Iterative Methods for Systems of Equations 30.5

Introduction

There are occasions when direct methods (like Gaussian elimination or the use of an LU decomposition) are not the best way to solve a system of equations. An alternative approach is to use an iterative method. In this Section we will discuss some of the issues involved with iterative methods.

Prerequisites

Before starting this Section you should ...

- revise matrices, especially the material in HELM 8
- revise determinants
- revise matrix norms

Learning Outcomes

On completion you should be able to ...

- approximate the solutions of simple systems of equations by iterative methods
- assess convergence properties of iterative methods

1. Iterative methods

Suppose we have the system of equations

$$AX = B.$$

The aim here is to find a sequence of approximations which gradually approach X. We will denote these approximations

$$X^{(0)}, X^{(1)}, X^{(2)}, \ldots, X^{(k)}, \ldots$$

where $X^{(0)}$ is our initial "guess", and the hope is that after a short while these successive **iterates** will be so close to each other that the process can be deemed to have **converged** to the required solution X.

Key Point 10

An **iterative** method is one in which a sequence of approximations (or **iterates**) is produced. The method is successful if these iterates converge to the true solution of the given problem.

It is convenient to split the matrix A into three parts. We write

$$A = L + D + U$$

where L consists of the elements of A strictly below the diagonal and zeros elsewhere; D is a diagonal matrix consisting of the diagonal entries of A; and U consists of the elements of A strictly above the diagonal. **Note that L and U here are not the same matrices as appeared in the LU decomposition! The current L and U are much easier to find.**

For example

$$\underbrace{\begin{bmatrix} 3 & -4 \\ 2 & 1 \end{bmatrix}}_{\substack{\uparrow \\ A}} \; = \; \underbrace{\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}}_{\substack{\uparrow \\ L}} \; + \; \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}_{\substack{\uparrow \\ D}} \; + \; \underbrace{\begin{bmatrix} 0 & -4 \\ 0 & 0 \end{bmatrix}}_{\substack{\uparrow \\ U}}$$

and

$$\underbrace{\begin{bmatrix} 2 & -6 & 1 \\ 3 & -2 & 0 \\ 4 & -1 & 7 \end{bmatrix}}_{\substack{\uparrow \\ A}} \; = \; \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \\ 4 & -1 & 0 \end{bmatrix}}_{\substack{\uparrow \\ L}} \; + \; \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 7 \end{bmatrix}}_{\substack{\uparrow \\ D}} \; + \; \underbrace{\begin{bmatrix} 0 & -6 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\substack{\uparrow \\ U}}$$

and, more generally for 3×3 matrices

$$\underbrace{\begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix}}_{\substack{\uparrow \\ A}} \underset{=}{=} \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ \bullet & 0 & 0 \\ \bullet & \bullet & 0 \end{bmatrix}}_{\substack{\uparrow \\ L}} \underset{+}{+} \underbrace{\begin{bmatrix} \bullet & 0 & 0 \\ 0 & \bullet & 0 \\ 0 & 0 & \bullet \end{bmatrix}}_{\substack{\uparrow \\ D}} \underset{+}{+} \underbrace{\begin{bmatrix} 0 & \bullet & \bullet \\ 0 & 0 & \bullet \\ 0 & 0 & 0 \end{bmatrix}}_{\substack{\uparrow \\ U.}}.$$

The Jacobi iteration

The simplest iterative method is called **Jacobi iteration** and the basic idea is to use the $A = L + D + U$ partitioning of A to write $AX = B$ in the form

$$DX = -(L+U)X + B.$$

We use this equation as the motivation to define the iterative process

$$DX^{(k+1)} = -(L+U)X^{(k)} + B$$

which gives $X^{(k+1)}$ as long as D has no zeros down its diagonal, that is as long as D is invertible. This is Jacobi iteration.

Key Point 11

The **Jacobi iteration** for approximating the solution of $AX = B$ where $A = L + D + U$ is given by

$$X^{(k+1)} = -D^{-1}(L+U)X^{(k)} + D^{-1}B$$

Example 18

Use the Jacobi iteration to approximate the solution $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ of

$$\begin{bmatrix} 8 & 2 & 4 \\ 3 & 5 & 1 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix}.$$

Use the initial guess $X^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Solution

In this case $D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ and $L + U = \begin{bmatrix} 0 & 2 & 4 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$.

First iteration.

The first iteration is $DX^{(1)} = -(L+U)X^{(0)} + B$, or in full

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & -2 & -4 \\ -3 & 0 & -1 \\ -2 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} + \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix} = \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix},$$

since the initial guess was $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$.

Taking this information row by row we see that

$$\begin{aligned} 8x_1^{(1)} &= -16 \quad \therefore \boxed{x_1^{(1)} = -2} \\ 5x_2^{(1)} &= 4 \quad \therefore \boxed{x_2^{(1)} = 0.8} \\ 4x_3^{(1)} &= -12 \quad \therefore \boxed{x_3^{(1)} = -3} \end{aligned}$$

Thus the first Jacobi iteration gives us $X^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} -2 \\ 0.8 \\ -3 \end{bmatrix}$ as an approximation to X.

Second iteration.

The second iteration is $DX^{(2)} = -(L+U)X^{(1)} + B$, or in full

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & -2 & -4 \\ -3 & 0 & -1 \\ -2 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} + \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix}.$$

Taking this information row by row we see that

$$\begin{aligned} 8x_1^{(2)} &= -2x_2^{(1)} - 4x_3^{(1)} - 16 = -2(0.8) - 4(-3) - 16 = -5.6 \quad \therefore \boxed{x_1^{(2)} = -0.7} \\ 5x_2^{(2)} &= -3x_1^{(1)} - x_3^{(1)} + 4 = -3(-2) - (-3) + 4 = 13 \quad \therefore \boxed{x_2^{(2)} = 2.6} \\ 4x_3^{(2)} &= -2x_1^{(1)} - x_2^{(1)} - 12 = -2(-2) - 0.8 - 12 = -8.8 \quad \therefore \boxed{x_3^{(2)} = -2.2} \end{aligned}$$

Therefore the second iterate approximating X is $X^{(2)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} -0.7 \\ 2.6 \\ -2.2 \end{bmatrix}$.

Solution (contd.)

Third iteration.

The third iteration is $DX^{(3)} = -(L+U)X^{(2)} + B$, or in full

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{bmatrix} = \begin{bmatrix} 0 & -2 & -4 \\ -3 & 0 & -1 \\ -2 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} + \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix}$$

Taking this information row by row we see that

$$\begin{aligned} 8x_1^{(3)} &= -2x_2^{(2)} - 4x_3^{(2)} - 16 = -2(2.6) - 4(-2.2) - 16 = -12.4 & \therefore \boxed{x_1^{(3)} = -1.55} \\ 5x_2^{(3)} &= -3x_1^{(2)} - x_3^{(2)} + 4 = -3(-0.7) - (2.2) + 4 = 8.3 & \therefore \boxed{x_2^{(3)} = 1.66} \\ 4x_3^{(3)} &= -2x_1^{(2)} - x_2^{(2)} - 12 = -2(-0.7) - 2.6 - 12 = -13.2 & \therefore \boxed{x_3^{(3)} = -3.3} \end{aligned}$$

Therefore the third iterate approximating X is $X^{(3)} = \begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{bmatrix} = \begin{bmatrix} -1.55 \\ 1.66 \\ -3.3 \end{bmatrix}$.

More iterations ...

Three iterations is plenty when doing these calculations by hand! But the repetitive nature of the process is ideally suited to its implementation on a computer. It turns out that the next few iterates are

$$X^{(4)} = \begin{bmatrix} -0.765 \\ 2.39 \\ -2.64 \end{bmatrix}, \quad X^{(5)} = \begin{bmatrix} -1.277 \\ 1.787 \\ -3.215 \end{bmatrix}, \quad X^{(6)} = \begin{bmatrix} -0.839 \\ 2.209 \\ -2.808 \end{bmatrix},$$

to 3 d.p. Carrying on even further $X^{(20)} = \begin{bmatrix} x_1^{(20)} \\ x_2^{(20)} \\ x_3^{(20)} \end{bmatrix} = \begin{bmatrix} -0.9959 \\ 2.0043 \\ -2.9959 \end{bmatrix}$, to 4 d.p. After about 40 iterations successive iterates are equal to 4 d.p. Continuing the iteration even further causes the iterates to agree to more and more decimal places. The method converges to the exact answer

$$X = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}.$$

The following Task involves calculating just two iterations of the Jacobi method.

Task

Carry out two iterations of the Jacobi method to approximate the solution of

$$\begin{bmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

with the initial guess $X^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Your solution

First iteration:

Answer

The first iteration is $DX^{(1)} = -(L+U)X^{(0)} + B$, that is,

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

from which it follows that $X^{(1)} = \begin{bmatrix} 0.75 \\ 1 \\ 1.25 \end{bmatrix}$.

Your solution

Second iteration:

Answer

The second iteration is $DX^{(1)} = -(L+U)X^{(0)} + B$, that is,

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

from which it follows that $X^{(2)} = \begin{bmatrix} 0.8125 \\ 1 \\ 1.1875 \end{bmatrix}$.

Notice that at each iteration the first thing we do is get a new approximation for x_1 and then we continue to use the old approximation to x_1 in subsequent calculations for that iteration! Only at the *next* iteration do we use the new value. Similarly, we continue to use an old approximation to x_2 even after we have worked out a new one. And so on.

Given that the iterative process is supposed to improve our approximations why not use the better values straight away? This observation is the motivation for what follows.

Gauss-Seidel iteration

The approach here is very similar to that used in Jacobi iteration. The only difference is that we use new approximations to the entries of X as soon as they are available. As we will see in the Example below, this means rearranging $(L+D+U)X = B$ slightly differently from what we did for Jacobi. We write

$$(D+L)X = -UX + B$$

and use this as the motivation to define the iteration

$$(D+L)X^{(k+1)} = -UX^{(k)} + B.$$

Key Point 12

The **Gauss-Seidel iteration** for approximating the solution of $AX = B$ is given by

$$X^{(k+1)} = -(D+L)^{-1}UX^{(k)} + (D+L)^{-1}B$$

Example 19 which follows revisits the system of equations we saw earlier in this Section in Example 18.

Example 19

Use the Gauss-Seidel iteration to approximate the solution $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ of

$$\begin{bmatrix} 8 & 2 & 4 \\ 3 & 5 & 1 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix}.$$ Use the initial guess $X^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Solution

In this case $D + L = \begin{bmatrix} 8 & 0 & 0 \\ 3 & 5 & 0 \\ 2 & 1 & 4 \end{bmatrix}$ and $U = \begin{bmatrix} 0 & 2 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

First iteration.

The first iteration is $(D+L)X^{(1)} = -UX^{(0)} + B$, or in full

$$\begin{bmatrix} 8 & 0 & 0 \\ 3 & 5 & 0 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & -2 & -4 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} + \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix} = \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix},$$

since the initial guess was $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$.

Taking this information row by row we see that

$$\begin{aligned} 8x_1^{(1)} &= -16 & \therefore\ & \boxed{x_1^{(1)} = -2} \\ 3x_2^{(1)} + 5x_2^{(1)} &= 4 \therefore 5x_2^{(1)} = -3(-2) + 4 & \therefore\ & \boxed{x_2^{(1)} = 2} \\ 2x_1^{(1)} + x_2^{(1)} + 4x_3^{(1)} &= -12 \therefore 4x_3^{(1)} = -2(-2) - 2 - 12 & \therefore\ & \boxed{x_3^{(1)} = -2.5} \end{aligned}$$

(Notice how the new approximations to x_1 and x_2 were used immediately after they were found.)

Thus the first Gauss-Seidel iteration gives us $X^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ -2.5 \end{bmatrix}$ as an approximation to X.

Solution

Second iteration.

The second iteration is $(D+L)X^{(2)} = -UX^{(1)} + B$, or in full

$$\begin{bmatrix} 8 & 0 & 0 \\ 3 & 5 & 0 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & -2 & -4 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} + \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix}$$

Taking this information row by row we see that

$$\begin{aligned} 8x_1^{(2)} &= -2x_2^{(1)} - 4x_3^{(1)} - 16 & \therefore\ & \boxed{x_1^{(2)} = -1.25} \\ 3x_1^{(2)} + 5x_2^{(2)} &= -x_3^{(1)} + 4 & \therefore\ & \boxed{x_2^{(2)} = 2.05} \\ 2x_1^{(2)} + x_2^{(2)} + 4x_3^{(2)} &= -12 & \therefore\ & \boxed{x_3^{(2)} = -2.8875} \end{aligned}$$

Therefore the second iterate approximating X is $X^{(2)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} -1.25 \\ 2.05 \\ -2.8875 \end{bmatrix}$.

Third iteration.

The third iteration is $(D+L)X^{(3)} = -UX^{(2)} + B$, or in full

$$\begin{bmatrix} 8 & 0 & 0 \\ 3 & 5 & 0 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{bmatrix} = \begin{bmatrix} 0 & -2 & -4 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} + \begin{bmatrix} -16 \\ 4 \\ -12 \end{bmatrix}.$$

Taking this information row by row we see that

$$\begin{aligned} 8x_1^{(3)} &= -2x_2^{(2)} - 4x_3^{(2)} - 16 & \therefore\ & \boxed{x_1^{(3)} = -1.0687} \\ 3x_1^{(3)} + 5x_2^{(3)} &= -x_3^{(2)} + 4 & \therefore\ & \boxed{x_2^{(3)} = 2.0187} \\ 2x_1^{(3)} + x_2^{(3)} + 4x_3^{(3)} &= -12 & \therefore\ & \boxed{x_3^{(3)} = -2.9703} \end{aligned}$$

to 4 d.p. Therefore the third iterate approximating X is

$$X^{(3)} = \begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{bmatrix} = \begin{bmatrix} -1.0687 \\ 2.0187 \\ -2.9703 \end{bmatrix}.$$

More iterations ...

Again, there is little to be learned from pushing this further by hand. Putting the procedure on a computer and seeing how it progresses is instructive, however, and the iteration continues as follows:

$$X^{(4)} = \begin{bmatrix} -1.0195 \\ 2.0058 \\ -2.9917 \end{bmatrix}, \quad X^{(5)} = \begin{bmatrix} -1.0056 \\ 2.0017 \\ -2.9976 \end{bmatrix}, \quad X^{(6)} = \begin{bmatrix} -1.0016 \\ 2.0005 \\ -2.9993 \end{bmatrix},$$

$$X^{(7)} = \begin{bmatrix} -1.0005 \\ 2.0001 \\ -2.9998 \end{bmatrix}, \quad X^{(8)} = \begin{bmatrix} -1.0001 \\ 2.0000 \\ -2.9999 \end{bmatrix}, \quad X^{(9)} = \begin{bmatrix} -1.0000 \\ 2.0000 \\ -3.0000 \end{bmatrix}$$

(to 4 d.p.). Subsequent iterates are equal to $X^{(9)}$ to this number of decimal places. The Gauss-Seidel iteration has converged to 4 d.p. in 9 iterations. It took the Jacobi method almost 40 iterations to achieve this!

Task

Carry out two iterations of the Gauss-Seidel method to approximate the solution of

$$\begin{bmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

with the initial guess $X^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Your solution

First iteration

Answer

The first iteration is $(D+L)X^{(1)} = -UX^{(0)} + B$, that is,

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 4 & 0 \\ -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

from which it follows that $X^{(1)} = \begin{bmatrix} 0.75 \\ 0.9375 \\ 1.1719 \end{bmatrix}$.

Your solution

Second iteration

Answer

The second iteration is $(D+L)X^{(1)} = -UX^{(0)} + B$, that is,

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 4 & 0 \\ -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

from which it follows that $X^{(2)} = \begin{bmatrix} 0.7773 \\ 0.9873 \\ 1.1912 \end{bmatrix}$.

2. Do these iterative methods always work?

No. It is not difficult to invent examples where the iteration fails to approach the solution of $AX = B$. The key point is related to matrix norms seen in the preceding Section.
The two iterative methods we encountered above are both special cases of the general form

$$X^{(k+1)} = MX^{(k)} + N.$$

1. For the Jacobi method we choose $M = -D^{-1}(L+U)$ and $N = D^{-1}B$.
2. For the Gauss-Seidel method we choose $M = -(D+L)^{-1}U$ and $N = (D+L)^{-1}B$.

The following Key Point gives the main result.

Key Point 13

For the iterative process $X^{(k+1)} = MX^{(k)} + N$ the iteration will converge to a solution if **the norm of M is less than 1**.

Care is required in understanding what Key Point 13 says. Remember that there are lots of different ways of defining the norm of a matrix (we saw three of them). If you can find a norm (*any norm*) such that the norm of M is less than 1, then the iteration will converge. It doesn't matter if there are other norms which give a value greater than 1, all that matters is that there is one norm that is less than 1.

Key Point 13 above makes no reference to the starting "guess" $X^{(0)}$. The convergence of the iteration is independent of where you start! (Of course, if we start with a really bad initial guess then we can expect to need lots of iterations.)

Task

Show that the Jacobi iteration used to approximate the solution of

$$\begin{bmatrix} 4 & -1 & -1 \\ 1 & -5 & -2 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

is certain to converge. (Hint: calculate the norm of $-D^{-1}(L+U)$.)

Your solution

Answer

The Jacobi iteration matrix is

$$\begin{aligned} -D^{-1}(L+U) &= \begin{bmatrix} 4 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 2 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & -0.2 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 2 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0.25 & 0.25 \\ -0.2 & 0 & 0.4 \\ 0.5 & 0 & 0 \end{bmatrix} \end{aligned}$$

and the infinity norm of this matrix is the maximum of $0.25+0.25$, $0.2+0.4$ and 0.5, that is

$$\| - D^{-1}(L+U)\|_\infty = 0.6$$

which is less than 1 and therefore the iteration will converge.

Guaranteed convergence

If the matrix has the property that it is **strictly diagonally dominant**, which means that the diagonal entry is larger in magnitude than the absolute sum of the other entries on that row, then both Jacobi and Gauss-Seidel are guaranteed to converge. The reason for this is that if A is strictly diagonally dominant then the iteration matrix M will have an infinity norm that is less than 1.

A small system is the subject of Example 20 below. A large system with slow convergence is the subject of Engineering Example 1 on page 62.

Example 20

Show that $A = \begin{bmatrix} 4 & -1 & -1 \\ 1 & -5 & -2 \\ -1 & 0 & 2 \end{bmatrix}$ is strictly diagonally dominant.

Solution

Looking at the diagonal entry of each row in turn we see that

$$\begin{aligned} 4 &> |-1| + |-1| = 2 \\ |-5| &> 1 + |-2| = 3 \\ 2 &> |-1| + 0 = 1 \end{aligned}$$

and this means that the matrix is strictly diagonally dominant.

Given that A above *is* strictly diagonally dominant it is certain that both Jacobi and Gauss-Seidel will converge.

What's so special about strict diagonal dominance?

In many applications we can be certain that the coefficient matrix A will be strictly diagonally dominant. We will see examples of this in HELM 32 and HELM 33 when we consider approximating solutions of differential equations.

Exercises

1. Consider the system

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

   (a) Use the starting guess $X^{(0)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ in an implementation of the Jacobi method to show that $X^{(1)} = \begin{bmatrix} 1.5 \\ -3 \end{bmatrix}$. Find $X^{(2)}$ and $X^{(3)}$.

   (b) Use the starting guess $X^{(0)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ in an implementation of the Gauss-Seidel method to show that $X^{(1)} = \begin{bmatrix} 1.5 \\ -3.25 \end{bmatrix}$. Find $X^{(2)}$ and $X^{(3)}$.

   (Hint: it might help you to know that the exact solution is $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$.)

2. (a) Show that the Jacobi iteration applied to the system

$$\begin{bmatrix} 5 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 \\ 0 & -1 & 5 & -1 \\ 0 & 0 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 7 \\ -10 \\ -6 \\ 16 \end{bmatrix}$$

   can be written

$$X^{(k+1)} = \begin{bmatrix} 0 & 0.2 & 0 & 0 \\ 0.2 & 0 & 0.2 & 0 \\ 0 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0.2 & 0 \end{bmatrix} X^{(k)} + \begin{bmatrix} 1.4 \\ -2 \\ -1.2 \\ 3.2 \end{bmatrix}.$$

   (b) Show that the method is certain to converge and calculate the first three iterations using zero starting values.

   (Hint: the exact solution to the stated problem is $\begin{bmatrix} 1 \\ -2 \\ 1 \\ 3 \end{bmatrix}$.)

Answers

1. (a) $$2x_1^{(1)} = 2 - 1x_2^{(0)} = 2$$

and therefore $x_1^{(1)} = 1.5$

$$2x_2^{(1)} = -5 - 1x_1^{(0)} = -6$$

which implies that $x_2^{(1)} = -3$. These two values give the required entries in $X^{(1)}$. A second and third iteration follow in a similar way to give

$$X^{(2)} = \begin{bmatrix} 2.5 \\ -3.25 \end{bmatrix} \quad \text{and} \quad X^{(3)} = \begin{bmatrix} 2.625 \\ -3.75 \end{bmatrix}$$

(b) $$2x_1^{(1)} = 2 - 1x_2^{(0)} = 3$$

and therefore $x_1^{(1)} = 1.5$. This new approximation to x_1 is used straight away when finding a new approximation to $x_2^{(1)}$.

$$2x_2^{(1)} = -5 - 1x_1^{(1)} = -6.5$$

which implies that $x_2^{(1)} = -3.25$. These two values give the required entries in $X^{(1)}$. A second and third iteration follow in a similar way to give

$$X^{(2)} = \begin{bmatrix} 2.625 \\ -3.8125 \end{bmatrix} \quad \text{and} \quad X^{(3)} = \begin{bmatrix} 2.906250 \\ -3.953125 \end{bmatrix}$$

where $X^{(3)}$ is given to 6 decimal places

2. (a) In this case $D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ and therefore $D^{-1} = \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.2 \end{bmatrix}$.

So the iteration matrix $M = D^{-1} \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0.2 & 0 & 0 \\ 0.2 & 0 & 0.2 & 0 \\ 0 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0.2 & 0 \end{bmatrix}$

and that the Jacobi iteration takes the form

$$X^{(k+1)} = MX^{(k)} + M^{-1} \begin{bmatrix} 7 \\ -10 \\ -6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 & 0.2 & 0 & 0 \\ 0.2 & 0 & 0.2 & 0 \\ 0 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0.2 & 0 \end{bmatrix} X^{(k)} + \begin{bmatrix} 1.4 \\ -2 \\ -1.2 \\ 3.2 \end{bmatrix}$$

as required.

Answers

2(b)

Using the starting values $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = x_4^{(0)} = 0$, the first iteration of the Jacobi method gives

$$\begin{aligned}
x_1^1 &= 0.2x_2^0 + 1.4 = 1.4 \\
x_2^1 &= 0.2(x_1^0 + x_3^0) - 2 = -2 \\
x_3^1 &= 0.2(x_2^0 + x_4^0) - 1.2 = -1.2 \\
x_4^1 &= 0.2x_3^0 + 3.2 = 3.2
\end{aligned}$$

The second iteration is

$$\begin{aligned}
x_1^2 &= 0.2x_2^1 + 1.4 = 1 \\
x_2^2 &= 0.2(x_1^1 + x_3^1) - 2 = -1.96 \\
x_3^2 &= 0.2(x_2^1 + x_4^1) - 1.2 = -0.96 \\
x_4^2 &= 0.2x_3^1 + 3.2 = 2.96
\end{aligned}$$

And the third iteration is

$$\begin{aligned}
x_1^3 &= 0.2x_2^2 + 1.4 = 1.008 \\
x_2^3 &= 0.2(x_1^2 + x_3^2) - 2 = -1.992 \\
x_3^3 &= 0.2(x_2^2 + x_4^2) - 1.2 = -1 \\
x_4^3 &= 0.2x_3^2 + 3.2 = 3.008
\end{aligned}$$

Engineering Example 1

Detecting a train on a track

Introduction

One means of detecting trains is the 'track circuit' which uses current fed along the rails to detect the presence of a train. A voltage is applied to the rails at one end of a section of track and a relay is attached across the other end, so that the relay is energised if no train is present, whereas the wheels of a train will short circuit the relay, causing it to de-energise. Any failure in the power supply or a breakage in a wire will also cause the relay to de-energise, for the system is fail safe. Unfortunately, there is always leakage between the rails, so this arrangement is slightly complicated to analyse.

Problem in words

A 1000 m track circuit is modelled as ten sections each 100 m long. The resistance of 100 m of one rail may be taken to be 0.017 ohms, and the leakage resistance across a 100 m section taken to be 30 ohms. The detecting relay and the wires to it have a resistance of 10 ohms, and the wires from the supply to the rail connection have a resistance of 5 ohms for the pair. The voltage applied at the supply is $4V$. See diagram below. What is the current in the relay?

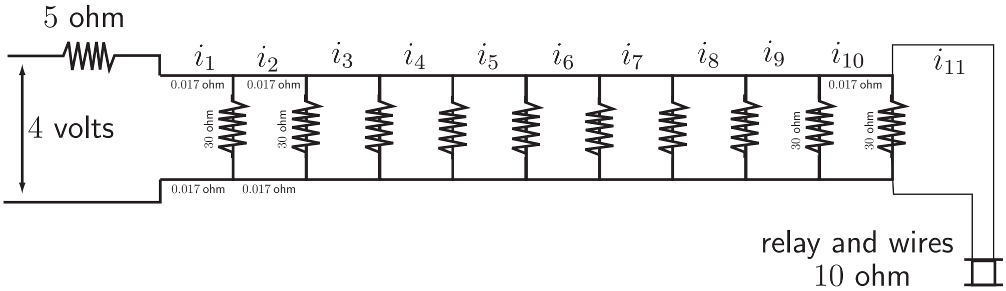

Figure 1

Mathematical statement of problem

There are many ways to apply Kirchhoff's laws to solve this, but one which gives a simple set of equations in a suitable form to solve is shown below. i_1 is the current in the first section of rail (i.e. the one close to the supply), i_2 , $i_3, \ldots i_{10}$, the current in the successive sections of rail and i_{11} the current in the wires to the relay. The leakage current between the first and second sections of rail is $i_1 - i_2$ so that the voltage across the rails there is $30(i_1 - i_2)$ volts. The first equation below uses this and the voltage drop in the feed wires, the next nine equations compare the voltage drop across successive sections of track with the drop in the (two) rails, and the last equation compares the voltage drop across the last section with that in the relay wires.

$$\begin{aligned}
30(i_1 - i_2) + (5.034)i_1 &= 4 \\
30(i_1 - i_2) &= 0.034i_2 + 30(i_2 - i_3) \\
30(i_2 - i_3) &= 0.034i_2 + 30(i_3 - i_4) \\
&\vdots \\
30(i_9 - i_{10}) &= 0.034i_{10} + 30(i_{10} - i_{11}) \\
30(i_{10} - i_{11}) &= 10i_{11}
\end{aligned}$$

These can be reformulated in matrix form as $A\underline{i} = \underline{v}$, where $\underline{v}$ is the 11×1 column vector with first entry 4 and the other entries zero, $\underline{i}$ is the column vector with entries $i_1, i_2, \ldots, i_{11}$ and A is the matrix

$$A = \begin{bmatrix} 35.034 & -30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -30 & 60.034 & -30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -30 & 60.034 & -30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -30 & 60.034 & -30 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -30 & 60.034 & -30 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -30 & 60.034 & -30 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -30 & 60.034 & -30 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -30 & 60.034 & -30 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -30 & 60.034 & -30 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -30 & 60.034 & -30 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -30 & 40 \end{bmatrix}$$

Find the current i_1 in the relay when the input is $4V$, by Gaussian elimination or by performing an L-U decomposition of A.

Mathematical analysis

We solve $A\underline{i} = \underline{v}$ as above, although actually we only want to know i_{11}. Letting M be the matrix A with the column $\underline{v}$ added at the right, as in Section 30.2, then performing Gaussian elimination on M, working to four decimal places gives

$$M = \begin{bmatrix} 35.0340 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.0000 \\ 0 & 34.3447 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3.4252 \\ 0 & 0 & 33.8291 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.9919 \\ 0 & 0 & 0 & 33.4297 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 2.6532 \\ 0 & 0 & 0 & 0 & 33.1118 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 2.3810 \\ 0 & 0 & 0 & 0 & 0 & 32.8534 & -30.0000 & 0 & 0 & 0 & 0 & 2.1572 \\ 0 & 0 & 0 & 0 & 0 & 0 & 32.6396 & -30.0000 & 0 & 0 & 0 & 1.9698 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 32.4601 & -30.0000 & 0 & 0 & 1.8105 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 32.3077 & -30.0000 & 0 & 1.6733 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 32.1769 & -30.0000 & 1.5538 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12.0296 & 1.4487 \end{bmatrix}$$

from which we can calculate that the solution $\underline{i}$ is

$$\underline{i} = \begin{bmatrix} 0.5356 \\ 0.4921 \\ 0.4492 \\ 0.4068 \\ 0.3649 \\ 0.3234 \\ 0.2822 \\ 0.2414 \\ 0.2008 \\ 0.1605 \\ 0.1204 \end{bmatrix}$$

so the current in the relay is 0.1204 amps, or 0.12 A to two decimal places.

You can alternatively solve this problem by an L-U decomposition by finding matrices L and U such that $M = LU$. Here we have

$$L = \begin{bmatrix} 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.8563 & 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.8735 & 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.8868 & 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.8974 & 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.9060 & 1.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.9131 & 1.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.9191 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.9242 & 1.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.9286 & 1.0000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.9323 & 1.0000 \end{bmatrix}$$

and

$$U = \begin{bmatrix} 35.0340 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 34.3447 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 33.8291 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 33.4297 & -30.0000 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 33.1118 & -30.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 32.8534 & -30.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 32.6395 & -30.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 32.4601 & -30.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 32.3076 & -30.0000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 32.1768 & -30.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12.0295 \end{bmatrix}$$

Therefore $U\underline{i} = \begin{bmatrix} 4.0000 \\ 3.4240 \\ 2.9892 \\ 2.6514 \\ 2.3783 \\ 2.1547 \\ 1.9673 \\ 1.8079 \\ 1.6705 \\ 1.5519 \\ 1.4464 \end{bmatrix}$ and hence $\underline{i} = \begin{bmatrix} 0.5352 \\ 0.4917 \\ 0.4487 \\ 0.4064 \\ 0.3644 \\ 0.3230 \\ 0.2819 \\ 0.2411 \\ 0.2006 \\ 0.1603 \\ 0.1202 \end{bmatrix}$ and again the current is found to be 0.12 amps.

Mathematical comment

You can try to solve the equation $A\underline{i} = \underline{v}$ by Jacobi or Gauss-Seidel iteration but in both cases it will take very many iterations (over 200 to get four decimal places). Convergence is very slow because the norms of the relevant matrices in the iteration are only just less than 1. Convergence is nevertheless assured because the matrix A is diagonally dominant.