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 . . .

① revise matrices, especially the material in Workbook 8

② revise determinants

③ revise matrix norms

Learning Outcomes

After completing this Section you should be able to . . .

✓ approximate the solutions of simple systems of equations by iterative methods

✓ assess convergence properties of these 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

An **iterative** method is one in which a sequence of approximations (or **iterates**) is produced. The method is a successful one 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

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

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

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 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.

Solution

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

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

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

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

Solution

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

$$X^{(4)} = \begin{bmatrix} -0.765 \\ 2.39 \\ -2.64 \end{bmatrix}, \quad X^{(5)} = \begin{bmatrix} -1.2775 \\ 1.787 \\ -3.215 \end{bmatrix}, \quad X^{(6)} = \begin{bmatrix} -0.83925 \\ 2.2095 \\ -2.808 \end{bmatrix},$$

to 4 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 decimal places.

After about 40 iterations successive iterates are equal to 4 decimal places. 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 exercise involves calculating just two iterations of the Jacobi method.

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

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

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 boils down to rearranging $(L+D+U)X = B$ slightly differently than 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

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

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

The example below revisits the system of equations we saw earlier in this section.

Example 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

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

(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

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

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

Solution

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

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

to 4 decimal places. Therefore the third iterate approximating X is

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

Solution

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

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

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

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 a 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

The iteration will converge to a solution if **the norm of M is less than 1**.

Care is required in understanding what this keypoint 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.
The Keyp Point 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.)

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

The Jacobi iteration matrix is

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

$$= \begin{bmatrix} 0 & 0.25 & 0.25 \\ -0.2 & 0 & 0.4 \\ 0.5 & 0 & 0 \end{bmatrix}$$

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.

Example 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 later Workbooks 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. 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
$$X^{(k+1)} = \begin{bmatrix} 0 & 0.2 & 0 & 0 \\ 0.2 & 0 & 0.2 & 0 \\ 0 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0.2 & 0 \end{bmatrix} X^{(k)} + \begin{bmatrix} 1.4 \\ -2 \\ -1.2 \\ 3.2 \end{bmatrix}.$$
   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 $\begin{bmatrix} 1 \\ -2 \\ 1 \\ 3 \end{bmatrix}$.)