u_2(t)]^T,$$
$$Y(t) = y(t);$$

and

$$A(\otimes) = \begin{bmatrix} \otimes_1 & 0 \\ \otimes_4 & 0 \end{bmatrix}, B(\otimes) = \begin{bmatrix} \otimes_2 & \otimes_3 \\ 0 & 0 \end{bmatrix},$$
$$C(\otimes) = \begin{bmatrix} 0 & \otimes_5 \end{bmatrix}.$$

1. Controllability. Let

$$\begin{aligned} L(\otimes) &= \begin{bmatrix} B(\otimes) & A(\otimes)B(\otimes) \end{bmatrix} \\ &= \begin{bmatrix} \otimes_2 & \otimes_3 & \otimes_1 \cdot \otimes_2 & \otimes_1 \cdot \otimes_3 \\ 0 & 0 & \otimes_4 \cdot \otimes_2 & \otimes_4 \cdot \otimes_3 \end{bmatrix}. \end{aligned}$$

From

$$\otimes_1 \cdot \otimes_2 \in [2,4] \cdot [0.8,1.2] = [1.6,4.8],$$

$$\otimes_1 \cdot \otimes_3 \in [2,4] \cdot [1,3] = [2,12],$$

$$\otimes_4 \cdot \otimes_2 \in [0.8,1.2] \cdot [0.8,1.2] = [0.64,1.44],$$

and

$$\otimes_4 \cdot \otimes_3 \in [0.8,1.2] \cdot [1,3] = [0.8,3.6],$$

it follows that the determinant of the sub-matrix

$$\det \begin{bmatrix} \otimes_3 & \otimes_1 \cdot \otimes_2 \\ 0 & \otimes_4 \cdot \otimes_2 \end{bmatrix} = \otimes_3 \cdot \otimes_4 \cdot \otimes_2$$

$$\in [1,3] \cdot [0.64,1.44] = [0.64,4.32]$$

can never take zero as its whitenization value. So,

$$\text{rank}(L(\otimes)) = 2.$$

From Theorem 13.2.1, it follows that the system is controllable.

2. Observability. Let

$$D(\otimes) = \begin{bmatrix} C(\otimes) \\ \\ C(\otimes)A(\otimes) \end{bmatrix} = \begin{bmatrix} 0 & \otimes_5 \\ \\ \otimes_5 \cdot \otimes_4 & 0 \end{bmatrix}.$$

Then

$$\det\left(D(\otimes)\right) = -\otimes_5 \cdot \otimes_5 \cdot \otimes_4 \in [-1.728, -0.512],$$

which will never take zero as its whitenization value. So,

$$\text{rank}(D(\otimes)) = 2.$$

From Theorem 13.2.1, it follows that the system is observable.

3. Stability. We have

$$\begin{aligned} \det\left[\lambda(\otimes)E - A(\otimes)\right] &= \det \begin{bmatrix} \lambda(\otimes) - \otimes_1 & 0 \\ \\ -\otimes_4 & \lambda(\otimes) \end{bmatrix} \\ &= \left[\lambda(\otimes) - \otimes_1\right] \cdot \lambda(\otimes). \end{aligned}$$

Solving

$$\det\left[\lambda(\otimes)E - A(\otimes)\right] = 0,$$

gives that

$$\lambda_1(\otimes) = \otimes_1,\ \lambda_2(\otimes) = 0.$$

Because $\otimes_1$ and 0 are real grey characteristic roots, and the lower bound of $\otimes_1$ is 2, which is greater than 0, the system is unstable.

13.3 Grey Transfer Functions and Special Links

In this section, we show how the concept of Laplace transformations can be applied to the study of grey control systems.

Definition 13.3.1. *Assume that the mathematical model for the nth-order linear system with grey parameters is*

$$\otimes_n \frac{d^n x}{dt^n} + \otimes_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \cdots + \otimes_1 \frac{dx}{dt} + \otimes_0 x = \otimes \cdot u(t)$$

Apply Laplace transformation to both sides of the equation, and denote

$$L[x(t)] = X(s), L[u(t)] = U(s).$$

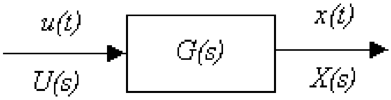

FIGURE 13.3. Relationship between the drive and response

Then

$$G(s) = \frac{X(s)}{U(s)} = \frac{\otimes}{\otimes_n s^n + \otimes_{n-1} s^{n-1} + \cdots + \otimes_1 s + \otimes_0}$$

is called a grey transfer function.

Each grey transfer function is the ratio of the Laplace transformation of the response $x(t)$ of the nth-order linear grey system and the Laplace transformation of the driving term $u(t)$. A grey system, which is described with an equation, is also called a *grey link* or *grey component*. When the transfer function of a certain link is known, the Laplace transformation of the response term can be solved through

$$X(s) = G(s) \cdot U(s)$$

using the Laplace transformation of the driving term. Now, by using inverse transformation, the response $x(t)$ can be obtained. The relationship between the drive and response is shown in Figure 13.3.

In the following, we discuss several typical transfer functions.

Definition 13.3.2. *The link or component with the driving term $u(t)$ and the response term $x(t)$ satisfying*

$$x(t) = K(\otimes)u(t)$$

is called a grey proportional link or component, where $K(\otimes)$ is a grey amplifying coefficient.

Proposition 13.3.1. *The transfer function of a grey proportional link is given by*

$$G(s) = K(\otimes).$$

The characteristics of a grey proportional link are that when a jump occurs in the driving quantity, the response value changes proportionally. This kind of change and relationship between the drive and response are shown in Figure 13.4.

Definition 13.3.3. *With a unit jump occurring in the drive, if the response satisfies*

$$x(t) = K(\otimes) \cdot \left(1 - e^{-tT}\right),$$

then the link is called a grey inertia link, where T is a time constant of the link.

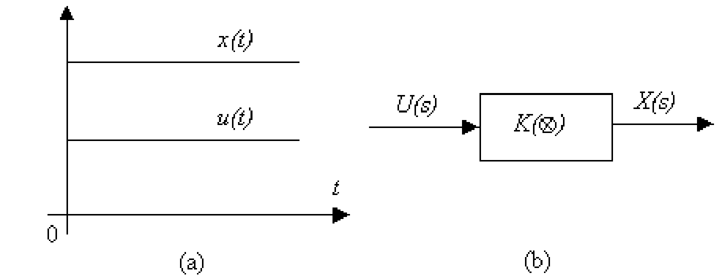

FIGURE 13.4. Relation of change and a relation between drive and response

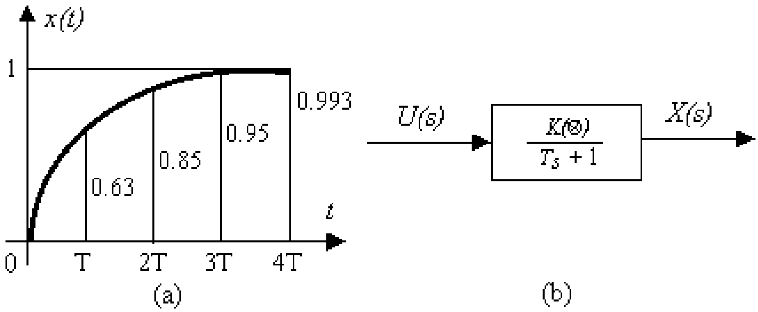

FIGURE 13.5. Curve of change and block diagram of a grey inertia link

Proposition 13.3.2. *The transfer function of a grey inertia link is given by*

$$G(s) = \frac{K(\otimes)}{Ts+1}.$$

The characteristics of a grey inertia link are that when a jump occurs in the driving quantity, the response needs a period of time to reach a new equilibrium state. Figure 13.5 gives a block diagram and the curve of change of the response of a grey inertia link when $\widetilde{K}(\otimes) = 1$.

Definition 13.3.4. *When the drive and response satisfy the relationship*

$$x(t) = \int K(\otimes)u(t)dt,$$

the link is called a grey integral link.

Proposition 13.3.3. *The transfer function of a grey integral link is given by*

$$G(s) = \frac{K(\otimes)}{s}.$$

For grey integral links, when a jump occurs in the drive, the response is

$$x(t) = K(\otimes) \cdot u \cdot t,$$

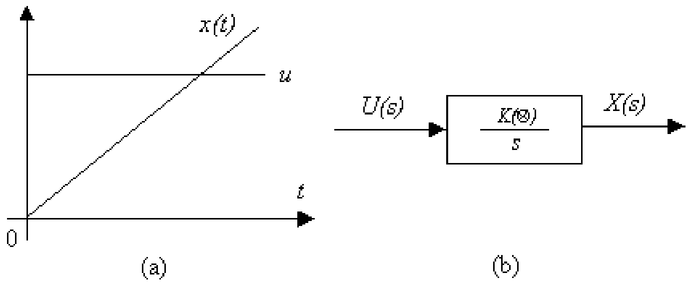

FIGURE 13.6. The block diagram and the response of a grey integral link with a drive of jumps

which is shown in Figure 13.6.

Definition 13.3.5. *When the drive and response satisfy*

$$x(t) = K(\otimes) \cdot \frac{du(t)}{dt},$$

the link is called a grey differential link.

Proposition 13.3.4. *The transfer function of a grey differential link is given by*

$$G(s) = K(\otimes) \cdot s.$$

The characteristics of a grey differential link are that when a jump occurs in the drive, the response is an impulse with an infinite amplitude.

Definition 13.3.6. *When the drive and response satisfy the relationship*

$$x(t) = u[t - \tau(\otimes)],$$

the link is called a grey postponing link, where $\tau(\otimes)$ *is a constant.*

Proposition 13.3.5. *The transfer function of a grey postponing link is given by*

$$G(s) = e^{-\tau(\otimes)\cdot s}.$$

For grey postponing links, when a jump occurs in the drive, it takes a period of time for the response to have a corresponding change. See Figure 13.7 for more details.

In the previous paragraphs, we only listed some typical and elementary links, combinations of which many complicated components or systems can be seen as. For example, a combination of a grey proportional link and a grey differential link can be used to construct a grey proportional differential link. A combination of a grey integral link and grey postponing link can produce a grey integral postponing link. A second level combination can also produce a proportional, differential, integral, and postponing link.

We can study systems problems of stability, controllability, etc., through the study of extremum points of the relevant grey transfer functions. From

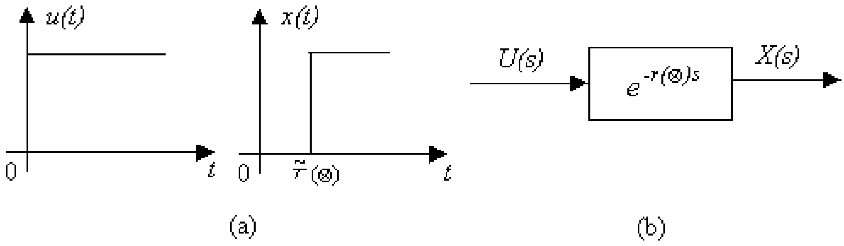

FIGURE 13.7. The block diagram and relationship between the drive and response of a grey postponing link

Theorem 13.3.1 below, it follows that each nth-order grey linear system can be converted to an equivalent first order grey linear system. Therefore, we can make use of the results in Section 13.2 to discuss problems of nth-order grey linear systems.

Theorem 13.3.1. *For the nth-order grey linear system given in Definition 13.3.1, there exists an equivalent first-order grey linear system.*

Proof: Assume that an nth-order grey linear system is given in the form

$$\otimes_n \frac{d^n x}{dt^n} + \otimes_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \cdots + \otimes_1 \frac{dx}{dt} + \otimes_0 x = \otimes \cdot u(t).$$

Let

$$x = x_1, \frac{dx}{dt} = \frac{dx_1}{dt} = x_2, \frac{d^2 x}{dt^2} = \frac{dx_2}{dt} = x_3,$$

$$\cdots, \ \frac{d^{n-1} x}{dt^{n-1}} = \frac{dx_{n-1}}{dt} = x_n.$$

Then, it follows that

$$\frac{dx_n}{dt} = -\frac{\otimes_0}{\otimes_n} x_1 - \frac{\otimes_1}{\otimes_n} x_2 - \frac{\otimes_2}{\otimes_n} x_3 - \cdots - \frac{\otimes_{n-1}}{\otimes_n} x_n + \frac{1}{\otimes_n} u(t).$$

So, the nth-order system is converted to a first-order system:

$$\dot{X}(t) = A(\otimes) X(t) + B(\otimes) U(t),$$

where

$$X(t) = [x_1, x_2, ..., x_n]^T, U(t) = u(t),$$

and

$$A(\otimes) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{\otimes_0}{\otimes_n} & -\frac{\otimes_1}{\otimes_n} & -\frac{\otimes_2}{\otimes_n} & \cdots & -\frac{\otimes_{n-1}}{\otimes_n} \end{bmatrix},$$

$$B(\otimes) = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \\ \frac{1}{\otimes_n} \end{bmatrix}. \square$$

13.4 Matrices of Grey Transfer Functions

Assume that a grey linear control system is given as follows,

$$\begin{cases} \dot{X}(t) = A(\otimes)X(t) + B(\otimes)U(t) \\ Y(t) = C(\otimes)X(t). \end{cases}$$

An application of Laplace transformation gives

$$\begin{cases} sX(s) = A(\otimes)X(s) + B(\otimes)U(s) \\ Y(s) = C(\otimes)X(s). \end{cases}$$

So, we have

$$\begin{cases} [sE - A(\otimes)]X(s) = B(\otimes)U(s) \\ Y(s) = C(\otimes)X(s). \end{cases}$$

If

$$[sE - A(\otimes)]$$

is nonsingular, then we further have

$$\begin{cases} X(s) = [sE - A(\otimes)]^{-1}B(\otimes)U(s) \\ Y(s) = C(\otimes)X(s). \end{cases}$$

That is,

$$Y(s) = C(\otimes)[sE - A(\otimes)]^{-1}B(\otimes)U(s).$$

Definition 13.4.1. *The $m \times n$ matrix*

$$G(s) = C(\otimes)[sE - A(\otimes)]^{-1}B(\otimes)$$

is called the matrix of transfer functions of the grey linear control system, or grey transfer matrix for short.

Definition 13.4.2. *For an nth-order grey linear system, when the state grey matrix $A(\otimes)$ of the corresponding equivalent first-order system is non-singular,*

$$\lim_{s\to 0} G(s) = -C(\otimes)A(\otimes)^{-1}B(\otimes)$$

is called a grey gain matrix.

If the grey gain matrix

$$-C(\otimes)A(\otimes)^{-1}B(\otimes)$$

is used to approximate the grey transfer matrix $G(s)$, then the system is simplified to a proportional link.

From

$$Y(s) = G(s)U(s),$$

it follows that when $m = s = n$, if $G(s)$ is nonsingular, we can also obtain

$$U(s) = G(s)^{-1}Y(s).$$

Definition 13.4.3. *The following*

$$G(s)^{-1} = B(\otimes)^{-1}\,[sE - A(\otimes)]\,C(\otimes)^{-1}$$

is called a grey structure matrix.

Under the condition that the grey structure matrix is known, in order for the output vector $Y(s)$ to reach or get closer to a predetermined objective $J(s)$, we can use $G(s)^{-1}J(s)$ to determine the system's control vector $U(s)$.

We can also discuss the controllability and observability of systems through uses of grey transfer matrices.

13.5 Control with Abandonment

The dynamical characteristics of grey systems are mainly determined by the grey transfer matrix $G(s)$. Therefore, in order to control effectively the dynamical characteristics of a system, a feasible method will be to modify and change the transfer matrix or the structure matrix.

Definition 13.5.1. *Assume that $G(s)^{-1}$ is a system's structure matrix, and $G_*(s)^{-1}$ an objective structure matrix. Then,*

$$\Delta^{-1} = G(s)^{-1} - G_*(s)^{-1}$$

is called a structural deviation matrix.

From

$$G(s)^{-1}Y(s) = U(s)$$

and

$$G_*(s)^{-1} = G(s)^{-1} + \Delta^{-1},$$

it follows that

$$\left[G_*(s)^{-1} - \Delta^{-1}\right] Y(s) = U(s).$$

That is,

$$G_*(s)^{-1}Y(s) - \Delta^{-1}Y(s) = U(s).$$

Definition 13.5.2. *The following*

$$-\Delta^{-1}Y(s)$$

is called a $superfluous$ term. The control through a feedback of $\Delta^{-1}Y(s)$ to cancel the superfluous term is called a control with abandonment.

The system

$$G(s)^{-1}Y(s) = U(s)$$

by a feedback of $\Delta^{-1}Y(s)$, can be converted to

$$G(s)^{-1}Y(s) + \Delta^{-1}Y(s) = U(s);$$

that is,

$$\left[G(s)^{-1} + \Delta^{-1}\right] Y(s) = U(s).$$

Therefore,

$$G_*(s)^{-1}Y(s) = U(s)$$

already has the desirable objective structure.

The number of entries in the structural deviation matrix Δ^{-1}, used in a control with abandonment, directly affects the number of components in the controlling equipment. From the angles of economics, reliability, being easy to realize technically, etc., under the guarantee that the system will possess

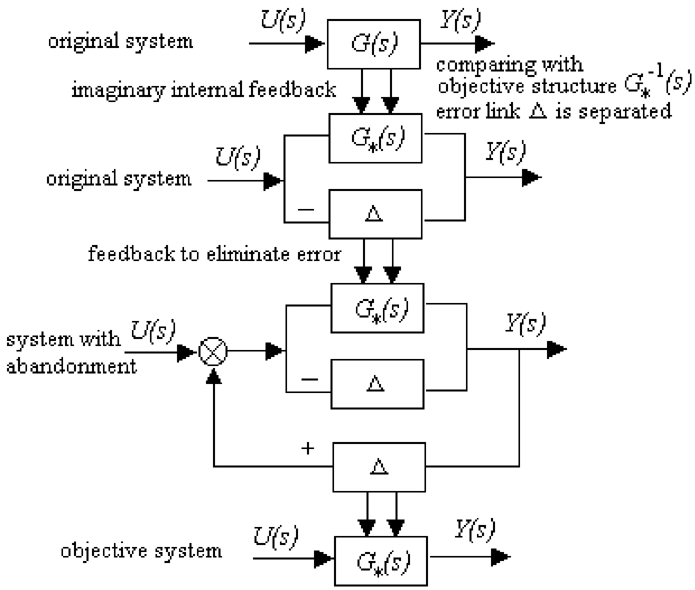

FIGURE 13.8. Flowchart for the idea of control with abandonment

desirable dynamical characteristics, we always try to keep the number of entries in the deviation matrix Δ^{-1} to a minimum. That is to say, in the objective structural matrix, we should try to keep the corresponding entries of the original structural matrix.

The concept, embedded in the control with abandonment, can be well depicted with the block diagram of Figure 13.8.

13.6 Control of Grey Incidences

In this section, we learn how to employ the concept of grey incidences in problems of grey control.

Definition 13.6.1. *Assume that*

$$Y = (y_1, y_2, ..., y_m)^T$$

is the output vector, and

$$J = (j_1, j_2, ..., j_m)^T$$

the objective vector. If the components of the control vector

$$U = (u_1, u_2, ..., u_s)^T$$

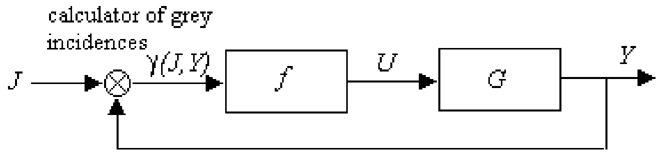

FIGURE 13.9. The block diagram for grey incidence control systems

satisfy

$$u_k = f_k\left[\gamma(J,Y)\right],$$

$k = 1, 2, ..., s$, *where* $\gamma(J,Y)$ *is the degree of grey incidence between the output vector* Y *and the objective vector* J, *then the systems control is called a grey incidence control.*

Grey incidence control systems are obtained by attaching grey incidence controllers to regular control systems. It determines the control vector U through the degree of grey incidence $\gamma(J,Y)$ so that the degree of incidence between the output vector and the objective vector does not go beyond a certain predetermined range.

Grey incidence control systems are shown graphically in Figure 13.9.

13.7 Control of Grey Predictions

All the controls discussed previously are done through checks to see whether the system's behavioral sequence satisfies some predetermined requirements. There obviously exist the following shortcomings with these after-event controls.

1. Results cannot be applied to prevent predicted disasters in the future.
2. There is no way to perform an on-time control.
3. Adaptability is not very strong.

The idea of *control of grey predictions* is used to predict a system's future behaviors based on a collection of data regarding the system's behavior in order to uncover the development law, if any, of the system, and to perform precontrols on relevant controlling decisions, by using the predicted future development tendency of the system. In this way, it becomes possible for us to prevent a predicted disaster before it actually occurs, and to impose controls in a timely fashion. Therefore, this method has a relatively stronger adaptability in practical applications.

A general grey prediction control system is shown in Figure 13.10.

The principle behind the concept of grey prediction control systems is: First, through the use of sampling equipment, collect and organize data for the output vector Y. Second, through the equipment of prediction, establish

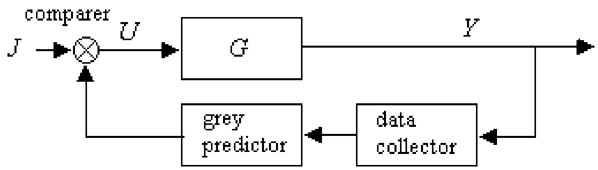

FIGURE 13.10. The flowchart for a general grey prediction control system

a model, from which predicted values after several steps are computed. Third, check with the objective, determine the control vector U, so that the future output vector Y will be as close to the objective J as possible.

Definition 13.7.1. *Assume that* $j_i(k)$, $y_i(k)$, *and* $u_i(k)$, $i = 1, 2, ..., m$, *are objective components, output components, and control components at time moment* k, *respectively. For* i = 1, 2, ..., m, *let*

$$j_i = (j_i(1), j_i(2), ..., j_i(n)),$$

$$y_i = (y_i(1), y_i(2), ..., y_i(n)),$$

and

$$u_i = (u_i(1), u_i(2), ..., u_i(n)).$$

For the control operator

$$f: \quad [j_i(\lambda), y_i(\lambda)] \to u_i(k),$$

that is,

$$u_i(k) = f[j_i(\lambda), y_i(\lambda)],$$

we have

1. *When* $k > \lambda$, *the system is an* $after$-$event$ $control$;
2. *When* $k = \lambda$, *the system is said to be an on-time control;* and
3. *When* $k < \lambda$, *the system is said to be a prediction control.*

Definition 13.7.2. *If the operator* f, *as defined in Definition 13.7.1, satisfies*

$$f[j_i(\lambda), y_i(\lambda)] = j_i(\lambda) - y_i(\lambda),$$

that is,

$$u_i(k) = j_i(\lambda) - y_i(\lambda),$$

then

1. *When* $k > \lambda$, *the system is said to be an error-*$afterward$ *control;*
2. *When* $k = \lambda$, *the system is said to be error-on-time control;* and
3. *When* $k < \lambda$, *the system is said to be error-prediction control.*

Definition 13.7.3. *Assume that*

$$y_i = (y_i(1), y_i(2), ..., y_i(n)),$$

for $i = 1, 2, \cdots, m$, *is an observational sequence of the output components, whose GM(1, 1) response sequence is*

$$\begin{cases} \widehat{y}_i^{(1)}(k+1) = \left[y_i(1) - \frac{b_i}{a_i}\right] \cdot e^{-a_i k} + \frac{b_i}{a_i} \\ \widehat{y}_i^{(0)}(k+1) = \widehat{y}_i^{(1)}(k+1) - \widehat{y}_i^{(1)}(k). \end{cases}$$

If the control operator f *satisfies*

$$u_i(n+k_0) = f\left[j_i(k),\ \widehat{y}_i^{(0)}(k)\right],\ n+k_0 < k,$$

$i = 1, 2, ..., m$, *then the system's control is called a grey prediction control.*

In a grey prediction control system, we often apply metabolic models to do predictions. Hence, the parameters of the prediction equipment vary with time. When a new data value is collected and is accepted by the sampling equipment, an older data value will be deleted so that a newer model is established, and a series of new predicted values will appear accordingly. This end guarantees a strong adaptability of the system.

13.8 Practical Applications

In this section, we look at three examples to see how our theory of grey control can be practically applied.

Example 13.8.1. Let us look at a grey control with abandonment of a boring lathe.

The block diagram of the input system of a horizontal precision boring lathe of the type $T618B$ is shown in Figure 13.11, and the dynamical block diagram of the system is shown in Figure 13.12.

Based on the work requirements and technical conditions of the boring lathe, the allowed region for extremum points is given as follows,

$$S^* = \{(u, v) \mid -\infty < u < -\delta,\ \ -\rho < v < \rho\},$$

where $|\delta| = 0.2$ and $|\rho| = 0.3$.

Before the system is equipped with additional modifying components, the characteristic equation is

$$f(s) = s^3 + 2.42 \cdot 10^2 s^2 + 5.1 \cdot 10^3 s + 1.3 \cdot 10^5 = 0$$

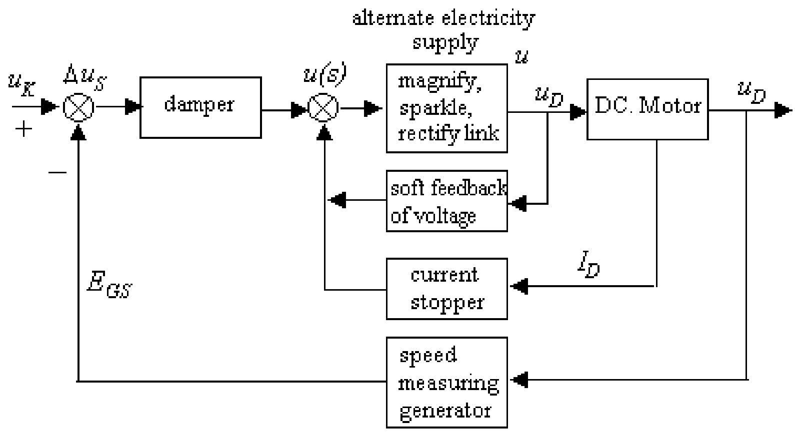

FIGURE 13.11. The block diagram of the input system of a horizontal precision boring lathe of type T618B

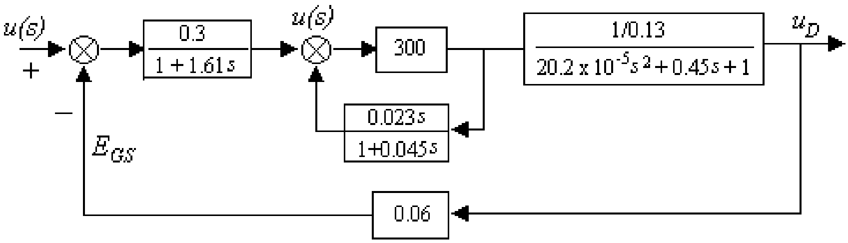

FIGURE 13.12. The dynamical block diagram of the input system of a horizontal precision boring lathe of type T618B

with characteristic roots

$$s_1 = -10.1821 + 21.9743i,$$

$$s_2 = -10.1821 - 21.9743i,$$

and

$$s_3 = -221.6357.$$

Obviously, $s_j \notin S^*$, for $j = 1, 2, 3$.

If the closed loop transfer function for prediction, satisfying the requirement S^*, is given as follows,

$$G_*(s) = \frac{21.2 \cdot 10^5}{s^3 + 3.4 \cdot 10^4 s^2 + 2.6 \cdot 10^4 s + 1.3 \cdot 10^5},$$

from the principle of control with abandonment, it then follows that the treated control with abandonment is

$$\frac{0.023s}{1 + 0.045s} = \frac{u_s}{u_D}.$$

After adding an appropriate term, the system's characteristic equation becomes

$$\begin{aligned} f^*(s) = \quad & s^4 + 0.2 \cdot 10^3 s^3 + 5 \cdot 10^3 s^2 + \\ & +5.2 \cdot 10^3 s + 27.5 \cdot 10^3 = 0 \end{aligned}$$

with characteristic roots

$$s_1^* = -0.42 + 2.35i, s_2^* = -0.42 - 2.35i,$$

$$s_3^* = -170.92, \text{ and } s_4^* = -28.24.$$

Obviously, $s_j \in S^*$, for $j = 1, 2, 3, 4$. So, the system possesses all the desirable properties.

Now, let us look at the rotation system of the main shaft, where there exist relatively large differences between the parameters and the input system; see Table 13.2 more details.

Table 13.2. Comparison between the rotation and the input systems

	Motor Type	Capacity (Kilowatts)	Rotat. Speed (Rotat./minute)	Current (Ampere)
Rotation system	Z_2-52	7.5	1500	41
Input system	Z_2-32	2.2	1500	12.5
	Resistance (ohm)	Electromagnetic time constant (sec.)	Mechanical and electrical time constant (sec.)	
Rotation system	0.78	0.006	0.065	
Input system	1.8	0.0045	0.045	

Before the control of abandonment is applied, the characteristic equation of the main shaft system is

$$f(s) = s^3 + 167.2878s^2 + 2667.6s + 70393.37 = 0,$$

with characteristic roots

$$s_1 = -7.2198 + 20.209i, s_1 = -7.2198 - 20.209i,$$

$$s_3 = -152.848.$$

After applying the controller transfer function with abandonment

$$\frac{0.023s}{1 + 0.045s},$$

which is identical to that of the input system, and applying control with abandonment, the characteristic equation is converted to

$$\begin{aligned} f^*(s) &= s^4 + 0.1674 \cdot 10^3 s^3 + 2.6916 \cdot 10^3 s^2 + \\ &\quad +2.4224 \cdot 10^3 s + 10.1353 \cdot 10^3 = 0 \end{aligned}$$

with characteristic roots

$$s_1^* = -0.343 + 1.9547i, s_2^* = -0.343 - 1.9547i,$$

$$s_3^* = -149.5, \text{ and } s_4^* = -17.2.$$

Obviously, $s_j^* \in S^*$, for $j = 1, 2, 3, 4$.

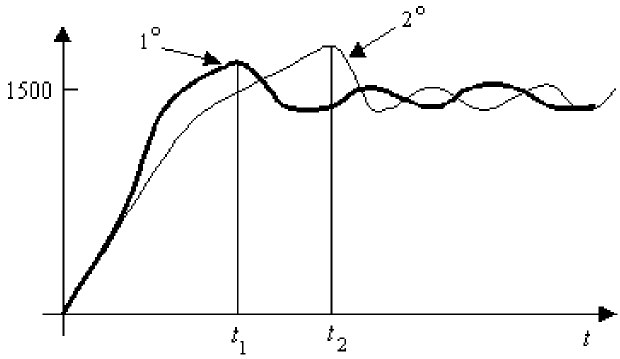

FIGURE 13.13. The dynamical responses of the main shaft system and the input system

Even though the parameters in the main shaft system and the input system are different, satisfactory properties are obtained by applying the same controller with abandonment. Their dynamical responses are shown in Figure 13.13.

In Figure 13.13, 1^o stands for the input system, with extra adjustment amount of about 150 rotations per minute, and the adjustment time t_1 being approximately 0.8 second. The curve 2^o represents the main shaft system, with extra adjustment amount of about 140 rotations per minute and adjustment time t_2 being approximately 0.9 seconds. □

Example 13.8.2. Let us look at a control of our biological prevention system of cotton aphids.

Cotton aphids are injurious insects to cotton production. Ladybugs are a natural enemy of cotton aphids. By planting rape in cotton fields, an existing biological prevention system of cotton aphids can be effectively controlled.

Ladybugs eat not only cotton aphids, but also rape aphids. At first, plant rape in cotton fields so that rape aphids can grow so that ladybugs will fly, stay, and be reproduced with the rape. When the growth of ladybugs has reached a certain level, cotton and cotton aphids start to grow. At this time, cut down the planting of the rape so that the ladybugs have to turn their attention to the cotton aphids to realize the objective of destroying the cotton aphids.

Suppose that $x_1(k)$, $x_2(k)$, and $x_3(k)$ stand for the numbers of ladybugs, rape aphids, and cotton aphids at the kth time phase, respectively. Then, these three numbers satisfy the following relation

$$\begin{cases} x_1(k+1) = a_{11}x_1(k) + a_{12}x_2(k) \\ x_2(k+1) = -a_{21}x_1(k) + a_{22}x_2(k) + \otimes_{23}x_3(k) \\ x_3(k+1) = -a_{31}x_1(k) + \otimes_{32}x_2(k) + a_{33}x_3(k). \end{cases}$$

That is,

$$X(k+1) = A(\otimes)X(k),$$

where

$$X(k+1) = \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix}, \; X(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

and

$$A(\otimes) = \begin{bmatrix} a_{11} & a_{12} & 0 \\ -a_{21} & a_{22} & \otimes_{23} \\ -a_{31} & \otimes_{32} & a_{33} \end{bmatrix}.$$

By taking $k = 0$, we have

$$X(1) = A(\otimes)X(0).$$

When $k = 1$, we have

$$X(2) = A(\otimes)X(1) = A(\otimes)^2X(0).$$

If we need to eliminate cotton aphids in the second phase, let $x_3(2) = 0$. From the preceding equation, it follows that

$$\begin{aligned} x_3(2) = \; & -[a_{31}a_{11} + a_{21}\otimes_{32} + a_{31}a_{33}]\,x_1(0) + \\ & +[-a_{31}a_{12} + a_{22}\otimes_{32} + a_{33}\otimes_{32}]x_2(0) \\ & +[\otimes_{32}\otimes_{23} + a_{33}^2]x_3(0). \end{aligned}$$

So,

$$\begin{aligned} & -[a_{31}a_{11} + a_{21}\otimes_{32} + a_{31}a_{33}]\,x_1(0) + \\ & +[-a_{31}a_{12} + a_{22}\otimes_{32} + a_{33}\otimes_{32}]x_2(0) \\ & +[\otimes_{32}\otimes_{23} + a_{33}^2]x_3(0) = 0. \end{aligned}$$

Let

$$\lambda_{13}(0) = \frac{x_1(0)}{x_3(0)} \text{ and } \lambda_{23}(0) = \frac{x_2(0)}{x_3(0)},$$

and take $\otimes_{32} = 0$. Then, it follows that

$$\lambda_{13}(0) = \frac{a_{33}^2}{a_{31}(a_{11} + a_{33})} - \frac{a_{31}a_{12}}{a_{31}(a_{11} + a_{33})}\lambda_{23}(0).$$

Therefore, we can choose the ratio of ladybugs and cotton aphids to satisfy

$$\lambda_{13} \in \left[\frac{a_{33}^2 - a_{31}a_{12}\lambda_{23}(0)}{a_{31}(a_{11} + a_{33})}, \frac{a_{33}^2}{a_{31}(a_{11} + a_{33})}\right].$$

Here λ_{13} is a control quantity of our biological prevention system of cotton aphids. It is a grey number with the following lower limit

$$\underline{\lambda}_{13} = \frac{a_{33}^2 - a_{31}a_{12}\lambda_{23}(0)}{a_{31}(a_{11} + a_{33})},$$

and the following upper limit

$$\overline{\lambda}_{13} = \frac{a_{33}^2}{a_{31}(a_{11} + a_{33})}.$$

If we let μ be the ratio of the production capacity of ladybugs and that of cotton aphids, that is, $\mu = \dfrac{a_{11}}{a_{33}}$, then

$$\overline{\lambda}_{13} = \frac{a_{33}}{a_{31}(1 + \mu)}.$$

Obviously, the greater μ is, the smaller $\overline{\lambda}_{13}$ will be. When μ is fixed, the greater a_{33} is, the greater $\overline{\lambda}_{13}$ will be. When a_{33} and μ are fixed, the greater a_{31} is, the smaller $\overline{\lambda}_{13}$ will be. In general, we take $\lambda_{13} = \overline{\lambda}_{13}$. Then, it is assured that in the second phase, all cotton aphids will be eliminated.

Example 13.8.3. Let us now look at a grey prediction control for a boiler water supply system.

The block diagram for the system control is shown in Figure 13.14, where grey prediction controlling components are applied to the controller of flow and the water level controller so that models can be established quickly with sensitive responses to stochastic disturbances appearing in the environment and the parameters and with strong self-adaptability.

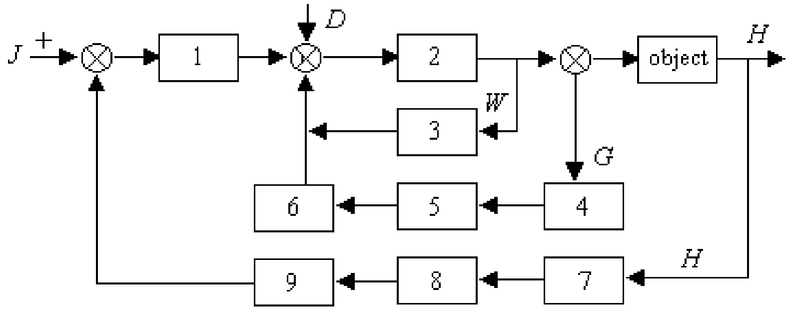

FIGURE 13.14. The block diagram of the system control for the boiler water supply system

In Figure 13.14,

W : Location signal	D : Steam flow
H : Water level	G : Input water flow
1 : (Electrical) potential shifter	2 : Execution component
3 : Location shifter	4 : Flow shifter
5 : GM(1, 1) controller	6 : Proportional decision maker
7 : Water level shifter	8 : GM(1, 1) controller
9 : Proportional decision maker	J : Objective

The properties and characteristics of this system are obviously superior to those of the PID system, which was ranked number one by the Electricity Bureau of Hubei Province. For more details, see Table 13.3.

Table 13.3. Differences between the adjustment times of PID and a grey controller

Disturbance		PID Adjustment Time (Sec.)	Grey Controller Adjustment Time (Sec.)
Internal	+40%	60	3
	−40%	63	4
External	±10%	120∼ 180	40

References

[1] R. L. Ackoff (1973) Science in the Systems Age: Beyond IE. OR. And MS. *Operations Research*, vol. 21, no. 3, pp. 661–671.

[2] D. R. Anderson (1982). *An Introduction to Management Science: Quantitative Approaches to Decision-Making*, St. Paul: West.

[3] L. V. Bertranffy (1968). *General Systems Theory: Foundation, Development, Applications*, New York: George Braziller.

[4] P. Checkland (1981). *Systems Thinking, Systems Practice,* New York: John Wiley.

[5] F. Chen (1986). Grey control of industrial processes. *Journal of Wuhan College of Technology*, vol. 2, no. 1, pp. 11–23.

[6] H-S. Chen and W-C Chang (2000). Optimization of GM(1,N) modeling. *Journal of Grey Systems*, vol. 12, no. 3.

[7] M-Y. Chen (1982). Grey dynamic control of a boring lathe. *Journal of Huazhong College of Technology*, vol. 10, no. 3, pp. 7–11.

[8] M-Y. Chen (1985). Stability and steadiness problems of grey systems. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 54–58.

[9] Q-B. Chen and R. Nie (1999). Some study on grey GM(1,3) growth model and applications. *Journal of Song-Liao* (Natural Science Edition), no. 4, pp. 32–35.

[10] Y-H. Chen (1999). Unequal difference time series grey systems and applications. *Journal of Jing-Zhou Normal College*, no. 5, pp. 30–32.

[11] Y-K. Chen (1996). GM(1,1) model applied in health administration. In *Current State of Grey Systems Research*. Wuhan: Huazhong University of Science and Technology Press, pp. 151–152.

[12] Y-K. Chen and X-R. Tan (1995). Grey relational analysis on serum markers of liver fibrosis, *The Journal of Grey Systems*, vol. 7, no. 1, pp. 63–68.

[13] Y-Y. Chen (1984). *Fuzzy Mathematics*. Wuhan: Huazhong University of Science and Technology Press.

[14] Z-B. Chen (1984). Remnant recursive recognition of grey interval functions. *Journal of Huazhong College of Technology*, vol. 12, no. 1, pp. 67–70.

[15] Z-B. Chen and J-G. Liu (1984). Float of extrema of grey systems state feedback. *Journal of Huazhong College of Technology*, vol. 12, no. 3, pp. 8–14.

[16] Q-Y. Cheng and W-H. Qiu (2001). Grey system forecast for firing accuracy of gun. *Journal of Systems Science and Systems Engineering*, vol. 10, no. 2.

[17] X-B. Ci (1995). Some methods for improved accuracy of GM(1,1) predictions. *Journal of Sho-Guan University* (Natural Science Edition), no. 2, pp. 34–39.

[18] G. W. Churchman (1979). *The Systems Approach and Its Enemies*, New York: Basic.

[19] H-Q. Cuo, L-S. Kong, and M-P. Chen (1999). Evolutionary models for dynamic systems. *Research and Development of Computers*, no. 8, pp. 923–931.

[20] W-Z. Dai (1997). nth accumulating grey models with remnant corrections. *Systems Engineering Theory and Practice*, no. 12, pp. 121–124.

[21] E. D. David (1975). *Science Development: An Evaluation Study*, Washington: NAS Pub. Office.

[22] J-L. Deng (1984). Grey dynamic modeling and its application in long-term prediction of food productions. *Exploration of Nature*, vol. 3, no. 3, pp. 7–43.

[23] J-L. Deng (1985). Grey systems incidence space. *Fuzzy Mathematics*, vol. 5, no. 2, pp.3 –10.

[24] J-L. Deng (1985). Grey systems generation functions. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 11–22.

[25] J-L. Deng (1985). GM(1,1) model in grey systems theory. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 23–32.

[26] J-L. Deng (1985). Five grey predictions. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 33–42.

[27] J-L. Deng (1985). Decision making for grey situations. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 43–50.

[28] J-L. Deng (1985). *Grey Systems (social and economic)*. Beijing: National Defense Industry Press.

[29] J-L. Deng (1985). *Grey Control Systems*. Wuhan: Huazhong University of Science and Technology Press.

[30] J-L. Deng (1986). *Grey Prediction and Decision Making*. Wuhan: Huazhong University of Science and Technology Press.

[31] J-L. Deng (1987). *Fundamental Methods of Grey Systems*. Wuhan: Huazhong University of Science and Technology Press.

[32] J-L. Deng (1987). Grey exponential laws of accumulating generation – Problem on optimally treating information using grey control systems. *Journal of Huazhong University of Science and Technology*, vol. 12, no. 5, pp. 11–16.

[33] J-L. Deng (1988). *Grey Colored Multi-Dimensional Programming*. Wuhan: Huazhong University of Science and Technology Press.

[34] J-L. Deng (1988). *A Course on Grey Systems Theory*. Wuhan: Huazhong University of Science and Technology Press.

[35] J-L. Deng (1989). *Selected Papers in Grey Systems Theory*. Beijing: Ocean Press.

[36] J-L. Deng (1989). Grey information space, *The Journal of Grey Systems*, vol. 1, no. 2, pp. 103–118.

[37] J-L. Deng (1989). Introduction to grey gystem theory, *The Journal of Grey Systems*, vol. 1, no. 1, pp. 1–24.

[38] J-L. Deng (1990). The point-set topological properties of grey series, *The Journal of Grey Systems*, vol. 2, no. 4, pp. 283–295.

[39] J-L. Deng (1990). *A Course in Grey Systems Theory*. Wuhan: Huazhong University of Science and Technology Press.

[40] J-L. Deng (1991). Representations of information in grey systems theory. *Grey Systems Theory and Applications*, vol. 1, no. 1, pp. 1–12.

[41] J-L. Deng (1991). The unit of information representation in grey system theory, *The Journal of Grey Systems*, vol. 3, no. 2, pp. 87–106.

[42] J-L. Deng (1992). *Introduction to Grey Numbers – Grey Fuzzy Sets*. Wuhan: Huazhong University of Science and Technology Press.

[43] J-L. Deng (1992). Comparability of representations of information. *Grey Systems Theory and Applications*, vol. 2, no. 1, pp. 1–10.

[44] J-L. Deng (1992). Grey Units, *The Journal of Grey Systems*, vol. 4, no. 1, pp. 3–11.

[45] J-L. Deng (1992). Grey Causality and Information Cover, *The Journal of Grey Systems*, vol. 4, no. 2, pp. 99–119.

[46] J-L. Deng (1992). Grey hazy sets, *The Journal of Grey Systems*, vol. 4, no. 2, pp. 13–29.

[47] J-L. Deng (1992). Grey inclusion in set, *The Journal of Grey Systems*, vol. 4, no. 3, pp. 195–220.

[48] J-L. Deng (1993). Grey differential equation, *The Journal of Grey Systems*, vol. 5, no. 1, pp. 1–14.

[49] J-L. Deng (1995). Extent information cover in grey system theory, *The Journal of Grey Systems*, vol. 7, no. 2, pp. 131–138.

[50] J-L. Deng (1996). Several problems in grey systems theory and applications. In *New Development in Grey Systems Research*, ed. by S-F. Liu and Z-X. Xu. Wuhan: Huazhong University of Science and Technology Press, pp. 1–10.

[51] J-L. Deng (2001). Negative power AGO in grey theory. *The Journal of Grey Systems*, vol. 13, no. 3.

[52] J-L. Deng (2001). On boundary of parameters in GM(1,1| tau ,r). *The Journal of Grey Systems*, vol. 13, no. 2.

[53] J-L. Deng (2001). On boundary of grey input b in GM(1,1). *The Journal of Grey Systems*, vol. 13, no. 1.

[54] J-L. Deng and & D. K.W. Ng (1996). Chaos in grey model GM(1,N), *The Journal of Grey Systems*, vol. 8, no. 1, pp. 1–10.

[55] J-L. Deng and C-S. Zhou (1986). Sufficient conditions for the stability of a class of Interconnected Dynamic Systems, *Systems & Control Letters*, vol. 5, no. 2, pp. 105–108.

[56] Q. Deng, Q-X. Feng, and J-L Deng (1985). Underground water control systems and optimal grey decision making. *Journal of Automation*, vol. 12, no. 2, pp. 125–132.

[57] J. Di and J-M. Shen (1997). MG<(1,N) grey model and its applications. *Systems Engineering Theory and Practice*, vol. 17, no. 5, pp. 109–113.

[58] X-Y. Dong (1993). An application of grey multi-target programming in the study of the structure of crop plantings. In *Selected Papers on Grey Systems*, Henan University Press, pp. 219–221.

[59] Y-G. Dong (1994). Studies on grey degree of slope incidence. *Agricultural Systems Science and Synthetic Research*, Vol.10, special issue, pp. 331–337.

[60] Y-G. Dong (1994). Studies on DGM models. *Agricultural Systems Science and Synthetic Research*, Vol. 10, special issue, pp. 387–390.

[61] Y-G. Dong (1995). Valuation of grey clustering with human experience, *The Journal of Grey Systems*, vol. 7, no. 2, pp. 179–184.

[62] Y-G. Dong and S-F. Liu (1995). Comprehensive analysis and optimal design of Henan Provincial transportational system. In *Proceedings of the 1st Annual Conference of Henan Young Scholars*, pp. 96–99.

[63] Y-G. Dong and S-F. Liu (1995). Prediction and suggested handling of surplus farm labors in Henan Province. *Journal of Henan Agricultural University*, vol. 29, no. 4, pp. 369–374.

[64] Y-G. Dong, S-F. Liu, B-J. Li, and Z. Chen (2000). Computational analysis of contribution rate of progress in technology on the first class businesses. *Systems Engineering Theory and Practice*, vol. 20, no. 8, pp. 95–99.

[65] Y-G. Dong and L-H. Wang (1996). Grey clustering analysis for the economy in the country. In *New Progress in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 221–223.

[66] R. A. Dunn (1981). *Management Science: A Practical Approach to Decision-Making*, New York: Macmillan.

[67] F. Engels (1971). *Dialectics of Nature* (in Chinese). Beijing: The Preople's Press.

[68] C-M. Fan, Y-S. Gao, P-H. Wu, and X-G. Xu (1996). Applications of grey systems theory in predictions of Yellow River. In *New Development in Grey Systems Research*, ed. S-F. Liu and Z-X, Xu. Wuhan: Huazhong University of Science and Technology Press, pp. 93–95.

[69] E-H. Fan, H-W. Chen, and Z-C. Qi (1998). Technique of regularity developed on the completeness of grey time series. *Journal of National Defense University of Science and Technology*, no. 4.

[70] J-F. Fan and R-Y. Wang (1995). A simply test for the stability of symmetric grey systems. *Engineering Mathematics*, no. 2, pp. 31–33.

[71] L-Z. Fan, S-F. Liu, and C-C. Li (1996). *An Improvement of BP Neural Network*. Shengyang: Hei-Long-Jiang Science and Technology Press, pp. 66–68.

[72] L-Z. Fan, Z-G. Zhang, and S-F. Liu (1996). *Value of An Investment Opportunity: A Binomial Model*. Shengyang: Hei-Long-Jiang Science and Technology Press, pp. 69–71.

[73] T-G. Fan, Y-C. Zhang, and X-Z. Quan (2001). Grey model GM(1,t) and its chaos feature. *The Journal of Grey Systems*, vol. 13, no. 2.

[74] Z-G. Fang and S-F. Liu (2002). Algorithm model research of the logical cutting tree on the network maximum flow . In *The 12th International WOSC Congress Proceedings. Advances in Systems Science and Application*, vol. 4, no. 1, pp. 132–139.

[75] Y-G. Feng (1992). Grey clustering and comprehensive water pollution evaluation. *Systems Engineering Theory and Practice*, vol. 12, no. 6, pp. 46–48.

[76] Y-C Feng, D. Li, Y-Z Wang and Z-S Cao (2001). A grey model for response time in database system DM3. *Journal of Grey Systems*, vol. 13, no. 3.

[77] Z-Y. Feng (1992). Direct grey modeling. *Journal of Applied Mathematics*, vol. 15, no. 3, pp. 345–354.

[78] J. W. Forrester (1968). *Principles of Systems*, Cambridge: Wright-Allen Press.

[79] P. L. Galison and B. W. Hevly (1991). *Creative Management*, London, Sage.

[80] S-J. Gan (1989). *Soft Sciences in China*. Wuhan: Huazhong University of Science and Technology Press.

[81] D-E. Gong (1988). *Brief Introduction to Economic Control Theory*. Beijing: People's University of China Press.

[82] J-F. Gu (1998). *Systems Engineering and Strategy for Sustainable Development*. Baijing: Literature of Science and Technology Press.

[83] J-F. Gu and F. Gao (1998). Systems thinking on management of large social projects. *Management Science of China*, vol. 16, no. 4, pp. 1–8.

[84] X-P. Guan (1981). *Outline of Science* (in Chinese). Tianjin: Tianjin Press of Science and Technology.

[85] H. Guao (1985). Distinguishing coefficients of grey degree of incidence. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 55–58.

[86] H. Guao, S-H. Ma, and K-J. Xiao (1986). Medical grey clustering diagnosis. *Exploration of Nature*, vol. 5, no. 4, pp. 40–44.

[87] R-L. Guao (1994). Introduction and application of grey breedings. *Agricultural Systems Science and Comprehensive Research*, vol. 10, no. 5, pp. 334–338.

[88] H. Guo (1991). Grey exponential law attributed to mutual complement of series, *The Journal of Grey Systems*, vol. 3, no. 2, pp. 153–162.

[89] Y. Guo, H-Y. Pen, and W-Y. Guo (1997). An improvement on grey modeling of unequal interval time series. *Management and Technology of Environment Monitor*, vol. 9, no. 4, pp. 31–33.

[90] Q. Hao, M-L. Zhu, and B. Li (1996). Ordering of fuzzy grey incidences and applications. *Systems Engineering Theory and Practice*, vol. 16, no. 6, pp. 44–49.

[91] Q-K. Hao (1998). Characteristics and problems of grey systems methods. *Journal of Ji-Lin University of Agriculture*, vol. 20, no. 4, pp. 92–94.

[92] Y-H. Hao and X-M. Wang (2000). Period residual modification of GM(1,1) modeling. *Journal of Grey Systems*, vol. 12, no. 2.

[93] M-X. He (1997). A new method to establish GM(1,1) prediction models. *Agricultural Systems Science and Comprehensive Research*, vol. 13, no. 4, pp. 241–244.

[94] X-J. He and G-Z. Sun (2001). A non-equil gap grey model NGM(1,1). *Journal of Grey Systems*, vol. 13, no. 2.

[95] Y. He (1993). Solution of GM(1,1) models and discussions on relevant accuracy tests. *Systems Engineering*, vol. 11, no. 3, pp. 36–70.

[96] Y. He (1994). Grey multi-level comprehensive evaluation model and its applications. *Systems Engineering Theory and Practice*, vol. 14, no. 4, pp. 72–74.

[97] Y. He and Y-D. Bao (1992). Grey Markov prediction model and its applications. *Systems Engineering Theory and Practice*, vol. 12, no. 4, pp. 59–63.

[98] Y-X. Ho and W. Lin (1995). Combined development model for Henan Provincial Transportation. In *Proceedings of the 1st Henan Provincial Annual Conference of Young Scholars*, pp. 221–223.

[99] Z-L. Ho and Z-Z. Xi (1992). *Input-Output Model for Defense Industry. Applications and Development of Modern Chinese Input-Output*. Beijing: China Statistics Press.

[100] W. Hofkirchner (1998). *The Quest for A Unified Theory of iInformation*. NY: Cooper Station Gordon and Breach.

[101] E. M. Hoover (1975). *An Introduction to Regional Economics*, New York: McGraw-Hill.

[102] C-H. Hsieh (2001). Grey data fitting model and its application to image coding. *Journal of Grey Systems*, vol.13, no.1.

[103] C-H. Hsieh, J-H Chou, andY-J Wu (2000).Taguchi-genetic algorithms for optimizing GM(1,1). *Journal of Grey Systems*, vol. 12, no. 3.

[104] S-G. Hu (1987). *Introduction to Modern Analysis*. Wuhan: Huazhong University of Science and Technology Press.

[105] S-H. Hu (1996). Multi-level grey evaluations of subjective index evaluation. *Systems Engineering Theory and Practice*, vol. 16, no. 1, pp. 12–20.

[106] Z-J. Hu (1987). *Introduction to Unbalanced Systems Economics*. Chang Zhueng: Ji-Lin People's Press.

[107] Z-W. Hu (1987). Food production tendencies and drought and flood disasters at Guong Shan County. *Agricultural Systems Science and Comprehensive Research*, vol. 3, no. 1, pp. 28–31.

[108] Y-P. Huang and H-C. Chu (1996). Practical consideration for grey modeling and its application to image processing, *The Journal of Grey Systems*, vol. 8, no. 3, pp. 217–233.

[109] F. Y. Huong (1994). Transformation method of grey systems modeling. *Systems Engineering Theory and Practice*, vol. 14, no. 6, pp. 35–38.

[110] H-F. Jia and Y-Q. Zhen (1998). Grey time series prediction model and applications in annual precipitation predictions. *Systems Engineering Theory and Practice*, vol. 18, no. 8, pp. 42–45.

[111] C-W. Jiang, Z-J. Hou, and Y-C. Zhang (2000). Chaos analysis of the grey forecasting model. *Journal of Grey Systems*, vol. 12, no. 4.

[112] C-Y. Jiang (1996). Grey Markov prediction model. *Journal of Zhong-Qing Construction University*, vol. 18, no. 3, pp. 116–122.

[113] P-R. Jie, X-Y. Hu, and D-Q. Xiong (1999). Some analysis and evaluation of grey prediction models. *Science of Hydraulic and Electric Energy Resources*, vol. 17, no. 2, pp. 42–44.

[114] P-S. Jie, W-S. Huong, and Y-Y. Hu (2001). Some study on the characteristics of grey prediction models. *Systems Engineering Theory and Practice*, vol. 21, no. 9, pp. 105–108.

[115] H. Jing (1993). Calculus simulation method and application of GM(1,1). *Systems Engineering*, vol. 11, no. 6, pp. 31-36.

[116] G. B. John (1979). *Information Systems: Theory and Practice*, New York: Wiley.

[117] J. Kendrick and A. David (1981). *Stochastic Control for Economic Models*, New York: McGraw-Hill.

[118] G. J. Klir (1991). *Facets of Systems Science*, New York: Plenum.

[119] J-Z. Ko (1996). Data sequence transformation and GM(1,1) model accuracy. In *New Development in Grey Systems Research*, ed. S-F. Liu and Z-X. Xu, Wuhan: Huazhong University of Science and Technology Press, pp. 233–235.

[120] W-M. Lee, Y-S. Liao, and Sh-H Chang (2001). Neural-based WEDM optimization via grey analysis. *Journal of Grey Systems*, vol. 13, no. 2.

[121] B-J. Li (1998). Agricultural automation and industrialization of countryside. *Ph.D. dissertation of Agricultural University of China*, Beijing.

[122] B-L. Li and J-L. Deng (1984). Grey model for the prevention system of cotton insects. *Exploration of Nature*, vol. 3, no. 3, pp. 44–49.

[123] B-J. Li, S-F. Liu, and Y-D. Zhu (2000). Determination of the reliability of grey intervals. *Systems Engineering Theory and Practice*, vol. 20, no. 4, pp. 104–106.

[124] B-J. Li, Y-D. Zhu, and H-Q. Wen (1997). Probability grey evaluation methods and applications in agricultural automation. *Journal of Agricultural Engineering*, vol. 13, no. 3, pp. 16–21.

[125] G-Y. Li, Q-F. Li, and G-F. Zhou (1992). An improvement of grey GM(1,1) prediction model. *Systems Engineering*, vol. 10, no. 6, pp. 27–31.

[126] J-W. Li, M-Y. Chen, P-Y. Dong, and S-H. Zhang (2000). Simplified SCGM(1,1) model and applications. *Journal of Wuhan Transportation University of Science and Technology*, vol. 24, no. 6, pp. 615–618.

[127] M-L. Li and S-G. Lin (1985). *Introduction to Quantitative Analysis of Microeconomics*. Wuhan: Huazhong University of Science and Technology Press.

[128] S-R. Li (1987). Ecological adaptability of main pieces of trees used in Chinese afforestation. *Journal of Henan Agricultural University*, vol. 21, no. 4, pp. 436–441.

[129] S-R. Li, Y. Zhao, and S-F. Liu (1994). Ecological classification of Henan Province and its stability analysis. *Journal of Henan Agricultural University*, vol. 28, no. 2, pp. 111–118.

[130] S-R. Li and C-L. Zhu (1999). Grey predictions applied in weapon testing. *Journal of Jiamusi University* (Natural Science Edition), vol. 17, no. 1, pp. 12–16.

[131] T-J. Li, W. Xiong, P. Wang, and Y-S. Zhen (1995). Grey prediction methods of tests without structure being damaged. *Journal of Huazhong University of Science and Technology*, vol. 23, no. 5, pp. 75–78.

[132] W-X. Li (1990). Grey algebraic curve models. *Systems Engineering*, vol. 8, no. 1, pp. 32–36.

[133] W-X. Li (1990). A clustering analysis method based on grey incidences and its applications. *Systems Engineering*, vol. 8, no. 3, pp. 37–44.

[134] W-X. Li (1991). Degrees of incidence of continuous processes. *Journal of Huazhong University of Science and Technology*, vol. 19, no. 3, pp. 113–116.

[135] X-L. Li (1999). Expansion of the application range of grey systems GM(1,1) model. *Systems Engineering Theory and Practice*, vol. 19, no. 1, pp. 97–101.

[136] X-L. Li (2000). Dynamic even difference grey prediction optimal models. *Survey Engineering*, vol. 8, no. 1, pp. 34–40.

[137] X-Q. Li (1995). A further study on the quantification of grey degrees of incidence models. *Systems Engineering*, vol. 13, no. 6, pp. 58–61.

[138] X-Q. Li (1998). Grey modelings in economic systems decision making. *Journal of Zhonog-Nan Industrial University*, vol. 29, no. 4, pp. 326–329.

[139] X-Q. Li, S-R. Li, and S-L. Han (1997). A generalization of applications of grey systems GM(n,h) model. *Systems Engineering Theory and Practice*, vol. 17, no. 8, pp. 82–85.

[140] Z-M. Li (1986). Elementary applications of grey systems theory in earthquake predictions. *Annual of Earthquakes*, no. 4.

[141] Z-N. Li, W-Y. Yuan, F. Zhang, and G-Q. Li (1996). Reliability analysis based on structural changes of grey systems theory. *Journal of Wuhan University of Hydraulic and Electric Power*, vol. 29, no. 6, pp. 163–165.

[142] Z-Z. Li (1991). *Expert's Simulation Systems for Business Management*. Wuhan: Huazhong University of Science and Technology Press.

[143] C-R. Lin and J-L. Deng (1996). *On Grey Prediction of Gas Pool, New Development in Grey Systems Research*. Wuhan: Huazhong University of Science and Technology Press, pp. 25–31.

[144] J-L. Lin, K-S. Wang, B-H. Yan, and B.H. Tarng (2000). Optimization of parameters in GM(1,1) model by Taguchi method. *Journal of Grey Systems*, vol. 11, no. 4.

[145] L-Z. Lin and W-J. Wu (2001). Exploration on optimal grey GM(1,1) model. *Systems Engineering Theory and Practice*, vol. 21, no. 8, pp. 92–96.

[146] S-G. Lin and M-L. Li (1993). *Concise Economic Statistics and Quantitative Economics*. Shanghai: Shanghai People' Press.

[147] W. Lin (1994). Optimal grey parametric linear programming for Henan Provincial transportation structure. *Agricultural Systems Science and Comprehensive Research*, vol. 10, special issue, pp. 348–350.

[148] W. Lin and Y-X. Huo (1996). Grey multi-target programming and applications. In *New Developments in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 373–375.

[149] Y. Lin (2001). Differential and integral calculus on discrete time series data. In *Proceedings of the 14th International Conference on Systems Science*, vol. 1: *Plenary and Invited Papers: Systems Theory and Control Theory*, pp. 123–131. Also, *Systems Science*, vol. 27, no. 3, pp. 49–58.

[150] Y. Lin (2001). Information, prediction and structural whole: an introduction. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 30, no. 4, pp. 305–364.

[151] Y. Lin, M-Y. Chen, and S-F. Liu (2004). Theory of grey systems: Capturing uncertainties of grey information. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 33, no. 2, pp. 196–218.

[152] Y. Lin, R. DeNu, and N. Patel (2002). True distance fit of exponential curves and tests of applications. *International Journal of Applied Mathematics*, vol. 9, no. 1, pp. 49–68.

[153] Y. Lin and T-H. Fan (1997). The fundamental structure of general systems and its relation to knowability of the physical world. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 26, no. 3, pp. 275–285.

[154] Y. Lin and S-F. Liu (1999). Regional economic planning based on systemic analysis of small samples (I). *Problems of Nonlinear Analysis in Engineering Systems*, vol. 10, no. 2, pp. 24–39.

[155] Y. Lin and S-F. Liu (1999). Several programming models with unascertained parameters and their application. *Journal of Multi-Criteria Decision Analysis*, vol. 8, pp. 206–220.

[156] Y. Lin and S-F. Liu (2000). Regional economic planning based on systemic analysis of small samples (II). *Problems of Nonlinear Analysis in Engineering Systems*, vol. 11, no. 6, pp. 33–49.

[157] Y. Lin and S-F. Liu (2000). Law of exponentiality and exponential curve fitting. *Systems Analysis Modelling Simulation*, vol. 38, pp. 621–636.

[158] Y. Lin and S-F. Liu (2000). A systemic analysis with data (I) and (II). *International Journal of General Systems*, vol. 29, pp. 989 – 999 and 1001–1013.

[159] Y. Lin, S-F. Liu, and D. McNeil (1998). Past, present and future of systems research. In *Systems Science and Sustainable Development*, Beijing: Literature of Science and Technology Press, pp. 245–252.

[160] Y. Lin and S-T. Wang (1998). Developing a mathematical theory of computability which speaks the language of levels. *Mathematical and Computer Modelings: An International Journal*, vol. 27, no. 6, pp. 23–32.

[161] F. H. Liu and J. Yi (1995). *Macroeconomics*. Wuhan: Huazhong University of Science and Technology Press.

[162] K. Liu (1982). Grey sets and stability of grey systems. *Journal of Huazhong University of Science and Technology*, vol. 10, no. 3, pp. 23–25.

[163] K-D. Liu, Y. Lin, and L-G. Gao (2001). Informational uncertainties and their mathematical expressions. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 30, no. 4, pp. 378–396.

[164] K-D. Liu, Y. Lin, and L-G. Gao (2002). Unascertained rationals and subjective uncertain information. *Systems Analysis Modelling Simulation*, vol. 42, pp. 343–358.

[165] K-L. Liu, Y-X. Lin, and G. Dal (2001). Grey predicting the structural failure load. *Journal of Grey Systems*, vol. 13, no. 2.

[166] L. S. Liu (1995). Progress in technology and its impact on productivity. In *Management Science and Systems Science*, Xian: Xian Jiaotong University Press, pp. 85–91.

[167] P-L Liu (2001). Stability of continuous and discrete time-delay grey systems. *International Journal of Systems Science*, vol. 32, no. 7.

[168] S-A. Liu, H-T. Du, and X-L. Wang (2001). Theoretical and applicational development of rough sets. *Systems Engineering Theory and Practice*, vol. 21, no. 10, pp. 77–82.

[169] S-F. Liu (1987). P-F theorems of grey non-negative matrices. *Journal of Henan Agricultural University*, vol. 21, no. 4, pp. 376–380.

[170] S-F. Liu (1989). On Perron-Frobenius theorem of grey nonnegative matrix, *The Journal of Grey Systems*, vol. 1, no. 2, pp. 157–166.

[171] S-F. Liu (1991) The three axioms of buffer operator and their application. *The Journal of Grey Systems*, vol. 3, no. 1, pp. 39–48.

[172] S-F. Liu (1991). A study on optimization of business structures in Henan countryside. *Journal of Henan Agricultural University*, vol. 25, no. 1, pp. 38–44.

[173] S-F. Liu (1992). Buffer operators and applications. *Grey Systems Theory and Applications*, vol. 2, no. 1, pp. 45–50.

[174] S-F. Liu (1992) Generalized Degree of Grey Incidence, In *Information and Systems*, ed. by Zhang Shengkai, Dalian: DMU Publishing House, pp. 113–116.

[175] S-F. Liu (1993). Positioned solution of grey parametric linear programming, *Grey Systems Theory and Applications*, vol. 3, no. 2, pp. 23–28.

[176] S-F. Liu (1993). Fixed weight grey cluster analysis and evaluation. In *New Methods of Grey Systems*, Beijing: Agriculture Press, pp. 178–184.

[177] S-F. Liu (1993). Von Neumann's economic growing turnpike of Henan Province, In *Systems Science and Systems Engineering*, ed. by Zheng Weimin, Beijing: International Academic, pp. 316–319.

[178] S-F. Liu (1994). Grey forecast of drought and inundation in Henan Province, *The Journal of Grey Systems*, vol. 6, no. 4, pp. 279–288.

[179] S-F. Liu (1995). On measure of grey information, *The Journal of Grey Systems*, vol. 7, no. 2, pp. 97–101.

[180] S-F. Liu (1995). Sequence operators and systems modeling. In *New Development in Fundamental Science Research*, Beijing: Chinese Science and Technology Press, pp. 16 –20.

[181] S-F. Liu (1996). Position and role of grey systems theory in scientific development. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 230–232.

[182] S-F. Liu (1996). Axiom on grey degree, *Journal of Grey Systems*, vol. 8, no. 4, pp. 397–400.

[183] S-F. Liu (1997). Traps in systems prediction affected by shockwaves and buffor operators. *Journal of Huazhong University of Science and Technology*, vol. 25, no. 1, pp. 28–31.

[184] S-F. Liu (1998). New methods in grey mathematics and analysis of science and technology systems. *Ph.D. dissertation of Huazhong University of Science and Technology.*

[185] S-F Liu (2002). A new method for small sample prediction. In *The 12th International WOSC Congress Proceedings. Advances in Systems Science and Application*, vol. 4, no. 1, pp. 91–102.

[186] S-F. Liu and K-J. Chen (2001). Analysis on Jiangsu provincial research funding and effectiveness of usage. *Journal of Nanjing University of Aeronautics and Astronautics*, vol. 3, no. 3, pp. 24–27.

[187] S-F. Liu and J-L. Deng (1996). Grey recognition on disorders of brain functions and evaluation models for psychological health. *Exploration of Nature*, vol. 15, no. 4, pp. 46–49.

[188] S-F. Liu and J-L. Deng (1999). GM(1,1) Coding for exponential series. *Journal of Grey Systems*, vol. 11, no. 2, pp.147–152.

[189] S-F. Liu and J-L. Deng (1999). The range suitable for GM(1,1), *The Journal of Grey Systems*, vol. 11, no. 1, pp. 131–138.

[190] S-F. Liu and J-L. Deng (2000). The applicability range of GM(1,1) model. *Systems Engineering Theory and Practice*, vol. 20, no. 5, pp. 121–124.

[191] S-F. Liu and Y-G. Dong (1997). The positioned solution of the linear programming with grey parameters, *Advance in Systems Science and Application*, vol. 3, no. 1, pp. 371–377.

[192] S-F. Liu and Y-G. Dong (1997). LPGP flotation and satisfactory degree of positioned solutions. *Journal of Huazhong University of Science and Technology*, vol. 25, no. 1, pp. 24–27.

[193] S-F. Liu, Y-G. Dong, and B-J. Li (1998). Computational analysis for contribution rates due to progress in technology of different time periods in Henan Province. *Journal of Henan Agricultural University*, vol. 32, no. 3, pp. 203–207.

[194] S-F. Liu, Y-G. Dong, and B-J. Li (1999). Computations of contribution rates due to progress in technology of Henan Province and horizontal comparison. *Journal of Agricultural University*, vol. 33, no. 1, pp. 40–43.

[195] S-F. Liu, Y-G. Dong, and B-J. Li (1999). G–C–D model and estimate on contribution rate due to progress in technology. *Management Science of China*, vol. 7, no. 2, pp. 76–80.

[196] S-F. Liu, Y-G. Dong, and Y. Lin (2004). Synthetic utility index method and venturous capital decision-making. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 33, no. 2, pp. 288–294.

[197] S-F. Liu and J. Forrest (1997). The role and position of grey system theory in science development, *Journal of Grey Systems*, vol. 9, no. 4, pp. 351–356.

[198] S-F. Liu, T-B. Guo, and Y-G. Dong (1999). *Grey Systems Theory and Applications* (2nd edition). Beijing: Science Press.

[199] S-F. Liu and B-J. Li (1998). Evaluation criteria for regional dominant businesses and mathematical modeling. *Management Science of China*, vol. 16, no. 2, pp. 8–13.

[200] S-F. Liu, B-J. Li, and Y-G. Dong (2000). The G–C–D model and technical change. *Systems Analysis Modelling Simulation*, vol. 38, pp. 435–447.

[201] S-F. Liu and Y. Lin (1999). *An Introduction to Grey Systems: Foundations, Methodology and Applications*. IIGSS Academic, Grove City, PA.

[202] S-F. Liu, Y. Lin, Y-G. Dong, and B-J. Li (2004). Technical change and funds for science and technology. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 33, no. 2, pp. 295–302.

[203] S-F. Liu and K-Q. Shi (1993). *Selected Papers in Grey Systems*. Kaifen: Henan University Press.

[204] S-F. Liu and S-A. Sun (1993). *Applied Mathematics*. Zhenzhou: Henan Press of Science and Technology.

[205] S-F. Liu and Z-Y. Wang (2000). Entropy of grey evaluation coefficient vector. *Journal of Grey Systems*, vol. 12, no. 3.

[206] S-F. Liu and Z-X. Xu (1996). *New Development in Grey Systems Research*. Wuhan: Huazhong University of Science and Technology Press.

[207] S-F. Liu and L. Yang (1993). Self-organizational mechanism in regional economics. In *Carefully Selected Papers of Chinese Young Scholars*.

[208] S-F. Liu and L. Yang (1994). *Evaluation, Warning and Control of Regional Economics*. Zhenzhou: Henan People's Press.

[209] S-F. Liu, L. Yang, and G-Q. Zhen (1994). Evaluation and warning of regional economics. *Henan Economics*, vol. 10, no. 1, pp. 41–44.

[210] S-F. Liu, L. Yang, and Y-D. Zhu (1994). A quantitative method for determination of dominant businesses of a region. *Agricultural Systems Science and Comprehensive Research*, vol. 10, special issue, pp. 355–358.

[211] S-F. Liu and R-S. Yang (1989). Some discussion on economic vitalization and starting mechanism. *Investigation Research and Exploration*, vol. 4, no. 10, pp. 1–16.

[212] S-F. Liu and L. Zhao (1989). Von Neumann turnpike and Chinese economic growth. *Investigation, Research and Exploration*, vol. 4, no. 11, pp. 1–12.

[213] S-F. Liu, L. Zhao, Y. Jia, and X. L. Guao (1985). Possible reasons for Henan's economic laggard. *Henan Economics*, vol. 1, no. 1, pp. 24–29.

[214] S-F. Liu, L. Zhao, and Z-Y. Wang (2001). A new method to evaluate risky investments. *Management Science of China*, vol. 9, no. 2, pp. 22–26.

[215] S-F. Liu, L. Zhao, and D-Y. Zhao (1989). An analysis for Henan provincial human resources. *Science and Technology Daily*, September 28, 1989.

[216] S-F. Liu, L. Zhao, and Y-D. Zhu (1993). Grey combined models for econometrics. In *Selected Papers in Grey Systems,* Kaifen: Henan University Press, pp. 133–138.

[217] S-F. Liu, L. Zhao, and Y-D. Zhu (1993). G-E prediction model established for Henan Provincial food production. *Journal of Henan Agricultural University*, vol. 27, no. 1, pp. 90–94.

[218] S-F. Liu and Y-D. Zhu (1993). Triangular membership function evaluation model for regional economic evaluations. *Journal of Agricultural Engineering*, vol. 9, no. 2, pp. 8–13.

[219] S-F. Liu and Y-D. Zhu (1994). Study on triangular model and indexes in synthetic evaluation of regional economy, In *Mianyun Chen, Systems Control Information Methodologies & Applications*, HUST Publishing House, pp. 1274–1279.

[220] S-F. Liu and Y-D. Zhu (1996). Grey-econometrics combined model, *The Journal of Grey Systems*, vol. 8, no. 1, pp. 103–110.

[221] S-F. Liu, Y-D. Zhu, and B-J. Li (1996). Main reasons and resolutions for low income earners in the countryside. *Research for Agricultural Modernization*, vol. 17, no. 3, pp. 154–158.

[222] T-C. Liu and E-Z. Liu (1996). *G*rey analysis for development in agriculture. In *New Development in Grey Systems Research*, Wuhan: PHuazhong University of Science and Technology Press, pp. 213–215.

[223] W-G. Liu and L-H. Jiong (1996). Grey GM(1,1) model in production of alloy iron. *Steel*, vol. 31, no. 9, pp. 52–56.

[224] X-Q. Liu and Z-M. Wang (1996). *Grey Economic Prediction Model and Applications*. Ji-Nan: Yellow River Press.

[225] X-Q. Liu, S-Z. Wang, and Z-M. Song (1993). *Introduction to Grey Incidence Spaces*. Guiyang: Guiyang People's Press.

[226] Y-B. Liu and S-J. Li (1996). About the definition and operations of grey numbers. *Journal of Hei-Long-Jiang University* (Natural Science Edition), vol. 13, no. 4, pp. 26–28.

[227] Z-Y. Liu (1995). Grey generation functions with n input factors. *Systems Engineering Theory and Practice*, vol. 15, no. 9, pp. 6–8.

[228] Z-Y. Liu (2000). Two problems about grey systems analysis. *Systems Engineering Theory and Practice*, vol. 20, no. 9, pp. 123–124.

[229] A-L. Lu (1996). A grey design for studying physiological digest system. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press.

[230] A-L. Lu and Y. Shao (1996). A diagnostic model for difficult and complicated illnesses. In *New Development in Grey Systems Research*. Wuhan: Huazhong University of Science and Technology Press.

[231] F. Lu (1997). A study on distinguishing coefficients of grey systems degrees of incidence. *Systems Engineering Theory and Practice*, vol. 17, no. 6, pp. 49–54.

[232] H-H. Lu, L. Zhong, and H-X. Xia (2000). Exploration on combined technology of grey systems and neural networks. *PC Development*, vol. 10, no. 3, pp. 3–5.

[233] Z-L. Lu and Z-P. Fan (1990). Grey systems modeling method for tendencies with level jumps. *Systems Engineering*, vol. 8, no. 5, pp. 55–58.

[234] B-X. Luo (1986). *Input-Output Analysis and Its Mathematical Foundation*. Wuhan: Huazhong University of Science and Technology Press.

[235] M. F. Luo (1994). Fault detection, diagnosis and prognosis using grey system theory, *Monash University, A dissertation submitted for the degree of Ph.D in engineering*.

[236] Q-Z. Luo (1989). Design and applications of grey input-output optimal models. *Systems Engineering Theory and Practice*, vol. 9, no. 5, pp. 55–59.

[237] Q-Z. Luo (1990). *Management of Agricultural and Forestry Systems*. Hongzhou: Zhejiang University Press.

[238] Q-Z. Luo and G-X. Xu (1989). *Grey Incidence Analysis and Its Applications*. Nanjing: Jiangsu Science and Technology Press.

[239] X-M. Luo and H-H. Yang (1994). Grey comprehensive evaluation models. *Systems Engineering and Electronic Technology*, vol. 16, no. 9, pp. 18–25.

[240] X-R. Ma and P. Zhen (1985). Predictions on capability of change of low-alloy steel casts. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 59–62.

[241] J. G. Miller (1978). *Living Systems*, New York: Mcgraw-Hill.

[242] E. D. Mills (1972). *Studies in the Structure of the Urban Economy*, Baltimore: Johns Hopkins University Press.

[243] T-Z. Ming and L. Ming (1996). Grey systems prediction and decision making for problems in environmental geology. In *New Development in Grey Systems Research*, ed. by S-F. Liu and Z-X. Xu, Wuhan: Huazhong University of Science and Technology Press, pp. 301–303.

[244] W-R. Mong (1985). Grey prediction for the technological improvement of publicly own businesses. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 63–66.

[245] H-S. Mu and J-G. Yang (1992). Grey systems theory and economic wavelet analysis. *Systems Engineering*, vol. 10, no. 2, pp. 25–28.

[246] Y. Mu and J-G. Liu (1996). Studies on grey systems modeling method. *Journal of Shan-Dong College of Construction Materials*, vol. 10, no. 3, pp. 67–71.

[247] D. Nachmias (1990). *The Practice of Policy Evaluation*, New York: St. Martin's.

[248] S. S. Nagel (1982). *Policy Evaluation: Making Optimum Decision*, New York: Praeger.

[249] A. C. Nelson (1995). *Growth Management Principles and Practices*, Washington DC: Planners.

[250] H. Nikaido (1968). *Convex Structures and Economic Theory*, New York: Academic.

[251] M. Norris (2001). Shades of grey superceded by precise imaging [AOI in PCB assembly]. *Electronics Manufacture & Test*, no. 12.

[252] OECD (1981). *The Measurement of Scientific Activities*, Paris: Frascati Manual.

[253] S. C. OuYang, Y. Lin and Y. Wu (2000). Prediction of natural disasters. *International Journal of General Systems,* vol. 29, pp. 897–912.

[254] H-P. Pan and Z-H. Huang (1996). Recognizing lithofaces via grey clustering, *The Journal of Grey Systems*, vol. 8, no. 1, pp. 93–102.

[255] G. Pask (1976). *Conversation Theory: Applications in Education and Epistemology*, Amsterdam: Elsevier.

[256] M. Q. Patton (1990). *Qualitative Evaluation and Research Methods*, Newberry Park, CA: Sage.

[257] Z. Pawlak (1991). *Rough Sets: Theoretical Aspects of Reasoning About Data*. Dordrecht: Kluwer Academic.

[258] H-Y. Pen and Y. Gao (1997). Improvement of grey reducible models: linear regression method after z-transformation. *Monitor of Drought Environment*, vol. 11, no. 1, pp. 6–7.

[259] X-L. Peng and X. Lu (1991). Algebraic criteria for stability and instability of grey discrete systems. *Applied Mathematics*, vol. 4, no. 2, pp. 123–128.

[260] X-S. Qian (1988). *About Systems Engineering*. Chang Sa: Hunan Science and Technology Press.

[261] X-J. Qiu (1995). Incidence analysis for grey clustering and its applications. *Systems Engineering Theory and Practice*, vol. 15, no. 1, pp. 15–21.

[262] P. Reitveld (1980). *Multiple Objective Decision Methods and Regional Planning*, Amsterdam: North-Holland.

[263] H. W. Richardson (1969). *Regional Economics*, New York: Praeger.

[264] A. D. Robert D. R. Kenneth (1981). *Management Science*, New York: Macmillan.

[265] S. P. Robert and L. R. Danial (1981). *Econometric Models and Economic Forecasts*, New York: McGraw-Hill.

[266] D-X. Shen (1993). *Stochastic processes*. Wuhan: Huazhong University of Science and Technology Press.

[267] K-Q. Shi and S-F. Liu (1994). Grey relation G(2) on (X, Γ), *The Journal of Grey Systems*, vol. 6, no. 2, pp. 109–118.

[268] K-Q. Shi and Z-L. Wang (1992). A( ) matrix and its grey characteristics, *Journal of Grey Systems Theory and Practice*, vol. 4, no. 1, pp. 11–17.

[269] N-X. Shui, T-H. Tong, and Z. Sa (1990). About the GM(2,1) model of grey systems. *Systems Engineering*, vol. 8, no. 6, pp. 31–34.

[270] N-X. Shui, T-H. Tong, and Z. Sa (1992). Some theoretical problems on grey degrees of incidence. *Systems Engineering*, vol. 10, no. 6, pp. 23–26.

[271] Q-J. Shi, B-T. Qin, and J. Chen (1985). *Technological Progress and Economic Growth*. Beijing: Science and Technology Literature Press.

[272] J. Sjoberg (200). A nonlinear grey-box example using a stepwise system identification approach. In *12th Symposium on System Identification*, Santa Barbara, CA.

[273] W. Shu (1998). Grey prediction models and applications. *Journal of Mathematics*, S1.

[274] N-X. Shui and S-C. Qing (1998). Some theoretical problems of grey systems GM(1,1) model. *Systems Engineering Theory and Practice*, vol. 18, no. 4, pp. 59–63.

[275] R. Slater (1996). *Soros: The Life, Times and Trading Secrets of the World's Greatest Investor*. Burr Ridge, IL: Irwin.

[276] N. L. Smith (1992). *Varieties of Investigative Evaluation*, San Francisco: Jossey-Bass.

[277] C-Q. Song and C-Q. Tan (1997). Grey systems analysis criteria and method developed for adequate evaluation of deposit layers of petroleum gas and applications. *Systems Engineering Theory and Practice*, vol. 17, no. 3, pp. 74–82.

[278] Z-M. Song, X-Q. Liu, and S-C. Wang (1997). Curvity simulation method of grey exponential curves. *Systems Engineering Theory and Practice*, vol. 17, no. 6, pp. 55–57.

[279] Z-M. Song, X-J. Tong, and X-P. Xiao (2001). Grey GM(1,1) model of center approximation style. *Systems Engineering Theory and Practice*, vol. 21, no. 5, pp. 110–113.

[280] Z-M. Song and Z-M. Wang (1998). Grey Engel model (GEM). *Journal of Yian-Tai Normal College* (Natural Science Edition), vol. 14, no 1, pp. 37–39.

[281] G. Soros (1998). *The Crisis of Global Capitalism: Open Society Endangered*. New York: Public Affairs.

[282] L. Su (1995). A simply computational method for using grey GM prediction. *Applied Technology of Energy Resources*, no. 3, pp. 5–10.

[283] H-Y. Sun and Q-J. Cuo (2001). Modeling and prediction of chaotic time series. *Systems Engineering Theory and Practice*, vol. 21, no. 5, pp. 106–109.

[284] J-J. Sun (1996). Study on development models which balance science and technology, economy and society. *Management Science of China*, no. 2.

[285] A. Sydow, Y. Lin, S-F. Liu, R. DeNu and W. Mennell (2001). Exponential prediction models based on sequence operators. *Kybernetes: The International Journal of Systems and Cybernetics*, vol. 30, no. 4, pp. 397–410.

[286] Y. Tamura, D-P. Zhang, N. Umeda, and K. Sakeshita (1992). Load forecasting using grey dynamic model, *The Journal of Grey Systems*, vol. 4, no. 4, pp. 49–58.

[287] X-R. Tang (1996). Grey incidence spaces and applications in medicine. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 130–132.

[288] X-R. Tang (1997). A study on grey medical incidence theory and its applications. *Ph.D. dissertation of Huazhong University of Science and Technology*, Wuhan.

[289] J-G. Tian and Y-X. Huong (1996). Design of a class of grey dynamic input-output optimal models and their applications in the planning of agricultural economic development. *Economic Mathematics*, vol. 13, no. 1, pp. 64–70.

[290] J. Tinbergen (1964). *Econometrics*, Englewood Cliffs, NJ: Prentice-Hall.

[291] J. Tinbergen (1967). *Development Planning*, London: Weidenfeld and Nicolson.

[292] J-L. Tong (1999). Stability of zero solution of nonlinear grey discrete systems. *Journal of Yang-Zhou University* (natural science edition), vol. 2, no. 2, pp. 9–12.

[293] W-X. Tong (1995). Degree of T-type incidence and its calculation. *Mathematical Statistics and Management*, vol. 14, no. 1, pp. 34–37.

[294] W-X. Tong (1995). A new method to estimate the GM(1,1) model parameters and hypothesis testing. *Systems Engineering Theory and Practice*, vol. 15, no. 3, pp. 20–49.

[295] X-W. Tong, Y. Zhen and Z-X. Cao (1993). Studies on varying weight combined prediction models. *Prediction*, vol. 12, no. 3, pp. 33–39.

[296] J. M. Utterback (1986). Innovation and corporate strategy, *International Journal of Technology Management*, vol. 1, no. 1 - 2, pp. 119–132.

[297] B-L. Wang (1993). Computation of degrees of incidence for a class of problems. In *Selected Papers on Grey Systems*, Kaifeng: Henan University Press, pp. 101–102.

[298] G-X. Wang, H-R. Bai, and H-X. Wu (1996). Grey incidence analysis applied to evaluate the factors affecting an elastic manufacturing system. In *New Development in Grey Systems Research*, ed. by S-F. Liu and Z-X. Xu, Wuhan: Huazhong University of Science and Technology Press, pp. 11–14.

[299] G-X. Wang, Y-F. Li, and H-X. Wu (1996). Weighted factor grey incidence evaluation of constraints on mechanic elements. In *New Development in Grey Systems Research*, ed. by S-F. Liu and Z-X. Xu, Wuhan: Huazhong University of Science and Technology Press, pp. 15–19.

[300] G-Y. Wang (1990). Unascertained information and its mathematical treatment. *Journal of College of Harbin Construction Engineering*, vol. 23, no. 4, pp. 11–19.

[301] H-X. Wang and J-W. Feng (1995). Comparison between grey prediction models and exponential curve models. *Journal of Nanjing University of Industry*, vol. 19, no. 3, pp. 235–238.

[302] M-X. Wang and W. Liu (1985). Grey extra long-term disaster predictions based on early frosts. *Fuzzy Mathematics*, vol. 5, no. 2, pp. 80–82.

[303] Q-Y. Wang and Y-M. Cui (2000). Some fundamental problems of uncertain systems prediction. *Journal of Systems Engineering Theory and Applications*. vol. 20, no. 3, pp. 93–98.

[304] Q-Y. Wang and K-D. Liu (1990). *Methods and Applications of Grey Mathematics*. Chengdu: Xi-Nan Jiao Tong University Press.

[305] Q-Y. Wang and R-H. Zhao (1992). Pan-grey matrices. *Journal of Hebei College of Mining Engineering*, vol. 9, no. 3, pp. 27–33.

[306] W-P. Wang (1993). Analysis of optimal values in GLP, *The Journal of Grey Systems*, vol. 5, no. 4, pp. 315–317.

[307] W-P. Wang (1994). A study on the theory and methods of grey information treatment. *Ph.D. dissertation of Huazhong University of Science and Technology*, Wuhan.

[308] W-P. Wang (1997). Study on grey linear programming, *The Journal of Grey Systems*, vol. 9, no. 1, pp. 41–46.

[309] W-P. Wang and J-L. Deng (1992). Measure of intension of information contents contained in propositions of grey systems. *Grey Systems Theory and Practice*, vol. 2, no. 1, pp. 41–43.

[310] X-M. Wang (1992). *Grey Systems Models Applied in Studies of Country Economy*. Wuhan: Huazhong University of Science and Technology Press.

[311] X-M. Wang (1993). Grey dynamic model established for economic growth and analysis for economic periods. *Systems Engineering Theory and Practice*, vol. 13, no. 1, pp. 42–47.

[312] X-M. Wang and J-J. Luo (1986). *Collection of Computer Programs for Grey Systems Prediction and Decision Making Models*. Beijing: Popularization of Science Press.

[313] X-Y. Wang and X. Yang (1997). Control of quality prediction and diagnostic grey systems method. *Systems Engineering Theory and Practice*, vol. 17, no. 5, pp. 105–108.

[314] Y-N. Wang, K-D. Liu, and Y-C. Li (2001). Optimal GM(1,1) modeling of the whitenization of grey derivatives. *Systems Engineering Theory and Practice*, vol. 21, no. 5, pp. 124–128.

[315] Y-N. Wang, G-Z. Liu, and K-D. Liu (2000). One step by step direct modeling based on GM(1,1). *Systems Engineering Theory and Practice*, vol. 20, no. 9, pp. 99–104.

[316] Y-Z. Wang and W-A. Zhou (1997). An improved grey statistics method for evaluating the safety of stoping work surfaces. *Systems Engineering Theory and Practice*, vol. 17, no. 9, pp. 118–123.

[317] Z-L. Wang (1993). Grey systems theory applied in selection analysis. In *Selected Papers on Grey Systems*, Kaifeng: Henan University Press, pp. 305–307.

[318] Z-L. Wang (1998). Theory on grey modeling techniques. *Ph.D. dissertation of Huazhong University of Science and Technology*, Wuhan.

[319] Z-L. Wang and X.Z. Li (1996). Ordering of grey numbers and inequalities. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 364–366.

[320] Z-L. Wang, Y-L. Liu, and K-Q. Shi (1995). On kenning grey degree, *The Journal of Grey Systems*, vol. 7, no. 2, pp. 103–110.

[321] G. M. Weinberg (1975). *An Introduction to General Systems Thinking*, New York: John Wiley.

[322] J-C Wen, J-L Chen, and J-S Yang (2001). Study of the non-equil gap GM(1,1) modeling. *Journal of Grey Systems*, vol. 13, no. 3.

[323] K-L. Wen, T-C Chang, W-C Chang, and M-L. You (2000). The study of missing point in GM(1,1) model. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*.

[324] H-Q. Wu and Z-A. Qiu (1990). *Introduction to Grey Mathematics*. Shijiazhuong: Hebei People's Press.

[325] J-H. Wu and K-L. Wen (2001). Rolling error in GM(1,1) modeling. *Journal of Grey Systems*, vol. 13, no. 1.

[326] R. Wu (1999). About applicability range of grey systems GM(1,1) model. *Journal of Beijing Industrial University*, vol. 25, no. 1, pp. 72–75.

[327] S-Z. Wu (1984). Criteria developed for forever observability of grey systems. *Journal of Huazhong University of Science and Technology*, vol. 12, no. 3, pp. 15–18.

[328] Y. Wu, S-Z. Yang, and H-Y. Zhao (1988). Grey prediction and discussion on time series predictions. *Journal of Huazhong University of Science and Technology*, vol. 16, no. 3, pp. 27–33.

[329] J. Xia and H-Y. Zhao (1996). Grey artificial neuro network model and its applications in short-term predictions of runoffs. *Systems Engineering Theory and Practice*, vol. 16, no. 11, pp. 82–90.

[330] K-W. Xia, J. Zhu, and Z-Q. Song (1995). Grey filtering of Mining signals' GM(1,1) model. *Mining Technology*, vol. 19, no. 3, pp. 172–177.

[331] X-P. Xiao (1997). Some study and comments on the theory of grey incidence models. *Systems Engineering Theory and Practice*, vol. 17, no. 8, pp. 76–81.

[332] X-P. Xiao (2000). On parameters in grey models. *Journal of Grey Systems*, vol. 11, no. 4.

[333] X-P. Xiao and S-Y. Li (1996). Some recent progress in applications of grey incidence analysis. *Exploration of Nature*, vol. 15, no. 4, pp. 50–56.

[334] X-P. Xiao, L-Z. Xie, and D-H. Huong (1995). An improved computation scheme for grey degrees of incidence and its applications. *Mathematical Statistics and Management*, vol. 14, no. 5, pp. 27–30.

[335] X-W. Xie and D. Wang (1997). An initial study on the theory and methods of grey input-output. *Systems Engineering Theory and Practice*, vol. 17, no. 1, pp. 56–61.

[336] M. Xing (2001). Combined grey neural network models for summer seasonal prediction. *Systems Engineering Theory and Practice*, no. 1, pp. 83–90.

[337] H-J. Xiong, M-Y. Chen, and T. Hao (1999). Two classes of grey models for control systems. *Journal of Wuhan Transportation University of Science and Technology*, vol. 30, no. 5, pp. 465–468.

[338] H-J. Xiong, Y-F. Xiong, M-Y. Chen, and T. Hao (1999). Generalizations of grey systems SCGM(1,h)_c model. *Journal of Huazhong University of Science and Technology*, vol. 27, no. 1, pp. 1–3.

[339] F-S. Xu (1996). Grey models for predicting foods. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 301–303.

[340] G-H. Xu (1993). Improved generalized degrees of incidence and criteria for multi-target grey incidence decision making. In *Selected Papers in Grey Systems*, Kaifeng: Henan University Press, pp. 96–100.

[341] J-P. Xu (1993). Grey prediction linear programming. *Management Science of China*, vol. 1, no. 4, pp. 35–42.

[342] J-P. Xu (1993). Optimizing theory on working procedure with grey numbers, *The Journal of Grey Systems*, vol. 5, no. 3, pp. 189–196.

[343] J-P. Xu (1995). On a kind of information grey number, *The Journal of Grey Systems*, vol. 7, no. 2, pp. 111–130.

[344] J-P. Xu (1997). Grey drift linear programming, *The Journal of Grey Systems*, vol. 9, no.3, pp. 243–246.

[345] W-X. Xu and Q-T. Zhang (2001). A synthetic algorithm based on grey theory and fuzzy mathematics. *Systems Engineering Theory and Practice*, no. 4, pp. 34–37.

[346] Y-C. Xu and X-R. Tang (1996). Analysis of medical grey fuzziness. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 161–162.

[347] Z-X. Xu and G-P. Wu (1993). *Grey Systems Theory and Grey Prediction of Mineral Deposits*. Wuhan: China Geological University Press.

[348] J-H. Yang and Z-L. Wang (1996). Degrees of grey information. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 360–361.

[349] J-H. Yang and J-L. Yang (1996). Grey information and its applications. In *New Development in Grey Systems Research*, Wuhan: Huazhong University of Science and Technology Press, pp. 362–363.

[350] X-H. Yang, G-T. Zhang, and J-L. Jing (1998). Parametric estimation for grey systems GM(1,1) model. *Monitor of Drought Environment*, vol. 12, no. 2, pp. 76–80.

[351] Y-L. Yeh, T-C. Chen, and C-N. Lin (2001). Effects of data characteristics for GM(1,1) modeling. *Journal of Grey Systems*, vol. 13, no. 2.

[352] D-S. Yi (1987). Grey models and human talent prediction. *Systems Engineering*, vol. 5, no. 1, pp. 36–43.

[353] L-Y. Yian and L-X. Mao (1998). Applications of GM(1,1) model. *Systems Engineering Theory and Practice*, vol. 18, no. 10, pp. 104–106.

[354] K-C. Ying, C-J. Liao, and Y-T. Hsu (2000). Generalized admissible region of class ratio for GM(1,1). *Journal of Grey Systems*, vol. 12, no. 2.

[355] D-J. Yu (1991). Some discussion on grey systems modeling method. *Systems Engineering*, vol. 9, no. 5, pp. 9–12.

[356] E-X. Yu (1988). *Matrix Theory Applied in Science and Technology*. Wuhan: Huazhong University of Science and Technology Press.

[357] J-Z. Yuan (1991). *Grey Systems Theory and Applications*. Beijing: Scientific Press.

[358] J- Z. Yuan (1991). *Theory and Applications of Grey Systems Theory*. Beijing: Science Press.

[359] D-H. Zhang, K-Q. Shi, and S-F. Jiang (2001). Grey predicting power load via GM(1,1) modified. *Journal of Grey Systems*, vol. 13, no. 1.

[360] J-X. Zhang, X-J. Guan, and Z-X. He (1989). *Manual of Soft Sciences*. Tianjing: Translation of Tianjing Science and Technology.

[361] P-L. Zhang and Y-W. Chen (1995). Applications of grey prediction in risk control of security investment. *Journal of Hubei College of Industry*, vol. 10, no. 4, pp. 93–96.

[362] Q-S. Zhang (1996). Difference information theory of grey fuzzy sets. *Ph.D. dissertation of Huazhong University of Science and Technology*.

[363] Q-S. Zhang (2001). Difference information entropy in grey theory. *Journal of Grey Systems*, vol. 13, no. 2.

[364] Q-S. Zhang, J-L. Deng, and G. Fu (1994). On grey clustering in grey hazy set, *The Journal of Grey Systems*, vol. 7, no. 4, pp. 377–390.

[365] Q-S. Zhang, X-H. Guo, and J-L. Deng (1996). An entropy analysis method based on grey incidences. *Systems Engineering Theory and Practice*, vol. 16, no. 8, pp. 7–11.

[366] Q-S. Zhang, W-Y. Han, and J-L. Deng (1994). Information entropy of discrete grey number, *The Journal of Grey Systems*, vol. 6, no. 4, pp. 303–314.

[367] S-L. Zhang and G-L. Zhang (1996). Comparison between grey incidences and existing problems. *Systems Engineering*, vol. 14, no. 3, pp. 45–49.

[368] X-W. Zhang, X-M. Wang, and H-S. Nie (1989). *Grey Systems Analysis for Country Economy*. Beijing: Beijing Scholastic Journals Press.

[369] D-Y. Zhao (1990). A complete grey GM(1,1) prediction on main lab materials of Henan Agricultural University. *Research on Higher Education*, vol. 9, no. 2, pp. 20–24.

[370] L. Zhao, S-F. Liu, and W. Lin (1992). An elementary study on the reform of Chinese price system. *Journal of Henan Agricultural University*, vol. 26, no. 1, pp. 84–88.

[371] R-J. Zhao and J-L. Ma (1994). Applications of grey systems remnant models in predictions. *Journal of Guang-Dong College of Pharmacy*, vol. 10, no. 1, pp. 13–16.

[372] Z-N. Zhen, Y-Y. Wu, and H-L. Bao (2001). Pathological problems of GM models. *Systems Engineering Theory and Practice*, no. 5, pp. 38–44.

[373] Y. Zheng (1998). Control on dynamic Smith estimate based on grey predictions. *Industrial Meters and Automatic Equipment*, no. 4, pp. 14–16.

[374] S-J. Zhong (1988). Multi-target grey situation decision making developed for port selections. *Systems Engineering Theory and Practice*, vol. 8, no. 4, pp. 59–66.

[375] C-S. Zhou (1986). Allocation of extrema regions of grey systems. *Journal of Huazhong University of Science and Technology*, vol. 14, no. 5, pp. 659–662.

[376] C-S. Zhou and J-L. Deng (1986). The stability of grey linear system, *Int. J.Control*, vol. 43, no. 1, pp. 313–320.

[377] C-S. Zhou and J-L. Deng (1989). Stability analysis of grey discrete-time systems, *IEEE Trans. AC*, vol. 34, no. 2, pp. 173–175.

[378] C-S. Zhou and Q-L. Que (1989). Some sufficient conditions for the stability of interconnected grey dynamic systems. *Journal of Huazhong University of Science and Technology*, vol. 17, no. 3, pp. 83–88.

[379] C-Q Zhu (2001). Grey prediction modeling via grey relational weighting. *Journal of Grey Systems*, vol.13, no.3.

[380] H-G. Zhu and G-Z. Lian (1992). Grey incidence analysis for systems with and without time series. *Information and Control*, vol. 21, no. 3, pp. 177–179.

[381] Y-D. Zhu, M-W. Jing, and Y-Z. Zhang (1993). Principle of unbalanced self-stimulated adjustment and its application. *Journal of Agricultural Engineering*, vol. 9, no. 4, pp. 16–21.

[382] Y-D. Zhu, S-F. Liu, and Y-Z. Zhang (1997). *Economic Management of the Middle Pointview*, Wuhan: Huazhong University of Science and Technology Press.

[383] Y-D. Zhu, J-S. Ma, and Y-Z. Zhang (1993). A theoretical model established for the mechanism of departmental economic development and practical tests. *Journal of Agricultural Engineering*, vol. 9, no. 3, pp. 1–6.

[384] Y-D. Zhu, Y-Z. Zhang, and S-F. Liu (1993). Corporation in regional economics. *Journal of Agricultural Engineering*, vol. 9, no. 2, pp. 1–7.

[385] Y-D. Zhu and G-D. Zhu (1987). *Foundations of Agricultural Systems Engineering*. Zhenzhou: Henan Press of Science and Technology.

Index