



አርባ ምንጭ ዩኒቨርሲቲ
Arba Minch University

Numerical Method(Math-2073/53)

Lecture note: Chapter-II

"Mathematics is the most beautiful and most powerful creation of the human spirit."

STEFAN BANACH

Dejen Ketema
Department of Mathematics
dejen.ketema@amu.edu.et

April, 2019

Contents

1	Solving System of Equations	2
1.1	Direct method	3
1.1.1	Vector equation	3
1.1.2	Matrix equation	3
1.1.3	Geometric interpretation	4
1.1.4	Cramer's rule	4
1.1.5	Explicit formulas for small systems	5
1.1.6	Inverse Matrix Method	7
1.1.7	Gaussian Elimination Method	8
1.1.8	Gauss-Jordan Elimination Method	11
1.1.9	Tri-diagonal Matrix	16
1.2	LU Decomposition Method	16
1.2.1	Crout, Doolittle's decomposition method	19
1.2.2	Cholesky Decomposition	19
1.3	Indirect Iteration Method	22
1.3.1	Introduction	22
1.3.2	Jacobi Method	22
1.3.3	Gauss-Seidel Method	25
1.4	Eigenvalue Problem	29
1.4.1	Finding Eigenvalues for 2×2 and 3×3	30
1.4.2	Power Method	32
1.4.3	Inverse Power Method	35
1.5	System of Non-linear Equations	37
1.5.1	Newton Raphson method	37

Chapter 1

Solving System of Equations

In this chapter we consider numerical methods for solving a system of linear equations $Ax = b$. We assume that the given matrix A is real, $n \times n$, and nonsingular and that b is a given real vector in $\mathbb{R}^n$, and we seek a solution x that is necessarily also a vector in $\mathbb{R}^n$. Such problems arise frequently in virtually any branch of science, engineering, economics, or finance.

There is really no single technique that is best for all cases. Nonetheless, the many available numerical methods can generally be divided into two classes: direct methods and iterative methods. The present chapter is devoted to these two methods. In the absence of roundoff error, direct method would yield the exact solution within a finite number of steps.

An example of a problem in electrical engineering that requires a solution of a system of equations is shown in Fig.1.1. Using Kirchhoff's law, the currents $i_1, i_2, i_3, & i_4$ can be determined by solving the following system of four equations:

$$\begin{aligned} 9i_1 - 4i_2 - 2i_3 &= 24 \\ -4i_1 + 17i_2 - 6i_3 - 3i_4 &= -16 \\ -2i_1 - 6i_2 + 14i_3 - 6i_4 &= 0 \\ -3i_2 - 6i_3 + 11i_4 &= 18 \end{aligned} \tag{1.1}$$

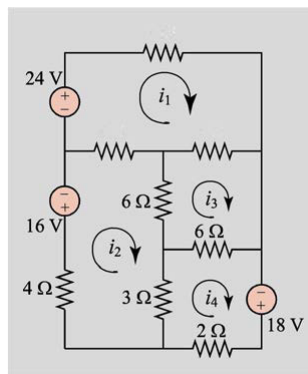


Figure 1.1: Electrical circuit.

1.1 Direct method

Definition 1.1

linear equation in the variables $x_1, x_2, \dots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

where the coefficients $a_1, a_2, \dots, a_n$ and b are constant real or complex numbers. The constant a_i is called the **coefficient** of x_i ; and b is called the **constant term** of the equation.

A **system of linear equations** (or **linear system**) is a finite collection of linear equations in same variables. For instance, a linear system of m equations in n variables $x_1, x_2, \dots, x_n$ can be written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (1.2)$$

where $x_1, x_2, \dots, x_n$ are the unknowns, $a_{11}, a_{12}, \dots, a_{mn}$ are the coefficients of the system, and $b_1, b_2, \dots, b_m$ the constant terms.

1.1.1 Vector equation

One extremely helpful view is that each unknown is a weight for a column vector in a linear combination.

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (1.3)$$

1.1.2 Matrix equation

The vector equation is equivalent to a matrix equation of the form

$$A\mathbf{x} = \mathbf{b}$$

where A is an $m \times n$ matrix, $\mathbf{x}$ is a column vector with n entries, and $\mathbf{b}$ is a column vector with m entries.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (1.4)$$

A **solution** of a linear system (1.2) is a tuple $(s_1, s_2, \dots, s_n)$ of numbers that makes each equation a true statement when the values $(s_1, s_2, \dots, s_n)$ are substituted for $x_1, x_2, \dots, x_n$, respectively. The set of all solutions of a linear system is called the **solution set** of the system.



Theorem 1.1

Any system of linear equations has one of the following exclusive conclusions.

- (a) No solution.
- (b) Unique solution.
- (c) Infinitely many solutions.

A linear system is said to be **consistent** if it has at least one solution; and is said to be **inconsistent** if it has no solution.

1.1.3 Geometric interpretation

For a system involving two variables (x and y), each linear equation determines a line on the xy -plane. Because a solution to a linear system must satisfy all of the equations, the solution set is the intersection of these lines, and is hence either a line, a single point, or the empty set.

For three variables, each linear equation determines a plane in three-dimensional space, and the solution set is the intersection of these planes. Thus the solution set may be a plane, a line, a single point, or the empty set.

For n variables, each linear equation determines a hyperplane in n -dimensional space. The solution set is the intersection of these hyperplanes, which may be a flat of any dimension.

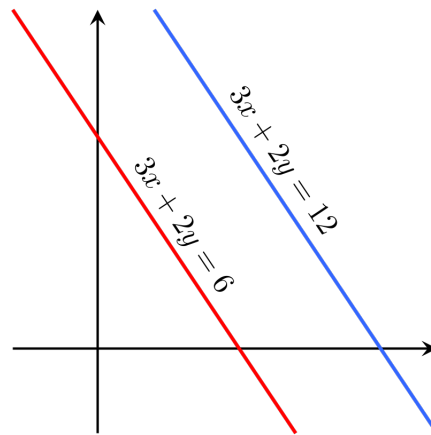


Figure 1.2: The equations $3x + 2y = 6$ and $3x + 2y = 12$ are (**inconsistent**).

1.1.4 Cramer's rule

Consider a system of n linear equations for n unknowns, represented in matrix multiplication form as follows:

$$Ax = b$$

where the $n \times n$ matrix A has a nonzero determinant, and the vector $x = (x_1, \dots, x_n)^T$ is the column vector of the variables. Then the theorem states that in this case the system has a unique solution, whose individual values for the unknowns are given by:

$$x_i = \frac{\det(A_i)}{\det(A)} \quad i = 1, \dots, n$$



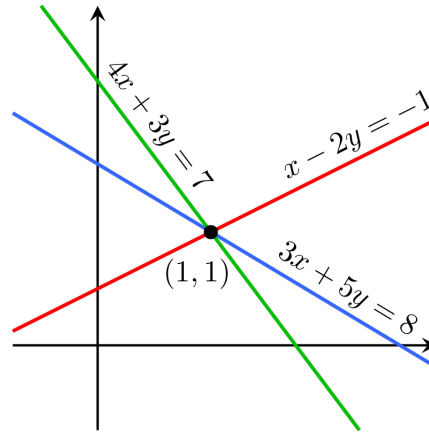


Figure 1.3: The equations $x - 2y = -1$, $3x + 5y = 8$, and $4x + 3y = 7$ are linearly dependent(**consistent**).

where A_i is the matrix formed by replacing the i -th column of A by the column vector b .

A more general version of Cramer's rule considers the matrix equation

$$AX = B$$

where the $n \times n$ matrix A has a nonzero determinant, and X, B are $n \times m$ matrices. Given sequences $1 \leq i_1 < i_2 < \dots < i_k \leq n$ and $1 \leq j_1 < j_2 < \dots < j_k \leq m$, let $X_{I,J}$ be the $k \times k$ submatrix of X with rows in $I := (i_1, \dots, i_k)$ and columns in $J := (j_1, \dots, j_k)$. Let $A_B(I, J)$ be the $n \times n$ matrix formed by replacing the i_s column of A by the j_s column of B , for all $s = 1, \dots, k$. Then

$$\det X_{I,J} = \frac{\det(A_B(I, J))}{\det(A)}.$$

In the case $k = 1$, this reduces to the normal Cramer's rule.

1.1.5 Explicit formulas for small systems

Consider the linear system

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$$

which in matrix format is

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Assume $a_1b_2 - b_1a_2$ nonzero. Then, with help of determinants, x and y can be found with Cramer's rule as

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} = \frac{c_1b_2 - b_1c_2}{a_1b_2 - b_1a_2}, \quad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} = \frac{a_1c_2 - c_1a_2}{a_1b_2 - b_1a_2}.$$

The rules for 3×3 matrices are similar. Given

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases}$$



which in matrix format is

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}.$$

Then the values of x, y and z can be found as follows:

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}, \quad \text{and } z = \frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}.$$

Example 1.1

Solve the following system by Cramer's rule.

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 19 \\ x_1 + 2x_2 + x_3 &= 4 \\ 3x_1 - x_2 + x_3 &= 9 \end{aligned}$$

Solution: The coefficient matrix is $A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \\ 3 & -1 & 1 \end{bmatrix}$ and column matrix $b = \begin{bmatrix} 19 \\ 4 \\ 9 \end{bmatrix}$, then

$\det(A) = \begin{vmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \\ 3 & -1 & 1 \end{vmatrix} = 4 + 9 - 4 - 24 - 3 + 2 = -16 \neq 0$ then the system has unique solution.

$$A_1 = \begin{bmatrix} 19 & 3 & 4 \\ 4 & 2 & 1 \\ 9 & -1 & 1 \end{bmatrix} \quad \& \quad \det(A_1) = \begin{vmatrix} 19 & 3 & 4 \\ 4 & 2 & 1 \\ 9 & -1 & 1 \end{vmatrix} = 38 + 27 - 16 - 72 - 12 + 19 = -16$$

$$A_2 = \begin{bmatrix} 2 & 19 & 4 \\ 1 & 4 & 1 \\ 3 & 9 & 1 \end{bmatrix} \quad \& \quad \det(A_2) = \begin{vmatrix} 2 & 19 & 4 \\ 1 & 4 & 1 \\ 3 & 9 & 1 \end{vmatrix} = 8 + 57 + 36 - 48 - 19 - 18 = 16$$

$$A_3 = \begin{bmatrix} 2 & 3 & 19 \\ 1 & 2 & 4 \\ 3 & -1 & 9 \end{bmatrix} \quad \& \quad \det(A_3) = \begin{vmatrix} 2 & 3 & 19 \\ 1 & 2 & 4 \\ 3 & -1 & 9 \end{vmatrix} = 36 + 36 - 19 - 114 - 27 + 8 = -80$$

$$\therefore x_1 = \frac{\det(A_1)}{\det(A)} = \frac{-16}{-16} = 1$$

$$x_2 = \frac{\det(A_2)}{\det(A)} = \frac{16}{-16} = -1$$

$$x_3 = \frac{\det(A_3)}{\det(A)} = \frac{-80}{-16} = 5.$$

This is the solution of the system.



Exercise 1.1

Use Cramer's Rule to solve each for each of the variables.

(a)

$$\begin{aligned} x - y &= 4 \\ -x + 2y &= -7 \end{aligned}$$

(b)

$$\begin{aligned} -2x + y &= -2 \\ x - 2y &= -2 \end{aligned}$$

(c)

$$\begin{aligned} 2x + y + z &= 1 \\ 3x \quad \quad + z &= 4 \\ x - y - z &= 2 \end{aligned}$$

1.1.6 Inverse Matrix Method

Let $AX = b$ is a system of $\mathbf{n}$ linear equations with $\mathbf{n}$ unknowns and A is invertible, then the system has unique solution given by inversion method $X = A^{-1}b$.

$$A^{-1} = \frac{\text{adj}(A)}{\det(A)}$$

Note:- When A is not square or is singular, the system may not have a solution or may have more than one solution.

Example 1.2

Use the inverse of the coefficient matrix to solve the following system

$$\begin{aligned} 3x_1 + x_2 &= 6 \\ -x_1 + 2x_2 + 2x_3 &= -7 \\ 5x_1 - x_3 &= 10 \end{aligned}$$

Solution: Okay, let's first write down the matrix form of this system.

$$\begin{bmatrix} 3 & 9 & 0 \\ -1 & 2 & 2 \\ 5 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -7 \\ 10 \end{bmatrix}$$

Now, we found the inverse of the coefficient matrix by using methods of finding Inverses and is the following;



Example

$$A = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 2 & 2 \\ 5 & 0 & -1 \end{bmatrix} \Rightarrow C_A = \begin{bmatrix} -2 & 9 & -10 \\ 1 & 3 & 5 \\ 2 & -6 & 7 \end{bmatrix} \Rightarrow \text{adj}(A) = \begin{bmatrix} 2 & -1 & 2 \\ 9 & -3 & -6 \\ -10 & 5 & 7 \end{bmatrix}$$

and $\det(A) = 3(-2) + 1(9) + 0(-10) = -6 + 9 = 3$, then

$$A^{-1} = 1/3 \begin{bmatrix} 2 & -1 & 2 \\ 9 & -3 & -6 \\ -10 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 & 2/3 \\ 3 & -1 & -2 \\ -10/3 & 5/3 & 7/3 \end{bmatrix}$$

$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 & 2/3 \\ 3 & -1 & -2 \\ -10/3 & 5/3 & 7/3 \end{bmatrix} \begin{bmatrix} 6 \\ -7 \\ 10 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 5 \\ -25/3 \end{bmatrix}$$

Now each of the entries of X are $x_1 = 1/3$, $x_2 = 5$ and $x_3 = -25/3$

1.1.7 Gaussian Elimination Method

In this section we show the following:

- How to solve linear equations when A is in upper triangular form. The algorithm is called backward substitution.
- How to transform a general system of linear equations into an upper triangular form, to which backward substitution can be applied. The algorithm is called Gaussian elimination.

A triangular matrix is a special kind of square matrix. A square matrix is called lower triangular if all the entries above the main diagonal are zero. Similarly, a square matrix is called upper triangular if all the entries below the main diagonal are zero. A triangular matrix is one that is either lower triangular or upper triangular. A matrix that is both upper and lower triangular is called a diagonal matrix.

A matrix of the form

$$L = \begin{bmatrix} \ell_{1,1} & & & & 0 \\ \ell_{2,1} & \ell_{2,2} & & & \\ \ell_{3,1} & \ell_{3,2} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n,1} & \ell_{n,2} & \dots & \ell_{n,n-1} & \ell_{n,n} \end{bmatrix}$$

is called a lower triangular matrix or left triangular matrix, and analogously a matrix of the form

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \dots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \dots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$



is called an upper triangular matrix or right triangular matrix.

Forward and back substitution

A matrix equation in the form $\mathbf{Lx} = \mathbf{b}$ or $\mathbf{Ux} = \mathbf{b}$ is very easy to solve by an iterative process called **forward substitution** for lower triangular matrices and analogously **back substitution** for upper triangular matrices. The process is so called because for lower triangular matrices, one first computes x_1 , then substitutes that forward into the next equation to solve for x_2 , and repeats through to x_n . In an upper triangular matrix, one works backwards, first computing x_n , then substituting that back into the previous equation to solve for x_{n-1} , and repeating through x_1 . Notice that this does not require inverting the matrix.

Forward substitution The matrix equation $Lx = b$ can be written as a system of linear equations

$$\begin{array}{rcl} \ell_{1,1}x_1 & & = b_1 \\ \ell_{2,1}x_1 + \ell_{2,2}x_2 & & = b_2 \\ \vdots & \ddots & \vdots \\ \ell_{m,1}x_1 + \ell_{m,2}x_2 + \cdots + \ell_{m,m}x_m & = & b_m \end{array}$$

Observe that the first equation ($\ell_{1,1}x_1 = b_1$) only involves x_1 , and thus one can solve for x_1 directly. The second equation only involves x_1 and x_2 , and thus can be solved once one substitutes in the already solved value for x_1 . Continuing in this way, the k -th equation only involves $x_1, \dots, x_k$, and one can solve for x_k using the previously solved values for $x_1, \dots, x_{k-1}$.

The resulting formulas are:

$$\begin{aligned} x_1 &= \frac{b_1}{\ell_{1,1}}, \\ x_2 &= \frac{b_2 - \ell_{2,1}x_1}{\ell_{2,2}}, \\ &\vdots \\ x_m &= \frac{b_m - \sum_{i=1}^{m-1} \ell_{m,i}x_i}{\ell_{m,m}}. \end{aligned}$$

A matrix equation with an upper triangular matrix U can be solved in an analogous way, only working backwards.

Backward Substitution.

Given an upper triangular matrix A and a right-hand-side $\mathbf{b}$,

$$\begin{aligned} &\text{for } k = n : -1 : 1 \\ &\quad x_k = b_k - \frac{\sum_{j=k+1}^n a_{kj}x_j}{a_{kk}} \\ &\text{end} \end{aligned}$$

Gauss elimination method is used to solve system of linear equations. In this method the linear system of equation is reduced to an upper triangular system by using successive elementary row operations. Finally we solve the value variables by using back ward substitution method. This method will be fail if any of the pivot element a_{ii} , $i = 1, 2, \dots, n$ becomes zero. In



such case we re-write equation in such manner so that pivots are non zero. This procedure is called pivoting.

Consider system $AX = b$

Step 1: Form the augmented matrix $[A|b]$

Step 2: Transform $[A|b]$ to row echelon form $[U|d]$ using row operations.

Step 3: Solve the system $UX = d$ by back substitution.

The following row operations on the augmented matrix of a system produce the augmented matrix of an equivalent system, i.e., a system with the same solution as the original one.

- Interchange any two rows.
- Multiply each element of a row by a nonzero constant.
- Replace a row by the sum of itself and a constant multiple of another row of the matrix.

For these row operations, we will use the following notations.

- $R_i \leftrightarrow R_j$ means: Interchange row i and row j .
- αR_i means: Replace row i with α times row i .
- $R_i + \alpha R_j$ means: Replace row i with the sum of row i and α times row j .

Example 1.3

Solve the following system using Gauss elimination method.

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 5 \\ 4x_1 + 14x_2 + 12x_3 &= 10 \\ 6x_1 + x_2 + 5x_3 &= 9 \end{aligned}$$

Solution: The augmented matrix of the system is

$$\begin{bmatrix} 2 & -3 & 1 & 5 \\ 4 & 14 & 12 & 10 \\ 6 & 1 & 5 & 9 \end{bmatrix}$$

Applying, elementary row operations on this matrix to change into its echelon form.

$$\begin{bmatrix} 2 & -3 & 1 & 5 \\ 4 & 14 & 12 & 10 \\ 6 & 1 & 5 & 9 \end{bmatrix} \begin{array}{l} R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1 \end{array} \begin{bmatrix} 2 & -3 & 1 & 5 \\ 0 & 20 & 10 & 0 \\ 0 & 10 & 2 & -6 \end{bmatrix}$$

$$R_3 \rightarrow R_3 - 1/2 R_2 \quad \begin{bmatrix} 2 & -3 & 1 & 5 \\ 0 & 20 & 10 & 0 \\ 0 & 0 & -3 & -6 \end{bmatrix}$$

Since $\text{rank}(A) = \text{rank}(A) = 3 = n$ the solution exists and is unique.

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 5 \\ 20x_2 + 10x_3 &= 0 \\ -3x_3 &= -6 \end{aligned}$$

From this we get $x_3 = 2$. And using back substitution we have $x_2 = -1$ and $x_1 = 0$

Hence $(0, -1, 2)$ is the solution of the system.



Exercise 1.2

Solve the following system of four equations using the Gauss elimination method.

$$\begin{aligned}4x_1 - 2x_2 - 3x_3 + 6x_4 &= 12 \\ -6x_1 + 7x_2 + 6.5x_3 - 6x_4 &= -6.5 \\ x_1 + 7.5x_2 + 6.25x_3 + 5.5x_4 &= 16 \\ -12x_1 + 22x_2 + 15.5x_3 - x_4 &= 17\end{aligned}$$

1.1.8 Gauss-Jordan Elimination Method

The Gauss-Jordan elimination method to solve a system of linear equations is described in the following steps.

1. Write the augmented matrix of the system.
2. Use row operations to transform the augmented matrix in the form described below, which is called the reduced row echelon form (RREF).
 - (a) The rows (if any) consisting entirely of zeros are grouped together at the bottom of the matrix.
 - (b) In each row that does not consist entirely of zeros, the leftmost nonzero element is a 1 (called a leading 1 or a pivot).
 - (c) Each column that contains a leading 1 has zeros in all other entries.
 - (d) The leading 1 in any row is to the left of any leading 1's in the rows below it.
3. Stop process in step 2 if you obtain a row whose elements are all zeros except the last one on the right. In that case, the system is inconsistent and has no solutions. Otherwise, finish step 2 and read the solutions of the system from the final matrix.

Note: When doing step 2, row operations can be performed in any order. Try to choose row operations so that as few fractions as possible are carried through the computation. This makes calculation easier when working by hand.



Example 1.4

Given the following linear system with corresponding augmented matrix:

$$\begin{aligned}3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15\end{aligned}$$

$$\begin{bmatrix}0 & 3 & -6 & 6 & 4 & -5 \\3 & -7 & 8 & -5 & 8 & 9 \\3 & -9 & 12 & -9 & 6 & 15\end{bmatrix}$$

To solve this system, the matrix has to be reduced into reduced echelon form.

Step 1: Switch row 1 and row 3. All leading zeros are now below non-zero leading entries.

$$\begin{bmatrix}3 & -9 & 12 & -9 & 6 & 15 \\3 & -7 & 8 & -5 & 8 & 9 \\0 & 3 & -6 & 6 & 4 & -5\end{bmatrix}$$

Step 2: Set row 2 to row 2 plus (-1) times row 1. In other words, subtract row 1 from row 2. This will eliminate the first entry of row 2.

$$\begin{bmatrix}3 & -9 & 12 & -9 & 6 & 15 \\0 & 2 & -4 & 4 & 2 & -6 \\0 & 3 & -6 & 6 & 4 & -5\end{bmatrix}$$



Example

Step 3: Multiply row 2 by 3 and row 3 by 2. This will eliminate the first entry of row 3.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 6 & -12 & 12 & 6 & -18 \\ 0 & 6 & -12 & 12 & 8 & -10 \end{bmatrix}$$

Step 4: Set row 3 to row 3 plus (-1) times row 2. In other words, subtract row 2 from row 3. This will eliminate the second entry of row 3.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 6 & -12 & 12 & 6 & -18 \\ 0 & 0 & 0 & 0 & 2 & 8 \end{bmatrix}$$

Step 5: Multiply each row by the reciprocal of its first non-zero value. This will make every row start with a 1.

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

The matrix is now in row echelon form: All nonzero rows are above any rows of all zeros (there are no zero rows), each leading entry of a row is in a column to the right of the leading entry of the row above it and all entries in a column below a leading entry are zeros.

As can and will be shown later, from this form one can observe that the system has infinitely many solutions. To get those solutions, the matrix is further reduced into reduced echelon form.

Step 6: Set row 2 to row 2 plus (-1) times row 3 and row 1 to row 1 plus (-2) times row 3. This will eliminate the entries above the leading entry of row 3.

$$\begin{bmatrix} 1 & -3 & 4 & -3 & 0 & -3 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$



Example

Step 7: Set row 1 to row 1 plus 3 times row 2. This eliminates the entry above the leading entry of row 2.

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

This is a reduced echelon form, since the leading entry in each nonzero row is 1 and each leading 1 is the only nonzero entry in its column.

From this the solution of the system can be read:

$$x_1 - 2x_3 + 3x_4 = -24$$

$$x_2 - 2x_3 + 2x_4 = -7$$

$$x_5 = 4$$

Those equations can be solved for x_1, x_2 and x_5 :

$$x_1 = 2x_3 - 3x_4 - 24$$

$$x_2 = 2x_3 - 2x_4 - 7$$

$$x_5 = 4$$

This is the solution of the system. The variables x_3 and x_4 can take any value and are so called free variables. The solution is valid for any x_3 and x_4 .

Exercise 1.3

Solve the following system by using the Gauss-Jordan elimination method.

1. $x + y + z = 5$

$$2x + 3y + 5z = 8$$

$$4x + 5z = 2$$

2. $x + 2y - 3z = 2$

$$6x + 3y - 9z = 6$$

$$7x + 14y - 21z = 13$$

3. $4y + z = 2$

$$2x + 6y - 2z = 3$$

$$4x + 8y - 5z = 4$$

4. Ava invests a total of \$10,000 in three accounts, one paying 5 interest, another paying 8 interest, and the third paying 9 interest. The annual interest earned on the three investments last year was \$770 . The amount invested at 9 was twice the amount invested at 5 . How much was invested at each rate?

1.1.9 Tri-diagonal Matrix

A tridiagonal system for n unknowns may be written as

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i,$$



where $a_1 = 0$ and $c_n = 0$.

$$\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{bmatrix}.$$

Exercise 1.4

1. Solve the following linear system of equation by using Cramer's rule, Gaussian elimination method, and inverse method.

$$\begin{aligned} 2x_1 + 5x_2 + 3x_3 &= 9 \\ (a) \quad 3x_1 + x_2 + 2x_3 &= 3 \\ x_1 + 2x_2 - x_3 &= 6 \end{aligned}$$

$$\begin{aligned} x + z &= 1 \\ (b) \quad 2x + y + z &= 0 \\ x + y + 2z &= 1 \end{aligned}$$

$$\begin{aligned} x + 2y + z &= 3 \\ (c) \quad 2x + 5y - z &= -4 \\ 3x - 2y - z &= 5 \end{aligned}$$

2. Use rank of matrix to determine the values of a , b and c so that the following system has:

a) no solution	b) more than one solution	c) a unique solution and solve it.
$1x + y - bz = 1$	$x + 2y - 3z = a$	$x - 2y + bz = 3$
i) $2x + 3y + az = 3$	ii) $2x + 6y - 11z = b$	iii) $ax + 2z = 2$
$x + ay + 3z = 2$	$x - 2y + 7z = c$	$5x + 2y = 2$

1.2 LU Decomposition Method

Definition 1.2

A permutation matrix is a square binary matrix that has exactly one entry of 1 in each row and each column and 0's elsewhere.

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$



Definition 1.3

Let A be a square matrix. **An LU factorization** refers to the factorization of A , with proper row and/or column orderings or permutations, into two factors, a unit lower triangular matrix L and an upper triangular matrix U ,

$$A = LU,$$

For example, for a 3-by-3 matrix A , its LU decomposition looks like this:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}.$$

Without a proper ordering or permutations in the matrix, the factorization may fail to materialize. For example, it is easy to verify (by expanding the matrix multiplication) that $a_{11} = l_{11}u_{11}$. If $a_{11} = 0$, then at least one of l_{11} and u_{11} has to be zero, which implies either L or U is singular. This is impossible if A is nonsingular (invertible). This is a procedural problem. It can be removed by simply reordering the rows of A so that the first element of the permuted matrix is nonzero. The same problem in subsequent factorization steps can be removed the same way; see the basic procedure below.



Example 1.5

We factorize the following 2-by-2 matrix:

$$\begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}.$$

One way to find the LU decomposition of this simple matrix would be to simply solve the linear equations by inspection. Expanding the matrix multiplication gives

$$\begin{aligned} l_{11} \cdot u_{11} + 0 \cdot 0 &= 4 \\ l_{11} \cdot u_{12} + 0 \cdot u_{22} &= 3 \\ l_{21} \cdot u_{11} + l_{22} \cdot 0 &= 6 \\ l_{21} \cdot u_{12} + l_{22} \cdot u_{22} &= 3. \end{aligned}$$

This system of equations is under determined. In this case any two non-zero elements of L and U matrices are parameters of the solution and can be set arbitrarily to any non-zero value. Therefore, to find the unique LU decomposition, it is necessary to put some restriction on L and U matrices. For example, we can conveniently require the lower triangular matrix L to be a unit triangular matrix (i.e. set all the entries of its main diagonal to ones). Then the system of equations has the following solution:

$$\begin{aligned} l_{21} &= 1.5 \\ u_{11} &= 4 \\ u_{12} &= 3 \\ u_{22} &= -1.5 \end{aligned}$$

Substituting these values into the LU decomposition above yields

$$\begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1.5 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 0 & -1.5 \end{bmatrix}.$$

Factorization is particularly useful in the repetitive solution of the linear system

$$Ax = b$$

where we vary the r.h.s column vector b , but keep A fixed. For example b may represent a signal that we want to process to get x . If we are handed a whole series of vectors b , it is useful to have an efficient way to get the output series of vectors x . With LU factorization, the solution is found in two steps:

$$LUx = b \quad (1)$$

implies

$$Ly = b \quad (2)$$

and

$$Ux = y.$$

Since L and U are triangular, these two problems are very easy to solve: First solve (1) to get y . Then solve (2) to get x .



LU Decomposition Algorithm:

Given a real nonsingular matrix A , apply LU decomposition first:

$$A = LU.$$

Given also a right-hand-side vector b :

1. Forward substitution: solve

$$Ly = b.$$

2. Backward substitution: solve

$$Ux = y.$$

Definition 1.4

An LDU decomposition is a decomposition of the form

$$A = LDU,$$

where D is a diagonal matrix and L and U are unit triangular matrices, meaning that all the entries on the diagonals of L and U are one.

1.2.1 Crout, Doolittle's decomposition method

The **Crout matrix decomposition algorithm** differs slightly from the Doolittle method. Doolittle's method returns a unit lower triangular matrix and an upper triangular matrix, while the Crout method returns a lower triangular matrix and a unit upper triangular matrix.

So, if a matrix decomposition of a matrix A is such that:

$$A = LDU$$

being L a unit lower triangular matrix, D a diagonal matrix and U a unit upper triangular matrix, then Doolittle's method produces

$$A = L(DU)$$

and Crout's method produces

$$A = (LD)U.$$

being L a lower triangular matrix, D a diagonal matrix and U a normalised upper triangular matrix.

1.2.2 Cholesky Decomposition**Definition 1.5**

Symmetric matrix is a square matrix that is equal to its transpose. Formally, matrix A is symmetric if

$$A = A^T.$$



Definition 1.6

A symmetric $n \times n$ real matrix M is said to be **positive definite** if the scalar $z^T M z$ is positive for every non-zero column vector z of n real numbers. Here z^T denotes the transpose of z .

This factorization is known as Cholesky's method, and A can be factored in the form

$$A = LL^T$$

where L is a lower triangular matrix. The construction of L is similar to the one used for Crout's method. Multiplying L by L^T and setting the result equal to A gives

$$l_{ii} = \left[a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right]^{1/2}, i = 1, 2, 3, \dots, n$$

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}}{l_{jj}}, \quad i = j+1, j+2, \dots, j+n, \quad j = 1, 2, 3, \dots, n$$

Example 1.6

Here is the Cholesky decomposition of a symmetric real matrix:

$$\begin{pmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 6 & 1 & 0 \\ -8 & 5 & 3 \end{pmatrix} \begin{pmatrix} 2 & 6 & -8 \\ 0 & 1 & 5 \\ 0 & 0 & 3 \end{pmatrix}.$$

And here is its LDL^T decomposition:

$$\begin{pmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -4 & 5 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} 1 & 3 & -4 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}.$$



Example 1.7

Solve the system of equations

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 8 & 22 \\ 3 & 22 & 82 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ -10 \end{bmatrix}$$

Using the cholesky method. We write

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 8 & 22 \\ 3 & 22 & 82 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix} = \begin{bmatrix} l_{11}^2 & l_{11}l_{21} & l_{11}l_{31} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & l_{21}l_{31} + l_{22}l_{32} \\ l_{31}l_{11} & l_{31}l_{21} + l_{32}l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}$$

Comparing the corresponding elements on both sides, we get First row $l_{11}^2 = 1$, or $l_{11} = 1$

$l_{11}l_{21} = 2$, or $l_{21} = 2$

$l_{11}l_{31} = 3$, or $l_{31} = 3$

second row $l_{21}^2 + l_{22}^2 = 8$, or $l_{22} = 2$

$l_{31}l_{21} + l_{32}l_{22} = 22$ or $l_{32} = 8$

Third row $l_{31}^2 + l_{32}^2 + l_{33}^2 = 82$ or $l_{33} = 3$

Hence we get $A = LL^T$

Where $L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 8 & 3 \end{bmatrix}$

We write the given system of equations as

$$LL^T x = b$$

$$Ly = b \quad \text{and} \quad L^T x = y.$$

From $Ly = b$, we obtain

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 8 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ -10 \end{bmatrix}$$

or $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix}$

From $L^T x = y$, we obtained

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 8 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 & -2 & -3 \end{bmatrix}.$$

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



Exercise 1.5

Determine if the following matrix is hermitian positive definite. Also find its Cholesky factorization if possible

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ 1 & 3 & 2 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 8 & 0 \\ 2 & 0 & 24 \end{bmatrix}$$

1.3 Indirect Iteration Method

Learning Outcomes

After studying this notebook you should be able to

- Give the motivation for using iterative methods to solve linear equations, as opposed to direct solver like Gaussian Elimination.
- Describe the algorithm for the Jacobi Iteration method
- Describe the algorithm for the Gauss-Seider and relaxed Gauss-Seider methods.
- Test whether a given linear equation can be solved by using the Jacobi or Gauss-Seider methods.

1.3.1 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.

1.3.2 Jacobi Method

Solving of the system of equations, we assume that the quantities a_{ii} in the system are pivot elements. The the system equation may be written as:

$$\begin{cases} a_{11}x_1 &= b_1 - (a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n) \\ a_{22}x_2 &= b_2 - (a_{21}x_1 + a_{23}x_3 + \cdots + a_{2n}x_n) \\ a_{33}x_3 &= b_3 - (a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n) \\ &\vdots \\ a_{nn}x_n &= b_n - (a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn-1}x_{n-1}) \end{cases} \quad (1.5)$$

The Jacobi iteration method can be defined as:

$$\begin{cases} x_1^{k+1} &= \frac{1}{a_{11}}(b_1 - a_{12}x_2^k + a_{13}x_3^k + \cdots + a_{1n}x_n^k) \\ x_2^{k+1} &= \frac{1}{a_{22}}(b_2 - a_{21}x_1^k + a_{23}x_3^k + \cdots + a_{2n}x_n^k) \\ x_3^{k+1} &= \frac{1}{a_{33}}(b_3 - a_{31}x_1^k + a_{32}x_2^k + \cdots + a_{3n}x_n^k) \\ &\vdots \\ x_n^{k+1} &= \frac{1}{a_{nn}}(b_n - a_{n1}x_1^k + a_{n2}x_2^k + \cdots + a_{nn-1}x_{n-1}^k) \end{cases} \quad (1.6)$$



Matrix form

Let

$$A\mathbf{x} = \mathbf{b}$$

be a square system of n linear equations, where:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Then A can be decomposed into a diagonal component D , and the remainder R :

$$A = D + R \quad \text{where} \quad D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix}.$$

The solution is then obtained iteratively via

$$\mathbf{x}^{(k+1)} = D^{-1}(\mathbf{b} - R\mathbf{x}^{(k)}),$$

where $\mathbf{x}^{(k)}$ is the k th approximation or iteration of $\mathbf{x}$ and $\mathbf{x}^{(k+1)}$ is the next or $k+1$ iteration of $\mathbf{x}$. The element-based formula is thus:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij} x_j^{(k)} \right), \quad i = 1, 2, \dots, n.$$

The computation of $x_i(k+1)$ requires each element in $x(k)$ except itself.

Convergence

The standard convergence condition (for any iterative method) is when the spectral radius of the iteration matrix is less than 1:

$$\rho(D^{-1}R) < 1.$$

A sufficient (but not necessary) condition for the method to converge is that the matrix A is strictly or irreducibly diagonally dominant. Strict row diagonal dominance means that for each row, the absolute value of the diagonal term is greater than the sum of absolute values of other terms:

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

The Jacobi method sometimes converges even if these conditions are not satisfied.



Example 1.8

A linear system of the form $Ax = b$ with initial estimate $x^{(0)}$ is given by

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 7 \end{bmatrix}, \quad b = \begin{bmatrix} 11 \\ 13 \end{bmatrix} \quad \text{and} \quad x^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

We use the equation $x^{(k+1)} = D^{-1}(b - Rx^{(k)})$, described above, to estimate x . First, we rewrite the equation in a more convenient form $D^{-1}(b - Rx^{(k)}) = Tx^{(k)} + C$, where $T = -D^{-1}R$ and $C = D^{-1}b$. Note that $R = L + U$ where L and U are the strictly lower and upper parts of A . From the known values

$$D^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/7 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 \\ 5 & 0 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

we determine $T = -D^{-1}(L + U)$ as

$$T = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/7 \end{bmatrix} \left\{ \begin{bmatrix} 0 & 0 \\ -5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \right\} = \begin{bmatrix} 0 & -1/2 \\ -5/7 & 0 \end{bmatrix}.$$

Further, C is found as

$$C = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/7 \end{bmatrix} \begin{bmatrix} 11 \\ 13 \end{bmatrix} = \begin{bmatrix} 11/2 \\ 13/7 \end{bmatrix}.$$

With T and C calculated, we estimate x as $x^{(1)} = Tx^{(0)} + C$:

$$x^{(1)} = \begin{bmatrix} 0 & -1/2 \\ -5/7 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 11/2 \\ 13/7 \end{bmatrix} = \begin{bmatrix} 5.0 \\ 8/7 \end{bmatrix} \approx \begin{bmatrix} 5 \\ 1.143 \end{bmatrix}.$$

The next iteration yields

$$x^{(2)} = \begin{bmatrix} 0 & -1/2 \\ -5/7 & 0 \end{bmatrix} \begin{bmatrix} 5.0 \\ 8/7 \end{bmatrix} + \begin{bmatrix} 11/2 \\ 13/7 \end{bmatrix} = \begin{bmatrix} 69/14 \\ -12/7 \end{bmatrix} \approx \begin{bmatrix} 4.929 \\ -1.714 \end{bmatrix}.$$

This process is repeated until convergence (i.e., until $\|Ax^{(n)} - b\|$ is small). The solution after 25 iterations is

$$x = \begin{bmatrix} 7.111 \\ -3.222 \end{bmatrix}.$$



Example 1.9

Suppose we are given the following linear system:

$$\begin{aligned} 10x_1 - x_2 + 2x_3 &= 6, \\ -x_1 + 11x_2 - x_3 + 3x_4 &= 25, \\ 2x_1 - x_2 + 10x_3 - x_4 &= -11, \\ 3x_2 - x_3 + 8x_4 &= 15. \end{aligned}$$

If we choose $(0, 0, 0, 0)$ as the initial approximation, then the first approximate solution is given by

$$\begin{aligned} x_1 &= (6 + 0 - 0)/10 = 0.6, \\ x_2 &= (25 - 0 - 0)/11 = 2.2727, \\ x_3 &= (-11 - 0 - 0)/10 = -1.1, \\ x_4 &= (15 - 0 - 0)/8 = 1.875. \end{aligned}$$

Using the approximations obtained, the iterative procedure is repeated until the desired accuracy has been reached. The following are the approximated solutions after five iterations.

x_1	x_2	x_3	x_4
0.6	2.27272	-1.1	1.875
1.04727	1.7159	-0.80522	0.88522
0.93263	2.05330	-1.0493	1.13088
1.01519	1.95369	-0.9681	0.97384
0.98899	2.0114	-1.0102	1.02135

The exact solution of the system is $(1, 2, -1, 1)$.

1.3.3 Gauss-Seidel Method

The Gauss-Seidel iteration method can be defined as:

$$\begin{cases} x_1^{k+1} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^k + a_{13}x_3^k + \cdots + a_{1n}x_n^k) \\ x_2^{k+1} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^{k+1} + a_{23}x_3^k + \cdots + a_{2n}x_n^k) \\ x_3^{k+1} = \frac{1}{a_{33}}(b_3 - a_{31}x_1^{k+1} + a_{32}x_2^{k+1} + \cdots + a_{3n}x_n^k) \\ \vdots \\ x_n^{k+1} = \frac{1}{a_{nn}}(b_n - a_{n1}x_1^{k+1} + a_{n2}x_2^{k+1} + \cdots + a_{nn-1}x_{n-1}^{k+1}) \end{cases} \quad (1.7)$$

The Gauss-Seidel method is an iterative technique for solving a square system of n linear equations with unknown x :

$$A\mathbf{x} = \mathbf{b}.$$

It is defined by the iteration

$$L_*\mathbf{x}^{(k+1)} = \mathbf{b} - U\mathbf{x}^{(k)},$$

where $\mathbf{x}^{(k)}$ is the k th approximation or iteration of $\mathbf{x}$, $\mathbf{x}^{(k+1)}$ is the next or $k + 1$ iteration of $\mathbf{x}$, and the matrix A is decomposed into a lower triangular component L_* , and a strictly upper triangular component U : $A = L_* + U$



In more detail, write out A , x and b in their components:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Then the decomposition of A into its lower triangular component and its strictly upper triangular component is given by:

$$A = L_* + U \quad \text{where} \quad L_* = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad U = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

The system of linear equations may be rewritten as:

$$L_* \mathbf{x} = \mathbf{b} - U \mathbf{x}$$

The Gauss-Seidel method now solves the left hand side of this expression for x , using previous value for x on the right hand side. Analytically, this may be written as:

$$\mathbf{x}^{(k+1)} = L_*^{-1}(\mathbf{b} - U \mathbf{x}^{(k)}).$$

However, by taking advantage of the triangular form of L_* , the elements of $x^{(k+1)}$ can be computed sequentially using forward substitution:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad i = 1, 2, \dots, n.$$

The procedure is generally continued until the changes made by an iteration are below some tolerance, such as a sufficiently small residual.

The element-wise formula for the Gauss-Seidel method is extremely similar to that of the Jacobi method. The computation of $x_i^{(k+1)}$ uses only the elements of $x^{(k+1)}$ that have already been computed, and only the elements of $x^{(k)}$ that have not yet to be advanced to iteration $k + 1$. This means that, unlike the Jacobi method, only one storage vector is required as elements can be overwritten as they are computed, which can be advantageous for very large problems. However, unlike the Jacobi method, the computations for each element cannot be done in parallel. Furthermore, the values at each iteration are dependent on the order of the original equations.

Convergence The convergence properties of the Gauss-Seidel method are dependent on the matrix A . Namely, the procedure is known to converge if either:

- A is symmetric positive-definite, or
- A is strictly or irreducibly diagonally dominant.

The Gauss-Seidel method sometimes converges even if these conditions are not satisfied.

Algorithm Since elements can be overwritten as they are computed in this algorithm, only one storage vector is needed, and vector indexing is omitted. The algorithm goes as follows:



Example 1.10: A

linear system shown as $A\mathbf{x} = \mathbf{b}$ is given by: $A = \begin{bmatrix} 16 & 3 \\ 7 & -11 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 11 \\ 13 \end{bmatrix}$. We want to use the equation

$$\mathbf{x}^{(k+1)} = L_*^{-1}(\mathbf{b} - U\mathbf{x}^{(k)})$$

in the form

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + C$$

where: $T = -L_*^{-1}U$ and $C = L_*^{-1}\mathbf{b}$. We must decompose A into the sum of a lower triangular component L_* and a strict upper triangular component U :

$L_* = \begin{bmatrix} 16 & 0 \\ 7 & -11 \end{bmatrix}$ and $U = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$. The inverse of L_* is:

$$L_*^{-1} = \begin{bmatrix} 16 & 0 \\ 7 & -11 \end{bmatrix}^{-1} = \begin{bmatrix} 0.0625 & 0.0000 \\ 0.0398 & -0.0909 \end{bmatrix}$$

Now we can find:

$$T = -\begin{bmatrix} 0.0625 & 0.0000 \\ 0.0398 & -0.0909 \end{bmatrix} \times \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.0625 & 0.0000 \\ 0.0398 & -0.0909 \end{bmatrix} \times \begin{bmatrix} 11 \\ 13 \end{bmatrix} = \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix}.$$

Now we have T and C and we can use them to obtain the vectors $\mathbf{x}$ iteratively. First of all, we have to choose $\mathbf{x}^{(0)}$: we can only guess. The better the guess, the quicker the algorithm will perform.

We suppose:

$$\mathbf{x}^{(0)} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}.$$

We can then calculate:

$$\begin{aligned} \mathbf{x}^{(1)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.5000 \\ -0.8636 \end{bmatrix}. \\ \mathbf{x}^{(2)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 0.5000 \\ -0.8636 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.8494 \\ -0.6413 \end{bmatrix}. \\ \mathbf{x}^{(3)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 0.8494 \\ -0.6413 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.8077 \\ -0.6678 \end{bmatrix}. \\ \mathbf{x}^{(4)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 0.8077 \\ -0.6678 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.8127 \\ -0.6646 \end{bmatrix}. \\ \mathbf{x}^{(5)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 0.8127 \\ -0.6646 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.8121 \\ -0.6650 \end{bmatrix}. \\ \mathbf{x}^{(6)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 0.8121 \\ -0.6650 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.8122 \\ -0.6650 \end{bmatrix}. \\ \mathbf{x}^{(7)} &= \begin{bmatrix} 0.000 & -0.1875 \\ 0.000 & -0.1193 \end{bmatrix} \times \begin{bmatrix} 0.8122 \\ -0.6650 \end{bmatrix} + \begin{bmatrix} 0.6875 \\ -0.7443 \end{bmatrix} = \begin{bmatrix} 0.8122 \\ -0.6650 \end{bmatrix}. \end{aligned}$$

As expected, the algorithm converges to the exact solution:

$$\mathbf{x} = A^{-1}\mathbf{b} \approx \begin{bmatrix} 0.8122 \\ -0.6650 \end{bmatrix}.$$

Example 1.11

Another linear system shown as $A\mathbf{x} = \mathbf{b}$ is given by:

$A = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$ and $b = \begin{bmatrix} 11 \\ 13 \end{bmatrix}$. We want to use the equation

$$\mathbf{x}^{(k+1)} = L_*^{-1}(\mathbf{b} - U\mathbf{x}^{(k)})$$

in the form

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + C$$

where: $T = -L_*^{-1}U$ and $C = L_*^{-1}\mathbf{b}$. We must decompose A into the sum of a lower triangular component L_* and a strict upper triangular component U : $L_* = \begin{bmatrix} 2 & 0 \\ 5 & 7 \end{bmatrix}$ and

$U = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$. is:

$$L_*^{-1} = \begin{bmatrix} 2 & 0 \\ 5 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} 0.500 & 0.000 \\ -0.357 & 0.143 \end{bmatrix}$$

Now we can find:

$$T = -\begin{bmatrix} 0.500 & 0.000 \\ -0.357 & 0.143 \end{bmatrix} \times \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.000 & -1.500 \\ 0.000 & 1.071 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.500 & 0.000 \\ -0.357 & 0.143 \end{bmatrix} \times \begin{bmatrix} 11 \\ 13 \end{bmatrix} = \begin{bmatrix} 5.500 \\ -2.071 \end{bmatrix}.$$

Now we have T and C and we can use them to obtain the vectors $\mathbf{x}$ iteratively.

First of all, we have to choose $\mathbf{x}^{(0)}$: we can only guess. The better the guess, the quicker will perform the algorithm.

We suppose:

$$x^{(0)} = \begin{bmatrix} 1.1 \\ 2.3 \end{bmatrix}.$$

x We can then calculate:

$$x^{(1)} = \begin{bmatrix} 0 & -1.500 \\ 0 & 1.071 \end{bmatrix} \times \begin{bmatrix} 1.1 \\ 2.3 \end{bmatrix} + \begin{bmatrix} 5.500 \\ -2.071 \end{bmatrix} = \begin{bmatrix} 2.050 \\ 0.393 \end{bmatrix}.$$

$$x^{(2)} = \begin{bmatrix} 0 & -1.500 \\ 0 & 1.071 \end{bmatrix} \times \begin{bmatrix} 2.050 \\ 0.393 \end{bmatrix} + \begin{bmatrix} 5.500 \\ -2.071 \end{bmatrix} = \begin{bmatrix} 4.911 \\ -1.651 \end{bmatrix}.$$

$$x^{(3)} = \dots$$

If we test for convergence we'll find that the algorithm diverges. In fact, the matrix A is neither diagonally dominant nor positive definite. Then, convergence to the exact solution

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -38 \\ 29 \end{bmatrix}$$

is not guaranteed and, in this case, will not occur.



Example 1.12

Suppose given k equations where x_n are vectors of these equations and starting point x_0 . From the first equation solve for x_1 in terms of $x_{n+1}, x_{n+2}, \dots, x_n$. For the next equations substitute the previous values of x s.

To make it clear let's consider an example.

$$\begin{aligned} 10x_1 - x_2 + 2x_3 &= 6, \\ -x_1 + 11x_2 - x_3 + 3x_4 &= 25, \\ 2x_1 - x_2 + 10x_3 - x_4 &= -11, \\ 3x_2 - x_3 + 8x_4 &= 15. \end{aligned}$$

Solving for x_1, x_2, x_3 and x_4 gives:

$$\begin{aligned} x_1 &= x_2/10 - x_3/5 + 3/5, \\ x_2 &= x_1/11 + x_3/11 - 3x_4/11 + 25/11, \\ x_3 &= -x_1/5 + x_2/10 + x_4/10 - 11/10, \\ x_4 &= -3x_2/8 + x_3/8 + 15/8. \end{aligned}$$

Suppose we choose $(0, 0, 0, 0)$ as the initial approximation, then the first approximate solution is given by

$$\begin{aligned} x_1 &= 3/5 = 0.6, \\ x_2 &= (3/5)/11 + 25/11 = 3/55 + 25/11 = 2.3272, \\ x_3 &= -(3/5)/5 + (2.3272)/10 - 11/10 = -3/25 + 0.23272 - 1.1 = -0.9873, \\ x_4 &= -3(2.3272)/8 + (-0.9873)/8 + 15/8 = 0.8789. \end{aligned}$$

Using the approximations obtained, the iterative procedure is repeated until the desired accuracy has been reached. The following are the approximated solutions after four iterations.

x_1	x_2	x_3	x_4
0.6	2.32727	-0.987273	0.878864
1.03018	2.03694	-1.01446	0.984341
1.00659	2.00356	-1.00253	0.998351
1.00086	2.0003	-1.00031	0.99985

The exact solution of the system is $(1, 2, -1, 1)$.

1.4 Eigenvalue Problem

Suppose that A is a square $(n \times n)$ matrix. We say that a nonzero vector v is an eigenvector and a number λ is its eigenvalue if

$$Av = \lambda v \quad (1.8)$$

Geometrically this means that Av is in the same direction as v , since multiplying a vector by a number changes its length, but not its direction.



1.4.1 Finding Eigenvalues for 2×2 and 3×3

If A is 2×2 or 3×3 then we can find its eigenvalues and eigenvectors by hand. Notice that Equation (1.8) can be rewritten as

$$Av - \lambda v = 0.$$

It would be nice to factor out the v from the right-hand side of this equation, but we can't because A is a matrix and λ is a number. However, since $Iv = v$, we can do the following:

$$\begin{aligned} Av - \lambda v &= Av - \lambda Iv \\ &= (A - \lambda I)v \\ &= 0 \end{aligned}$$

If v is nonzero, then the matrix $(A - \lambda I)$ must be singular. By the same theorem, we must have

$$\det(A - \lambda I) = 0.$$

This is called the **characteristic equation**.



Example 1.13

Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Taking the determinant to find characteristic polynomial of A ,

$$\begin{aligned} |A - \lambda I| &= \left| \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix}, \\ &= 3 - 4\lambda + \lambda^2. \end{aligned}$$

Setting the characteristic polynomial equal to zero, it has roots at $\lambda = 1$ and $\lambda = 3$, which are the two eigenvalues of A .

For $\lambda = 1$, the CXC equation becomes,

$$(A - I)v_{\lambda=1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Any non-zero vector with $v_1 = -v_2$ solves this equation. Therefore,

$$v_{\lambda=1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

is an eigenvector of A corresponding to $\lambda = 1$, as is any scalar multiple of this vector.

For $\lambda = 3$, CXC Equation becomes

$$(A - 3I)v_{\lambda=3} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Any non-zero vector with $v_1 = v_2$ solves this equation. Therefore,

$$v_{\lambda=3} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is an eigenvector of A corresponding to $\lambda = 3$, as is any scalar multiple of this vector.

Thus, the vectors $v_{\lambda=1}$ and $v_{\lambda=3}$ are eigenvectors of A associated with the eigenvalues $\lambda = 1$ and $\lambda = 3$, respectively.

As mentioned above, the eigenvalues and eigenvectors of an $n \times n$ matrix where $n \geq 4$ must be found numerically instead of by hand. The numerical methods that are used in practice depend on the geometric meaning of eigenvalues and eigenvectors which is equation (1.8). The essence of all these methods is captured in the Power method, which we now introduce.

Definition 1.7: L

t $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ be the eigenvalues of an matrix A . λ_1 is called the dominant eigenvalue of A if $|\lambda_1| > |\lambda_i|$ for $i = 2, 3, 4, \dots, n$. The eigenvectors corresponding to λ_1 are called dominant eigenvectors of A .



1.4.2 Power Method

Like the Jacobi and Gauss-Seidel methods, the power method for approximating eigenvalues is iterative.

Theorem 1.2

If A is an diagonalizable matrix with a dominant eigenvalue, then there exists a nonzero vector x_0 such that the sequence of vectors given by

$$Ax_0, A^2x_0, A^3x_0, A^4x_0, A^5x_0, A^6x_0, \dots, A^nx_0, \dots$$

approaches a multiple of the dominant eigenvector of A .

Proof Because A is diagonalizable, you know that it has n linearly independent eigenvectors $x_1, x_2, x_3, \dots, x_n$ with corresponding eigenvalues of $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$. Assume that these eigenvalues are ordered so that λ_1 is the dominant eigenvalue (with a corresponding eigenvector of x_1). Because the n eigenvectors $x_1, x_2, x_3, \dots, x_n$ are linearly independent, they must form a basis for R^n . For the initial approximation x_0 choose a nonzero vector such that the linear combination

$$x_0 = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

has nonzero leading coefficients. (If $c_1 = 0$ the power method may not converge, and a different x_0 must be used as the initial approximation.) Now, multiplying both sides of this equation by A produces

$$Ax_0 = A(c_1x_1 + c_2x_2 + \dots + c_nx_n)$$

$$Ax_0 = c_1(Ax_1) + c_2(Ax_2) + \dots + c_n(Ax_n)$$

$$Ax_0 = c_1(\lambda_1x_1) + c_2(\lambda_2x_2) + \dots + c_n(\lambda_nx_n)$$

Repeated multiplication of both sides of this equation by A produces

$$A^kx_0 = c_1(\lambda_1^kx_1) + c_2(\lambda_2^kx_2) + \dots + c_n(\lambda_n^kx_n)$$

which implies that

$$A^kx_0 = \lambda_1^k[c_1x_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^kx_2 + \dots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^kx_n]$$

Now, from the original assumption that λ_1 is larger in absolute value than the other eigenvalues it follows that each of the fractions

$$\frac{\lambda_2}{\lambda_1}, \frac{\lambda_3}{\lambda_1}, \dots, \frac{\lambda_n}{\lambda_1}$$

is less than 1 in absolute value. So each of the factors

$$\left(\frac{\lambda_2}{\lambda_1}\right)^k, \left(\frac{\lambda_3}{\lambda_1}\right)^k, \dots, \left(\frac{\lambda_n}{\lambda_1}\right)^k$$

must approach 0 as k approaches infinity. This implies that the approximation

$$A^kx_0 \approx c_1\lambda_1x_1$$

improves as k increases. Because x_1 is a dominant eigenvector, it follows that any scalar multiple of x_1 is also a dominant eigenvector, so showing that A^kx_0 approaches a multiple of the dominant eigenvector of A .

Note The power method will converge quickly if $\frac{\lambda_i}{\lambda_1}$, $i = 2, 3, \dots, n$ is small, and slowly if



$\frac{\lambda_i}{\lambda_1}$, $i = 2, 3, \dots, n$ is close to 1.

Hence first assume that the matrix A has a dominant eigenvalue with corresponding dominant eigenvectors. Then choose an initial approximation x_0 of one of the dominant eigenvectors of A . This initial approximation must be a nonzero vector in R^n . Finally, form the sequence given by

$$\begin{aligned} x_1 &= Ax_0 \\ x_2 &= Ax_1 = A(Ax_0) = A^2x_0 \\ x_3 &= Ax_2 = A(A^2x_0) = A^3x_0 \\ &\vdots \\ &\vdots \\ &\vdots \\ x_k &= Ax_{k-1} = A(A^{k-1}x_0) = A^kx_0 \end{aligned}$$

For large powers of k , and by properly scaling this sequence, you will see that you obtain a good approximation of the dominant eigenvector of A . This procedure is illustrated in the following Example

Example 1.14

Approximating a Dominant Eigenvector by the Power Method Complete six iterations of the power method to approximate a dominant eigenvector of

$$\begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix}$$

by the Power Method

Solution

Begin with an initial nonzero approximation of

$$x_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Then obtain the following approximations.

$$x_1 = Ax_0 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \\ 2 \end{bmatrix} \Rightarrow 7 \begin{bmatrix} 0.5714 \\ 1 \\ 0.2857 \end{bmatrix}$$

$$x_2 = Ax_1 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.5714 \\ 1 \\ 0.2857 \end{bmatrix} = \begin{bmatrix} 3.7143 \\ 7.1429 \\ 4 \end{bmatrix} \Rightarrow 7.1429 \begin{bmatrix} 0.52 \\ 1.00 \\ 0.56 \end{bmatrix}$$

$$x_3 = Ax_2 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.52 \\ 1.00 \\ 0.56 \end{bmatrix} = \begin{bmatrix} 2.96 \\ 7.52 \\ 2.8 \end{bmatrix} \Rightarrow 7.52 \begin{bmatrix} 0.3936 \\ 1.000 \\ 0.3723 \end{bmatrix}$$

$$x_4 = Ax_3 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.3936 \\ 1.000 \\ 0.3723 \end{bmatrix} = \begin{bmatrix} 2.8298 \\ 7.5851 \\ 2.2979 \end{bmatrix} \Rightarrow 7.5851 \begin{bmatrix} 0.3731 \\ 1.00 \\ 0.4348 \end{bmatrix}$$



$$x_5 = Ax_4 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.3731 \\ 1.00 \\ 0.4348 \end{bmatrix} = \begin{bmatrix} 2.6227 \\ 7.6886 \\ 3.0070 \end{bmatrix} \Rightarrow 7.6886 \begin{bmatrix} 0.3411 \\ 1.00 \\ 0.3911 \end{bmatrix}$$

$$x_6 = Ax_5 = \begin{bmatrix} 4 & 2 & -2 \\ -2 & 8 & 1 \\ 2 & 4 & -4 \end{bmatrix} \begin{bmatrix} 0.3411 \\ 1.00 \\ 0.3911 \end{bmatrix} = \begin{bmatrix} 2.5197 \\ 7.7401 \\ 3.0760 \end{bmatrix} \Rightarrow 7.7401 \begin{bmatrix} 0.3255 \\ 1.00 \\ 0.3974 \end{bmatrix}$$

The results show that the differences between the vector $[x_i]$ and the normalized vector $[x_{i+1}]$ are getting smaller. The value of the multiplicative factor (7.7401) is an estimate of the largest eigenvalue.

Theorem 1.3

Determining an Eigenvalue from an Eigenvector If x is an eigenvector of a matrix A , then its corresponding eigenvalue is given by

$$\lambda = \frac{Ax * x^t}{x * x^t}$$

This quotient is called the Rayleigh quotient

Proof Because x is an eigenvector of A , you know that $Ax = \lambda x$ and can write

$$\frac{(Ax) * x^t}{x * x^t} = \frac{(\lambda x) * x^t}{x * x^t} = \lambda \frac{x * x^t}{x * x^t} = \lambda$$

In cases for which the power method generates a good approximation of a dominant eigenvector, the Rayleigh quotient provides a correspondingly good approximation of the dominant eigenvalue

Example 1.15

The Power Method with Scaling Calculate seven iterations of the power method with scaling to approximate a dominant eigenvector of the matrix

$$\begin{bmatrix} 1 & 2 & 0 \\ -2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix}$$

Use $x_0 = (1, 1, 1)^t$ as the initial approximation.

Solution:

One iteration of the power method produces

$$Ax_0 = \begin{bmatrix} 1 & 2 & 0 \\ -2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} = 5 \begin{bmatrix} 0.60 \\ 0.20 \\ 1.00 \end{bmatrix}$$

and by scaling you obtain the approximation

$$x_1 = \frac{1}{5} \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0.60 \\ 0.20 \\ 1.00 \end{bmatrix}$$



A second iteration yields

$$Ax_1 = \begin{bmatrix} 1 & 2 & 0 \\ -2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0.60 \\ 0.20 \\ 1.00 \end{bmatrix} = \begin{bmatrix} 1.00 \\ 1.00 \\ 2.20 \end{bmatrix} = 2.2 \begin{bmatrix} 0.45 \\ 0.45 \\ 1.00 \end{bmatrix}$$

and

$$x_2 = \frac{1}{2.2} \begin{bmatrix} 0.45 \\ 0.45 \\ 1.00 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.45 \\ 1.00 \end{bmatrix}$$

Continuing this process, you obtain the sequence of approximations shown in the following Table

x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.6 \\ 0.2 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.45 \\ 0.45 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.48 \\ 0.55 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.51 \\ 0.51 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 0.49 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 0.50 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 0.50 \\ 1 \end{bmatrix}$

From the Table above you can approximate a dominant eigenvector of A to be $\begin{bmatrix} 0.50 \\ 0.50 \\ 1 \end{bmatrix}$ Using the Rayleigh quotient, you can approximate the dominant eigenvalue of A to be $\lambda = 3$ (For this example you can check that the approximations of x and λ are exact.)

x_1	x_2	x_3	x_4	x_5	x_6	x_7
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
5.00	2.20	2.82	3.13	3.02	2.99	3.00

are approaching the dominant eigenvalue $\lambda = 3$

1.4.3 Inverse Power Method

Inverse power method can give approximation to any eigenvalue. However, it is used usually to find the smallest eigenvalue in magnitude and the corresponding eigenvector of a given matrix A. The eigenvectors are computed very accurately by this method. Further, the method is powerful to calculate accurately the eigenvectors, when the eigenvalues are not well separated. In this case, power method converges very slowly.

If λ is an eigenvalue of A, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} corresponding to the same eigenvector. The smallest eigenvalue λ in magnitude of A is the largest eigenvalue $\frac{1}{\lambda}$ in magnitude of A^{-1} . Then choose an initial approximation x_0 of one of the dominant eigenvectors of A^{-1} . This initial approximation must be a nonzero vector in R^n . Finally, Applying the power method on A^{-1} , we have

$$\begin{aligned} x_1 &= A^{-1}x_0 \\ x_2 &= A^{-1}x_1 = A^{-1}(A^{-1}x_0) = (A^{-1})^2x_0 \\ x_3 &= A^{-1}x_2 = A^{-1}((A^{-1})^2x_0) = (A^{-1})^3x_0 \\ &\vdots \\ x_k &= A^{-1}x_{k-1} = A^{-1}((A^{-1})^{k-1}x_0) = (A^{-1})^kx_0 \end{aligned}$$

For large powers of k , and by properly scaling this sequence, you will see that you obtain a good approximation of the dominant eigenvector of A. This procedure is illustrated in the following.



Then using Rayleigh quotient we can find the dominant eigenvalue of A^{-1}

$$\frac{1}{\lambda} = \frac{A^{-1}x * x^t}{x * x^t}$$

Example 1.16

Find the smallest eigenvalue in magnitude of the matrix

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

use four iteration of the inverse power method.

Solution:

The smallest eigenvalue in magnitude of A is the largest eigenvalue in magnitude of A^{-1} . We have

$$A^{-1} = \frac{1}{4} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Then use $x_0 = (1, 1, 1)^t$ and apply inverse power method with scaling.

First approximation

$$A^{-1}x_0 = A^{-1} = \frac{1}{4} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 2 \\ 1.5 \end{bmatrix} \Rightarrow x_1 = \begin{bmatrix} 1 \\ 1.333 \\ 1 \end{bmatrix}$$

Second Approximation

$$A^{-1}x_1 = \frac{1}{4} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1.333 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.6667 \\ 2.3333 \\ 1.6667 \end{bmatrix} \Rightarrow x_2 = \begin{bmatrix} 1.0000 \\ 1.4000 \\ 1.0000 \end{bmatrix}.$$

Third approximation

$$A^{-1}x_2 = \frac{1}{4} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1.0000 \\ 1.4000 \\ 1.0000 \end{bmatrix} = \begin{bmatrix} 1.7000 \\ 2.4000 \\ 1.7000 \end{bmatrix} \Rightarrow x_3 = \begin{bmatrix} 1.0000 \\ 1.4118 \\ 1.0000 \end{bmatrix}.$$

Fourth approximation

$$A^{-1}x_3 = \frac{1}{4} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1.0000 \\ 1.4118 \\ 1.0000 \end{bmatrix} = \begin{bmatrix} 1.7059 \\ 2.4118 \\ 1.7059 \end{bmatrix} \Rightarrow x_4 = \begin{bmatrix} 1.0000 \\ 1.4138 \\ 1.0000 \end{bmatrix}.$$

From the above we can approximate a dominant eigenvector of A^{-1} to be $\begin{bmatrix} 1.0000 \\ 1.4138 \\ 1.0000 \end{bmatrix}$. After four iteration using the Rayleigh quotient, you can approximate the dominant eigenvalue of A^{-1} is

$$\frac{1}{\lambda} = \frac{\frac{1}{4} \left(\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1.0000 \\ 1.4138 \\ 1.0000 \end{bmatrix} \right)^t \begin{bmatrix} 1.0000 \\ 1.4138 \\ 1.0000 \end{bmatrix}}{\left(\begin{bmatrix} 1.0000 & 1.4138 & 1.0000 \end{bmatrix} \right) \begin{bmatrix} 1.0000 \\ 1.4138 \\ 1.0000 \end{bmatrix}} = 1.7071.$$



Therefore $\lambda = 0.5858$ is required eigenvalue. The corresponding eigenvector is $\begin{bmatrix} 1.0000 & 1.4138 & 1.0000 \end{bmatrix}^t$. The smallest eigenvalue of A is $2 - \sqrt{2} = 0.5858$.

1.5 System of Non-linear Equations

Recall that at the end of Chap. 2 we presented an approach to solve two nonlinear equations with one unknowns. This approach can be extended to the general case of solving n simultaneous nonlinear equations.

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0 \\ f_2(x_1, x_2, \dots, x_n) &= 0 \\ f_3(x_1, x_2, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, x_2, \dots, x_n) &= 0 \end{aligned}$$

The solution of this system consists of the set of x values that simultaneously result in all the equations equaling zero.

1.5.1 Newton Raphson method

One approach to solving such systems is based on a multidimensional version of the Newton-Raphson method. Thus, a Taylor series expansion is written for each equation about the point $(x_1^k, x_2^k, \dots, x_n^k)$ we get,

$$\begin{aligned} f_1(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) &= f_1(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) \\ &+ \left[\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right] f_1(x_1^k, x_2^k, \dots, x_n^k) + \\ &\frac{1}{2!} \left[\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right]^2 f_1(x_1^k, x_2^k, \dots, x_n^k) + \dots = 0 \\ f_2(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) &= f_2(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) \\ &+ \left[\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right] f_2(x_1^k, x_2^k, \dots, x_n^k) + \\ &\frac{1}{2!} \left[\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right]^2 f_2(x_1^k, x_2^k, \dots, x_n^k) + \dots = 0 \\ &\dots \\ &\dots \\ f_n(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) &= f_n(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) \\ &+ \left[\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right] f_n(x_1^k, x_2^k, \dots, x_n^k) + \\ &\frac{1}{2!} \left[\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n} \right]^2 f_n(x_1^k, x_2^k, \dots, x_n^k) + \dots = 0 \end{aligned} \quad (1.9)$$



Neglecting 2^{nd} and higher powers of $\Delta x_1, \Delta x_2, \dots$ and Δx_n , we obtain

$$\begin{aligned}
 f_1(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) &= f_1(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) \\
 &+ [\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n}] f_1(x_1^k, x_2^k, \dots, x_n^k) \\
 f_2(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) &= f_2(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) \\
 &+ [\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n}] f_2(x_1^k, x_2^k, \dots, x_n^k) \\
 &\dots \\
 &\dots \\
 f_n(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) &= f_n(x_1^k + \Delta x_1, x_2^k + \Delta x_2, \dots, x_n^k + \Delta x_n) \\
 &+ [\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \dots + \Delta x_n \frac{\partial}{\partial x_n}] f_n(x_1^k, x_2^k, \dots, x_n^k)
 \end{aligned} \tag{1.10}$$

Since $x_1^{k+1} = x_1^k + \Delta x_1, x_2^{k+1} = x_2^k + \Delta x_2, \dots$ and $x_n^{k+1} = x_n^k + \Delta x_n$ writing the equation in matrix form, we get

$$J_k \Delta X^k = -F(X^k) \tag{1.11}$$

$$\begin{aligned}
 \text{where } J_k &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad \text{at } (x_1^k, x_2^k, \dots, x_n^k) \quad , \Delta X^k = \begin{bmatrix} x_1^{k+1} - x_1^k \\ x_2^{k+1} - x_2^k \\ \dots \\ x_n^{k+1} - x_n^k \end{bmatrix} \quad \text{and} \\
 F(X^k) &= \begin{bmatrix} f_1(x_1^k, x_2^k, x_3^k, \dots, x_n^k) \\ f_2(x_1^k, x_2^k, x_3^k, \dots, x_n^k) \\ \dots \\ f_n(x_1^k, x_2^k, x_3^k, \dots, x_n^k) \end{bmatrix} \quad \text{Therefore equation can be written as } \Delta X^k = -J_k^{-1} F(X^k) \\
 [X^{k+1}] &= [X^k] - J_k^{-1} F(X^k) \quad , k = 0, 1, 2, 3, \dots \tag{1.12}
 \end{aligned}$$

The convergence of the method depends on the initial approximation X_0 . A sufficient condition for convergence is that for each k

$$\|J_k^{-1}\| < 1.$$

whereas a necessary and sufficient condition for convergence is

$$\rho(J_k^{-1}) < 1$$

Where $\|\cdot\|$ is suitable norm and $\rho(J_k^{-1})$ is the spectral radius (large eigenvalue in magnitude) of the matrix J_k^{-1}

if the method converges, then its rate of convergence is two. The iterations stopped when

$$\|X^{k+1} - X^k\| < \epsilon$$

Where ϵ is the given error tolerance.



Example 1.17

perform three iterations of the Newton-Raphson Method to solve the system of equations

$$\begin{aligned}x^2 + xy + y^2 &= 7 \\x^3 + y^3 &= 9\end{aligned}$$

Take initial approximation as $x_0 = 1.5$ and $y_0 = 0.5$. The exact solution is $x = 2, y = 1$

Solution: We have

$$\begin{aligned}f(x) &= x^2 + xy + y^2 - 7 = 0 \\g(x) &= x^3 + y^3 - 9 = 0\end{aligned}$$

$$J_k = \begin{bmatrix} f_x(x_k, y_k) & f_y(x_k, y_k) \\ g_x(x_k, y_k) & g_y(x_k, y_k) \end{bmatrix} = \begin{bmatrix} 2x_k + y_k & x_k + 2y_k \\ 3x_k^2 & 3y_k^2 \end{bmatrix}$$

$$J_k^{-1} = \frac{1}{D_k} \begin{bmatrix} 3y_k^2 & -(x_k + 2y_k) \\ -3x_k^2 & 2x_k + y_k \end{bmatrix}$$

Where $D_k = |J_k| = 3y_k^2(2x_k + y_k) - 3x_k^2(x_k + 2y_k)$. Know we can write the method as

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \frac{1}{D_k} \begin{bmatrix} 3y_k^2 & -(x_k + 2y_k) \\ -3x_k^2 & 2x_k + y_k \end{bmatrix} \begin{bmatrix} x_k^2 + x_k y_k + y_k^2 - 7 \\ x_k^3 + y_k^3 - 9 \end{bmatrix} \quad k = 0, 1, 2, 3, \dots$$

Using $(x_0, y_0) = (1.5, 0.5)$, we get

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} - \frac{1}{-14.25} \begin{bmatrix} 0.75 & -2.5 \\ -6.75 & 3.5 \end{bmatrix} \begin{bmatrix} -3.75 \\ -5.5 \end{bmatrix} = \begin{bmatrix} 2.2675 \\ 0.9254 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2.2675 \\ 0.9254 \end{bmatrix} - \frac{1}{-49.4951} \begin{bmatrix} 2.5691 & -4.1183 \\ -15.4247 & 5.4604 \end{bmatrix} \begin{bmatrix} 1.0963 \\ 3.4510 \end{bmatrix} = \begin{bmatrix} 2.0373 \\ 0.9645 \end{bmatrix}$$

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2.0373 \\ 0.9645 \end{bmatrix} - \frac{1}{-35.3244} \begin{bmatrix} 2.7908 & -3.9663 \\ -12.4518 & 5.0391 \end{bmatrix} \begin{bmatrix} 0.0458 \\ 0.3532 \end{bmatrix} = \begin{bmatrix} 2.0013 \\ 0.9987 \end{bmatrix}$$

