3

System of Linear Equations

3.1 Introduction

A completely general system of m linear equations in n unknowns is of the following form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ &. \\ &. \\ &. \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m \end{aligned}$$

The a_{1j} is the coefficient of x_j in the ith equation. The data for this system of equations are all the numbers a_{ij} and b_i. Now consider the four matrices.

$$A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ . & & & \\ . & & & \\ . & & & \\ a_{m1} & a_{m2} & ... & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{bmatrix}, \qquad [A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & b_2 \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{m1} & a_{m2} & ... & a_{mn} & b_n \end{array}\right]$$

In this context of a system of equations; A is called the coefficient matrix, x is called vector of unknowns, b is called righthandside vector, $[A|b]$ is called the augmented matrix. So, the **matrix notation** of the system of linear equations is $Ax = b$.

3.2 LU Factorization Method

Theorem 1. Let $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & & & \\ . & & & \\ . & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ be a singular matrix. Then if A is a non-singular matrix then A can be factorized as $A = LU$, where

- L is a lower triangular matrix, and
- U is an upper triangular matrix

It is the standard result of linear algebra that such a factorization, when exists, is *unique*.

Procedure for finding L and U

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$ Multiplying the matrices on the right side and equating the coefficients, we get

$$\begin{aligned} u_{11} &= a_{11}, & u_{12} &= a_{12}, & u_{13} &= a_{13} \\ l_{21}u_{11} &= a_{21}, & l_{21}u_{22} + u_{22} &= a_{22}, & l_{21}u_{23} + u_{23} &= a_{23} \\ l_{31}u_{11} &= a_{31}, & l_{31}u_{22} + l_{32}u_{22} &= a_{32}, & l_{31}u_{23} + l_{32}u_{33} &= a_{33} \end{aligned}$$

From these equations we obtain

$$l_{21} = \frac{a_{21}}{a_{11}}, \qquad l_{31} = \frac{a_{31}}{a_{11}}, \qquad u_{22} = a_{22} - a_{12}\frac{a_{21}}{a_{11}}, \dots \tag{3.1}$$

This procedure is a systematic one. First we determine first row of U and first column of L, then we determine second row of U and second column of L, and finally third row of U.

3.2.1 Method

Let $Ax = b$ be a system of linear equations. Let $A = LU$ be the LU factorization of A. Then we have,

$$Ax = b \implies LUx = B \implies Ly = b \tag{3.2}$$

where, $y = Ux$.
Since L is an lower triangular matrix y can be solved conveniently by forward substituion in 3.2. Then x can be solved using backward substituion in $Ux = y$, as U is an upper triangular matrix.

3.2.2 Tridiagonal Systems

The matrix of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & ... & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & ... & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & ... & 0 & 0 & 0 & 0 \\ ... \\ 0 & 0 & 0 & 0 & ... & 0 & a_{(n-1)(n-2)} & a_{(n-1)(n-1)} & a_{(n-1)n} \\ 0 & 0 & 0 & 0 & ... & 0 & 0 & a_{n(n-1)} & a_{nn} \end{bmatrix}$$

is called tridiagonal matrix. It is a square matrix with nonzero elements only on the main diagonal, subdiagonal, superdiagonal. The method of LU factorization can conveniently apllied to the system having a tridiagonal matrix as the coefficient matirx.

3.2.3 Exercise

Solve the system:

1. $2x - y = 0, \quad -x + 2y - z = 0, \quad -y + 2z - u = 0, \quad -z + 2u = 1.$
2. $3x_1 - x_2 = -1, \quad -x_1 + 3x_2 - x_3 = 7, \quad -x_2 + 3x_3 = 7$

3.3 Vector and Matrix Norms

Vector Norm

The distance between a vector and the null vector is a measure of the *length* of the vector. This is called a norm of the vector. Mathematically, a norm of a vector in a vector space V is defined as a function $\|.\| : V \mapsto \mathbb{R}$ that satisfies the following conditions:

i) $\|x\| \geq 0$

ii) $\|x\| = 0$ if and only if $x = 0$.

iii) $\alpha\|x\| = |\alpha|\|x\|$ for any real α.

iv) $\|x + y\| \leq \|x\| + \|y\|$.

There are different types of length of a vector. These gives rise to different types of norms. Some useful norms are as follows:

1. $L1$ norm also called
   Rectangular norm or Manhattan norm: $\|x\|_1 = |x_1| + |x_2| + ...|x_n|$.
2. $L2$ norm also called **Euclidean norm**: $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + ...|x_n|^2}$.
3. $L\infty$ norm also called
   Maximum norm or Uniform norm: $\|x\|_\infty = max\{|x_1|, |x_2|, ..., |x_n|\}$.

Matrix Norm

Corresponding to the vector norms we have the three matrix norms defned as follows:

1. **Column Norm**: $\|A\|_1 = \max_j \sum_i |a_{ij}|$.

2. **Euclidean Norm**: $\|A\|_e = \left[\sum_{i,j} |a_{ij}|^2\right]^{1/2}$.

3. **Column Norm**: $\|A\|_\infty = \max_i \sum_j |a_{ij}|$.

3.3.1 Exercise

1. Find $L1,\ L2,\ L\infty$ norms of $x = (1,\ -13,\ 5,\ 3,\ -4)$.

2. Find the Column, the Euclidean and the Row norm of $A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 4 & 3 \\ 1 & 3 & 2 \end{bmatrix}$

3.4 Ill-conditioned Linear Systems

1. In practical applications, one usually encouners system of equations in which small changes in the coefficients of the system produce large changes in th solution. Such system are said to be *ill-conditioned.*

2. On the other hand, if the corespoinding changes in the solution is also small, then the system is *well-conditioned.*

Definition 1 (Condition Number)**.** The quantity $c(A) = \|A\|\|A^{-1}\|$, gives the measure of the condition of the matrix A, and is called the condition number. Here $\|.\|$ is any matrix norm.

Ill-conditioning can usually be expected when $det(A)$, in the system $Ax = b$, is small. Let $A = [a_{ij}]$ and $s_i = \left[a_{i1}^2 + a_{i2}^2 + ... + a_{in}^2\right]^{1/2}$. Let $k = \dfrac{det(A)}{s_1 s_2 ... s_n}$. Then the system whose coefficient matrix is A is ill-conditioned if k is very small compared to unity. Otherwise, it is well-conditioned.

3.4.1 Exercise

1. Show that the matrix $A = \begin{bmatrix} 25 & 24 & 10 \\ 66 & 78 & 37 \\ 92 & -73 & -80 \end{bmatrix}$ is ill-conditioned.

3.4.2 Method for Ill-conditioned Systems

In general, the accuracy of an approximate solution can be improved upon by an iterative procedure.

3.5 Iterative Method

The general system of linear equations with n unknowns and n linear equations is:

$$
\begin{gathered}
a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\
. \\
. \\
. \\
a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n
\end{gathered}
$$

For the iterative this system has to be rearranged as following

$$
\begin{aligned}
x_1 &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}}x_2 - \frac{a_{13}}{a_{11}}x_3 - ... - \frac{a_{1n}}{a_{11}}x_n \\
x_2 &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1 - \frac{a_{23}}{a_{22}}x_3 ... - \frac{a_{2n}}{a_{22}}x_n \\
. \\
. \\
. \\
x_n &= \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - ... - \frac{a_{n(n-1)}}{a_{nn}}x_{n-1}
\end{aligned}
$$

We can write this system as $x = Bx + C$. This equation is a iterative in nature as x is a function of x-itself. So the iteration formula is

$$
x^{n+1} = Bx^n + C \tag{3.3}
$$

The necessary condition for the convergence of the iterative method is that the diagonal elements a_{ii} do not vanish. If this is not the case then, the equations should be rearranged so that, the system satisfies this condition.

1. **Method of simultaneous displacements**

   This method 3.3 is due to *Jacobi* and is called method of simultaneous displacements, because $x_1, ..., x_n$ values are displaced simultaneous in 3.3. It can be shown that a sufficient condition for the convergence of this method is $\|B\| < 1$.

2. **Method of successive displacements** But if the values $x_1, ..., x_n$ are displaced sucessively; obtain x_1 first from the initialization, then use this value to obtain x_2 and so on, then this medthod is called method of successive displacements which is also commonly known as *Gauss-Seidel method*.

The Jacobi and Gauss-Seidel methods converge for any choice of the first approximation $(x_j^{(1)})$ $j = 1, 2, ..., n$, if every equation of the system satisfies the condition that

$$\sum_{j=1, j\neq i}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| \leq 1, \qquad i = 1, 2, ..., n \tag{3.4}$$

It is found that the Gauss-Seidel method converges **twice as fast** as the Jacobi method.

3.5.1 Exercise

1. $10x - 2y - z - u = 3, \quad -2x + 10y - z - u = 15, \quad -x - y + 10z - 2u = 27,$ $-x - y - 2z + 10u = 9$

2. $2x - y = 1, \quad -x + 3y - z = 8, \quad -y + 2z = -5$

3. An approximate solution of the system $10x_1 + x_2 + x_3 = 12, \quad x_1 + 10x_2 + x_3 = 12, \quad x_1 + x_2 + 10x_3 = 12$ is given as $x_1^{(0)} = 0.4, \quad x_2^{(0)} = 0.6, \quad x_3^{(0)} = 0.8$. Use the iterative method to improve this solution.

4. Solve the system: $10x + 2y + z = 9, \quad 2x + 20y - 2z = -44, \quad -2x + 3y + 10z = 22.$

5. Apply upto six iterations of Gauss-Seidel to solve: $28x + 4y - z = 32, \quad 2x + 17y + 4z = 35, \quad x + 3y + 10z = 24.$

6. **Cholesky's method** A matrix A is called a symmetric matrix if $a_{ij} = a_{ji}$. For symmetric matrix the LU decomposition is more conveniently obtained, since $U = L^T$. This method is called Cholesky's method. Solve the system by cholesky's method : $5x_1 + x_3 = 8, \quad -2x_2 = -4, \quad x_1 + 5x_3 = 16$