# Chapter 1

# PROLOGUE

## 1.1 Introduction

The subject of LINEAR ALGEBRA has its origin in the study of sytems of linear equations. Therefore, it is natural that we begin our course with a discussion on linear systems of equations. We shall begin our discussion not from the point of view of a rigorous treatment but restrict ourselves, for the moment, only to motivate the basic and fundamental questions that arise and that we will be discussing in this course. We shall begin our discussion with a simple system of equations involving two equations in two unknowns. Consider the simple system of equation

$$x + 3y = 6 \tag{1.1.1}$$

$$x - y = 2 \tag{1.1.2}$$

It can be easily verified that $x = 3, y = 1$ is a solution of this sytem, and moreover, it can be shown that this is the ONLY solution to this system. We can also geometrically interpret this as the point of intersection of the two lines represented by the above two equations, the solution, $(x, y) = (3, 1)$, being the coordinates of the point of intersection of these two lines. (See Figure 1.1).

When we look at the system from a geometric point of view, we see that we can, in general, look at two lines represented by the two equations

$$a_1x + a_2y = c \tag{1.1.3}$$

$$b_1x + b_2y = d \tag{1.1.4}$$

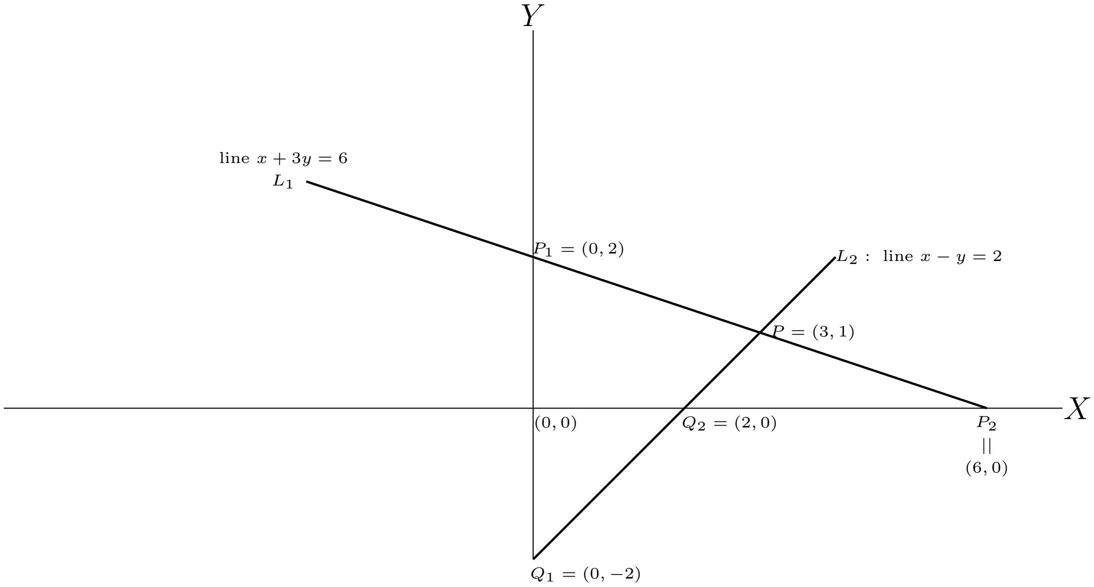

Figure 1.1

We can then interpret the problem of finding a solution to the system as that of finding the point of intersection of these two lines. However, this could lead to several possibilities, namely,

1. The two lines intersect at exactly one point, or,

2. The two lines coincide with each other, or,

3. The two lines are parallel

This leads to the following possibilities regarding the solution of the corresponding system.

1. In the first case when the two lines intersect at exactly one point, the coordinates of this point give us a UNIQUE solution of the system

2. In the second case, every point on the (common) line is a point of intersection and hence the coordinates of every point on the line give us a solution of the system. Thus in this case we have an INFINITE number of solutions of the system

3. In the third case, when the two lines are parallel, they do not have any point of intersection, and hence the system has NO solution

Thus, in general, a system of two linear equations in two unknowns may have exactly one solution, or an infinite number of solutions, or may not have any solution at all.
We would be, in general, interested in considering $m$ equations in $n$ unknowns, which is written in the form

$$\left.\begin{array}{cccccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1j}x_j & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2j}x_j & + & \cdots & + & a_{2n}x_n & = & b_2 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{i1}x_1 & + & a_{i2}x_2 & + & \cdots & + & a_{ij}x_j & + & \cdots & + & a_{in}x_n & = & b_i \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mj}x_j & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}\right\} (1.1.5)$$

In the above system, $x_1, x_2, \cdots, x_n$ are the $n$ unknowns to be determined. In the coefficients $a_{ij}$, the first index $i$ refers to the equation number and the second index $j$ refers to the unknown whose coefficient is $a_{ij}$. Thus $a_{34}$ refers to the fact that it is the coefficient of the unknown $x_4$ in the 3rd equation. Note that the $a_{ij}$ and $b_j$ are known and we have to determine the $x_j$. We write this in matrix notation as

$$Ax = b \tag{1.1.6}$$

where $A$ is the $m \times n$ matrix $A = (a_{ij})_{m\times n}$, $b$ is the $m \times 1$ column vector $b = (b_i)_{m\times 1}$ and $x$ is the $n \times 1$ column vector $x = (x_j)_{n\times 1}$.
Our main problem is the following:
**Main Problem**:
**Given the matrix $A$ find the solution(s) of the system for different $b$.**
Since we have already seen that we may have situations where the system may not have any solution, the first fundamental question that arises is the following:
**First Question**
**What is the condition that $b$ should satisfy in order that there exists a solution to the system $Ax = b$ ?**
Any such condition(s) is called **Consistency Condition**, and we shall denote these by [C]. So, given a $b$ the first question that we have to ask is the

following:
Does $b$ satisfy the condition(s) [C]?
Obviously there are two possible answers, namely "YES" and "NO", and each answer leads to several other natural questions. Let us first look at the basic questions that arise when $b$ satisfies the consistency condition, that is, when the answer to the above question is "YES". All the questions that arise are described in Figure 1.2
Let us next look at the case when $b$ does not satisfy the consistency condition [C]. In this case we can only conclude that there is NO SOLUTION to the system. What can we do in such a situation? In order to understand the question let us first understand the following question:
**What do we mean by the fact that the system has no solution**?
The system does not have any solution
$\Longrightarrow$
Whatever $x$ we choose and calculate $Ax$ this is not going to be equal to $b$. So if we take an $x$ as a solution we should have got $b$, but we get only $Ax$ which is not equal to $b$. So taking $x$ as a solution leads to an "Error", namely $b - Ax$. Note that the error, $b - Ax$, is an $m \times 1$ column vector. What we would like to do is to "**minimize**" this error. For this purpose we quantify the error $b - Ax$, by the sum of the squares of its components, that is,

$$\sum_{i=1}^{m} (b_i - (Ax)_i)^2$$

What we want to do is to minimize this error. Is it possible? What do we mean by this? Can we find an $x_\ell$ such that the error corresponding to $x_\ell$, namely,

$$\sum_{i=1}^{m} (b_i - (Ax_\ell)_i)^2$$

is the smallest error, that is,

$$\sum_{i=1}^{m} (b_i - (Ax_\ell)_i)^2 \;\; \leq \;\; \sum_{i=1}^{m} (b_i - (Ax)_i)^2 \text{ for all } x$$

We shall see later that this is possible and this leads us to the notion of "**Least Square Solution**".

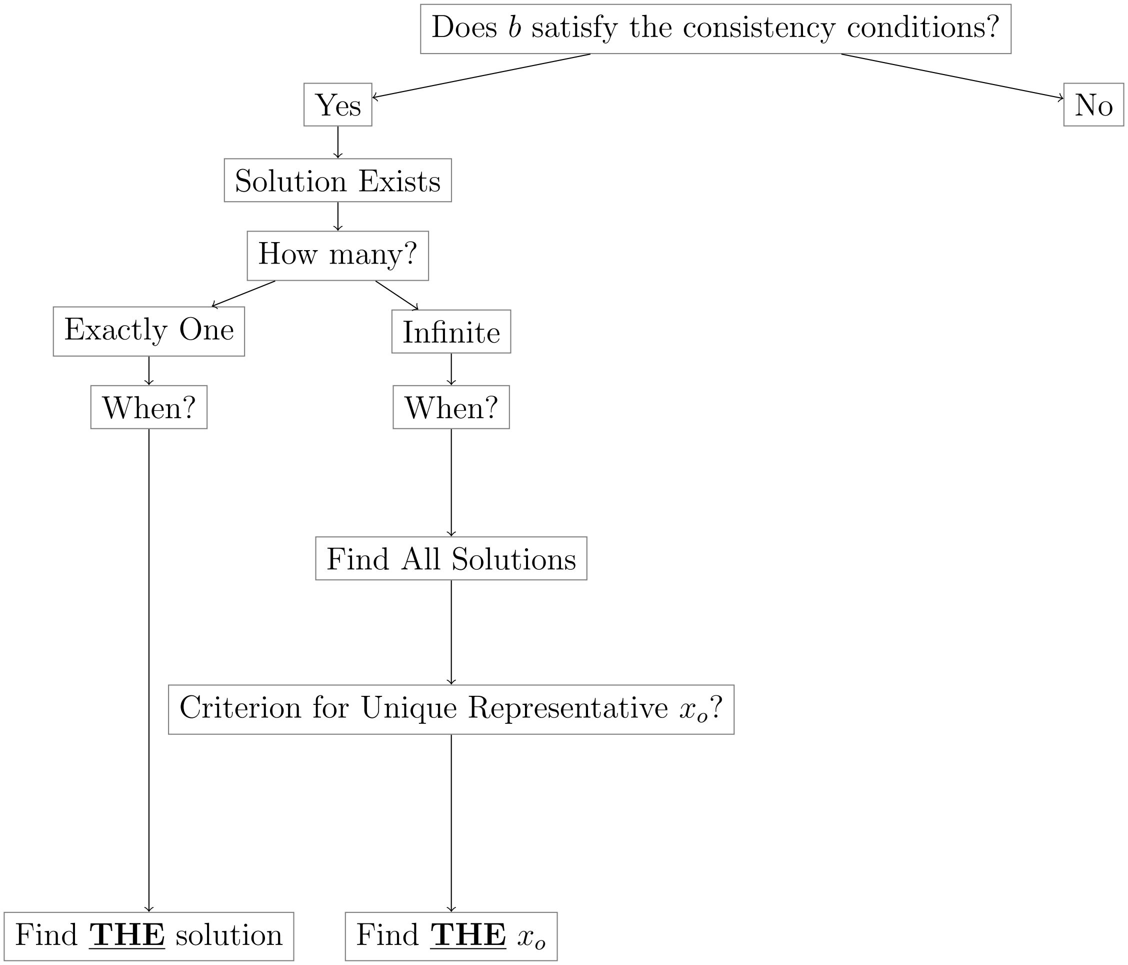
Does $b$ satisfy the consistency conditions?
Yes
No
Solution Exists
How many?
Exactly One
Infinite
When?
When?
Find All Solutions
Criterion for Unique Representative $x_o$?
Find **THE** solution
Find **THE** $x_o$

Figure 1.2

**Definition 1.1.1** If $b$ does not satisfy the consistency criterion [C], then any $n \times 1$ vector $x_\ell$ such that

$$\sum_{i=1}^{m} (b_i - (Ax_\ell)_i)^2 \quad \leq \quad \sum_{i=1}^{m} (b_i - (Ax)_i)^2 \text{ for all } x \tag{1.1.7}$$

is called a **Least Square Solution** for the system.

We can show that we can always get a least square solution if $b$ does not satisfy the consistency criterion [C]. The basic questions in the case when $b$ does not satisfy the cosistency criterion [C] are described in Figure 1.3.
We should develop mathematical theory to answer these questions and find techniques to find the solution

## 1.2 Simple Systems

What exactly do we mean by "easy" or "simple" systems? To understand this let us rephrase the question as follows: Why is it "difficult" to solve a general linear system? The main difficulty is the fact that each equation will involve all or at least several of the unknowns. Consequently no equation on its own is able to help us in solving for one of the unknowns. This means that the variables are "coupled" in a general system. When we say we want to solve the system, we actually mean that we want to get $x_1$,$x_2 \cdots, x_n$ separately, that is we want to uncouple the system. It is the uncoupling process that becomes difficult as more and more variables get coupled in each equation. So an easy or simple system is one in which such an uncoupling is easy. Clearly a system in which, to start with, there is no coupling, is an easy system since the uncoupling is already done. Each equation in such a system involves exactly one unknown. Let us look at the simple situation where we have $n$ equations in $n$ unknowns where the $i$th equation involves only the $i$th unknown $x_i$. Then the system is of the form

$$\left.\begin{array}{rcl} d_1x_1 & = & b_1 \\ d_2x_2 & = & b_2 \\ \cdots & \cdots & \cdots \\ d_ix_i & = & b_i \\ \cdots & \cdots & \cdots \\ d_nx_n & = & b_n \end{array}\right\} \tag{1.2.1}$$

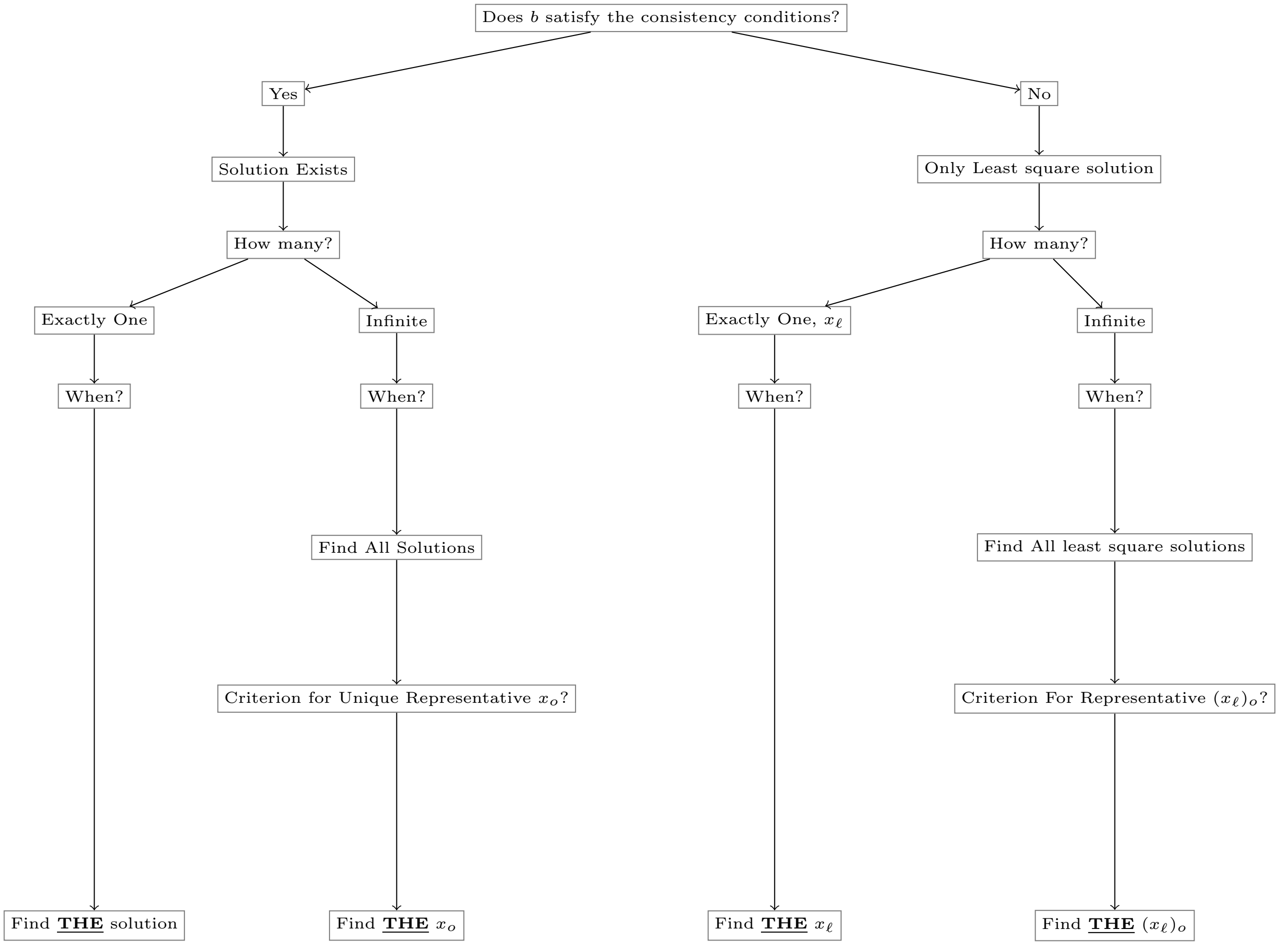
Does $b$ satisfy the consistency conditions?
Yes
No
Solution Exists
Only Least square solution
How many?
How many?
Exactly One
Infinite
Exactly One, $x_\ell$
Infinite
When?
When?
When?
When?
Find All Solutions
Find All least square solutions
Criterion for Unique Representative $x_o$?
Criterion For Representative $(x_\ell)_o$?
Find **THE** solution
Find **THE** $x_o$
Find **THE** $x_\ell$
Find **THE** $(x_\ell)_o$

Figure 1.3

If we now use the $i$th equation to solve for the $i$th unknown $x_i$, then we have to divide by $d_i$ which will be a problem if $d_i = 0$. So those equations of the above type, in which one or more of the $d_i$ happen to be zero will have some problem. Without loss of generality let us assume that we have arranged the system in such a way that the first $\rho$ of the $d_i$ are not zero and the remaining are zero, that is

$$\begin{aligned} d_1, d_2, \cdots, d_\rho &\neq 0 \qquad (1.2.2) \\ d_{\rho+1}, d_{\rho+2}, \cdots, d_n &= 0 \qquad (1.2.3) \end{aligned}$$

Then from the first $\rho$ equations we get

$$x_i = \frac{b_i}{d_i} \text{ for } 1 \leq i \leq \rho \qquad (1.2.4)$$

For $i = \rho + 1$ onwards the equations look like

$$0 = b_i \text{ for } \rho + 1 \leq i \leq n \qquad (1.2.5)$$

Thus we see that the Consistency Conditions are

$$b_i = 0 \text{ for } \rho + 1 \leq i \leq n \qquad (1.2.6)$$

When $b$ satisfies these conditions we see that whatever values we choose for $x_i$, for $\rho + 1 \leq i \leq n$, the equations from the $\rho + 1$th equation are satisfied. Thus we see that in this case we have the solution as

$$\left.\begin{aligned} x_i &= \tfrac{b_i}{d_i} \text{ for } 1 \leq i \leq \rho \\ x_i &= \text{any arbitrary value, for } \rho + 1 \leq i \leq n \end{aligned}\right\} \qquad (1.2.7)$$

Thus even if one of the $d_i$ is zero, that is if $\rho < n$, then one of the unknowns can take arbitrary values and hence the system has an infinite number of solutions. Hence we get the following result:
When $b$ satisfies the consistency condition (1.2.6), the system has a solution. Further,
i) If all the $d_i$ are nonzero, (that is $\rho = n$), then the system has a unique solution, and
ii) If $\rho < n$ of the $d_i$ are zero then the system has a solution if $b$ satisfies the consistency condition (1.2.6), and
iii) when $b$ satisfies the consistency condition (1.2.6), the system has infinite

number of solutions and we get all solutions from (1.2.7), by giving all possible values for the $x_i$ for $\rho+1 \leq i \leq n$.

Suppose next that the system (1.2.1) is such that (1.2.2) and (1.2.3) hold and the system is not consistent. This means that at least one of the $b_j$ for $\rho+1 \leq j \leq n$ must be nonzero. In this case we can look for only least square solutions. How do the least square solutions look like? For any vector $x$ we have

$$\sum_{i=1}^{n}(b_i - (Ax)_i)^2 \;=\; \sum_{i=1}^{\rho}(b_i - d_i x_i)^2 + \sum_{i=\rho+1}^{n} b_i^2$$

The rhs above will be minimum when

$$b_i - d_i x_i \;=\; 0 \text{ for } 1 \leq i \leq \rho$$

Whatever values of $x_j$ we choose for $\rho+1 \leq j \leq n$, the error above is not affected. Thus the vectors that give the least square error must be chosen such that

$$\left.\begin{array}{rcl} x_i & = & \frac{b_i}{d_i} \text{ for } 1 \leq i \leq \rho \\ x_i & = & \text{any arbitrary value, for } \rho+1 \leq i \leq n \end{array}\right\} \qquad (1.2.8)$$

Thus we can get all least square solutions by giving all possible values for the $x_j$ for $\rho+1 \leq j \leq n$.

The system that we have discussed above is of the form

$$d_i x_i \;=\; b_i \text{ for } 1 \leq i \leq n \qquad (1.2.9)$$

The matrix of the system is given by the diagonal matrix

$$A = \begin{pmatrix} d_1 & & & & & \\ & d_2 & & & & \\ & & \ddots & & & \\ & & & d_i & & \\ & & & & \ddots & \\ & & & & & d_n \end{pmatrix}$$

We see that the matrix is a diagonal matrix. Hence we call such systems as diagonal systems. The above discussion, therefore, tells us that diagonal systems are easy to handle.

In addition to diagonal systems, we also have other types of systems which are easy to handle. These are

1. **Lower Triangular Systems**, in which the coefficient matrix $A$ is a lower triangular matrix, that is

$$A = L \quad = \quad \begin{pmatrix} \ell_{11} & 0 & 0 & \cdots & 0 & 0 \\ \ell_{21} & \ell_{22} & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \ell_{(n-1)1} & \ell_{(n-1)2} & \ell_{(n-1)3} & \cdots & \ell_{(n-1)(n-1)} & 0 \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & \ell_{n(n-1)} & \ell_{nn} \end{pmatrix}$$

   Such systems can be analysed by forward substitution.

2. **Upper Triangular Systems**, in which the coefficient matrix $A$ is an upper triangular matrix, that is

$$A = U \quad = \quad \begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1(n-1)} & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2(n-1)} & u_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & u_{(n-1)(n-1)} & u_{(n-1)n} \\ 0 & 0 & 0 & \cdots & 0 & u_{nn} \end{pmatrix}$$

   Such systems can be analysed by backward substitution

We shall next see what we can do with general systems.

## 1.3 General Systems

Given any general system $Ax = b$ of $n$ equations in $n$ unknowns, we now see whether we can reduce the analysis to that of a diagonal system. To this end we look for a change of variables which could lead us to this. Let us consider a system

$$Ax \quad = \quad b \tag{1.3.1}$$

where $A$ is an $n \times n$ matrix, $b$ is an $n \times 1$ column vector and $x$ is the $n \times 1$ column vector that is to be determined satisfying the equation (1.3.1). Suppose we now introduce a change of variables,

$$y \quad = \quad Kx \tag{1.3.2}$$

Since we should be able to move from the new variable $y$ to the old variable $x$, and vice versa, we need to assume that our $K$ should be an invertible $n \times n$ matrix. Then we can write (1.3.2) as

$$x \;=\; Py \text{ where } P = K^{-1} \tag{1.3.3}$$

We shall similarly write

$$b \;=\; Pz \text{ where } z = Kb \tag{1.3.4}$$

Then the given system (1.3.1) can be written as

$$A(Py) \;=\; Pz$$

which can be rewritten as

$$Ty \;=\; z \tag{1.3.5}$$

where

$$T \;=\; P^{-1}AP \tag{1.3.6}$$

If we can solve for $y$ from (1.3.5) then we can get our $x$ from (1.3.3) as $x = Py$. So the question is whether we can find our transformation matrix $P$ such that it is easy to solve (1.3.5). From our discussion of the previous section we know that we can solve (1.3.5) easily if the matrix $T$ is a diagonal matrix. Thus we have the following conclusion:

**Suppose** we can find $P$ an invertible $n \times n$ matrix such that $T = P^{-1}AP$ is a DIAGONAL MATRIX. Then this $\Longrightarrow$

We can solve for $y$ easily from (1.3.5) (because it is a diagonal system). This $\Longrightarrow$

We can solve for $x$ by using (1.3.3). This $\Longrightarrow$

The original system (1.3.1) can be handled.

Thus if we can find such a $P$ we could handle the given system. Hence a fundamental question is the following:

**Diagonalization Problem**

**Given an $n \times n$ matrix $A$, can we find an invertible $n \times n$ matrix $P$ such that $P^{-1}AP$ is a diagonal matrix**?

The search for an answer to this question leads to some interesting canonical forms of matrices. We shall see through some examples the hurdles that we may face in getting the transformation $P$.

## 1.4 Questions Arising From Diagonalizability

We first look at some simple examples.

**Example 1.4.1** Consider the matrix $A \in \mathbb{R}^{2\times 2}$ given below:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \tag{1.4.1}$$

Then the matrix

$$P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{1.4.2}$$

is in $\mathbb{R}^{2\times 2}$ and is inveretible. In fact

$$P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \tag{1.4.3}$$

Further we have

$$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \tag{1.4.4}$$

a diagonal matrix in $\mathbb{R}^{2\times 2}$

**Example 1.4.2** Consider the matrix $A \in \mathbb{R}^{2\times 2}$ given below:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{1.4.5}$$

There is no invertible matrix $P \in \mathbb{R}^{2\times 2}$ such that $P^{-1}AP$ is diagonal matrix in $\mathbb{R}^{2\times 2}$. For suppose there exists such a $P$, say

$$P = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \tag{1.4.6}$$

Then we must have

$$P^{-1}AP = D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$

This $\Longrightarrow AP = PD$
$\Longrightarrow$

$$\begin{aligned}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix} &= \begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \\ &\Longrightarrow \\ \begin{pmatrix} r & s \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} d_1p & d_2q \\ d_1r & d_2s \end{pmatrix}\end{aligned}$$

Comparison of the entries on both sides leads to the conclusion that $P$ is not invertible, a contradiction since we assumed to start with that $P$ is invertible. Thus there is no $P \in \mathbb{R}^{2\times 2}$ invertible such that $P^{-1}AP$ is a diagonal matrix in $\mathbb{R}^{2\times 2}$

**Example 1.4.3** Consider the matrix $A \in \mathbb{R}^{2\times 2}$ given below:

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{1.4.7}$$

Again we can show that there is no $P \in \mathbb{R}^{2\times 2}$ invertible such that $P^{-1}AP$ is a diagonal matrix in $D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. For suppose there exists a $P = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \mathbb{R}^{2\times 2}$ which is invertible and such that $P^{-1}AP$ is a diagonal matrix. This then gives

$$\begin{aligned} AP &= PD \\ &\Longrightarrow \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix} &= \begin{pmatrix} d_1p & d_2q \\ d_1r & d_2s \end{pmatrix} \\ &\Longrightarrow \\ \begin{pmatrix} -r & -s \\ p & q \end{pmatrix} &= \begin{pmatrix} d_1p & d_2q \\ d_1r & d_2s \end{pmatrix} \\ &\Longrightarrow \\ -r &= d_1p \\ p &= d_1r \\ -s &= d_2q \\ q &= d_2s \end{aligned}$$

$$\begin{aligned}
&\Longrightarrow \\
(1+d_1^2)r &= 0 \\
(1+d_2^2)s &= 0 \\
&\Longrightarrow \\
r, s &= 0 \text{ since } d_1 \text{ and } d_2 \text{ are assumed to be real} \\
&\Longrightarrow
\end{aligned}$$

$P$ is not invertible, a contradiction. Hence there is no $P \in \mathbb{R}^{2\times 2}$ such that $P^{-1}AP$ is a diagonal matrix $D \in \mathbb{R}^{2\times 2}$.

**Example 1.4.4** On the other hand consider the above matrix given in (1.4.7) as a matrix in $\mathbb{C}^{2\times 2}$. Then the matrix

$$P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \tag{1.4.8}$$

is in $\mathbb{C}^{2\times 2}$, is invertible with

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{i}{2} \\ & \\ \frac{1}{2} & \frac{i}{2} \end{pmatrix} \tag{1.4.9}$$

and

$$P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{1.4.10}$$

which is a diagonal matrix in $\mathbb{C}^{2\times 2}$

From the above examples it is clear that

1. not all matrices can be diagonalised and
2. we must be precise about the the field $\mathcal{F}$ in which we want the matrices to be diagonalized.

We shall for most part of the course assume $\mathcal{F}$ to be either $\mathbb{R}$ or $\mathbb{C}$.
We have the first fundamental question about diagonalizability as,
**Question**
**What condition(s) should the matrix $A$ satisfy in order that there exists an invertible matrix $P$ such that $P^{-1}AP$ is a diagonal matrix**?
If we find this criterion, say [C], then then given any matrix $A$ we ask whether $A$ satisfies [C]. The various questions that arise out of this analysis are described in Figure 1.4

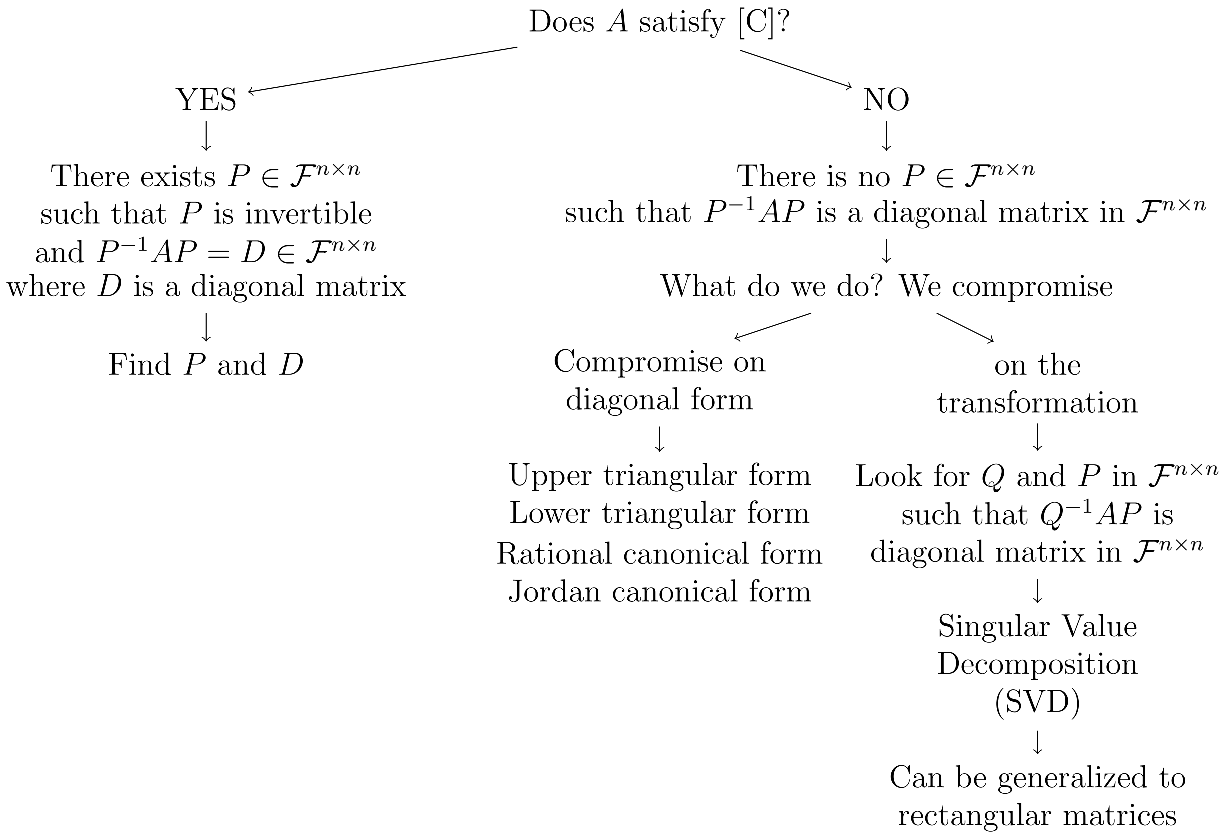
Does $A$ satisfy [C]?
YES
There exists $P \in \mathcal{F}^{n\times n}$ such that $P$ is invertible and $P^{-1}AP = D \in \mathcal{F}^{n\times n}$ where $D$ is a diagonal matrix
Find $P$ and $D$
NO
There is no $P \in \mathcal{F}^{n\times n}$ such that $P^{-1}AP$ is a diagonal matrix in $\mathcal{F}^{n\times n}$
What do we do? We compromise
Compromise on diagonal form
Upper triangular form
Lower triangular form
Rational canonical form
Jordan canonical form
on the transformation
Look for $Q$ and $P$ in $\mathcal{F}^{n\times n}$ such that $Q^{-1}AP$ is diagonal matrix in $\mathcal{F}^{n\times n}$
Singular Value Decomposition (SVD)
Can be generalized to rectangular matrices

Figure 1.4

## 1.5 Third Problem

We have discussed so far two important problems, namely,

1. Solving a system of equations, the basic questions regarding this being summarized in Figure 1.3, and

2. Diagonalization of a matrix

We have also seen that these two problems are interrelated. We shall now discuss a third problem which is also connected to these problems.

Consider two $n \times 1$ vectors,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

where $x_j$ and $y_j$ are all real. We then define the "**inner product**" of two such vectors, denoted by $(x, y)$, as

$$\begin{aligned} (x,y) &\overset{def}{=} y^T x \\ &= x_1y_1 + x_2y_2 + \cdots + x_ny_n \\ &= \sum_{i=1}^{n} x_iy_i \end{aligned}$$

Note that the inner product of two such real vectors is a real number. Let us look at a simple example

**Example 1.5.1** Consider the vectors

$$x = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ and } y = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

Then we have

$$\begin{aligned} (x,y) &= (1 \times (-1)) + (0 \times 0) + (1 \times 2) \\ &= 1 \end{aligned}$$

We shall now look at another type of product of two vectors.
Consider two $n \times 1$ vectors,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

where $x_j$ and $y_j$ are real. We then define the "**outer product**" or also known as "**Tensor Product**" of two such vectors, denoted by $x \otimes y$, as

$$x \otimes y \overset{def}{=} xy^T \tag{1.5.1}$$

Since $x$ is $n \times 1$ and $y^T$ is $1 \times n$ we see that the tensor product of two $n \times 1$ vectors is an $n \times n$ matrix. Clearly the $(i, j)$ th entry of the tensor product matrix $x \otimes y$ is given by $x_i y_j$. We have taken two vectors of the same size. We could generalise this by taking $x$ to be $m \times 1$ vector and $y$ to be an $n \times 1$ vector. Then we see that the tensor product $x \otimes y = xy^T$ is now an $m \times n$ matrix. Note that in general the tensor product $x \otimes y$ need not be equal to the tensor product $y \otimes x$.

**Example 1.5.2** Consider the two vectors,

$$x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \text{ and } y = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Then we have

$$\begin{aligned} x \otimes y &= xy^T \\ &= \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 3 \\ 2 & 3 \\ 4 & 6 \end{pmatrix} \end{aligned}$$

which is a $3 \times 2$ matrix since $x$ is $3 \times 1$ and $y^T$ is $1 \times 2$. Similarly we see that

$$\begin{aligned} y \otimes x &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -2 & 4 \\ 3 & -3 & 6 \end{pmatrix} \end{aligned}$$

which is a $2 \times 3$ matrix

Note that every row of the tensor product $x \otimes y$ is a scalar multiple of the vector $y^T$. Thus $x \otimes y$ has a simple structure. Taking the tensor product $x \otimes y$, where $x$ is $m \times 1$ and $y$ is $n \times 1$, is a simple way of generating an $m \times n$ matrix. We further observe for any scalar $\alpha \in \mathbb{R}$ we have $\alpha x \otimes y$ is also an $m \times n$ matrix. We can generate more $m \times n$ matrices as follows: Instead of taking one pair of vectors $x$ and $y$ and one scalar $\alpha$ , let us now take any positive interger $k$, and $k$ pairs of vectors, say, $u_1, v_1; u_2, v_2; \cdots, u_k, v_k$ where $u_j$ are all $m \times 1$ and $v_j$ are all $n \times 1$, and $k$ nonzero scalars $\alpha_1, \alpha_2, \cdots, \alpha_k$. For each pair $u_i, v_i$ and scalar $\alpha_i$ we can form the tensor product $u_i \otimes v_i$ and get an $m \times n$ matrix $\alpha_i u_i \otimes v_i$. We can now add all these to get

$$\sum_{i=1}^{k} \alpha_i u_i \otimes v_i$$

which is also an $m \times n$ matrix. By varying $k$ and varying the pairs $u_i, v_i$ and the scalars $\alpha_i$, we can construct, this way, a large collection of $m \times n$ matrices. We now ask the following question:

**Question**

**Does the above construction exhaust all $m \times n$ matrices**?

We shall see that the answer to this question is "YES". This means that every $m \times n$ matrix can be expressed as the sum of a finite number of tensor products as above. We would prefer to have the vectors $u_1, u_2, \cdots, u_k$ as orthonormal vectors in $\mathbb{R}^m$ and the vectors $v_1, v_2, \cdots, v_k$ as orthoormal vectors in $\mathbb{R}^n$ and the scalars $\alpha_1, \alpha_2, \cdots, \alpha_k$ as $> 0$. This then leads to the following questions: Given an $m \times n$ matrix

1. What is the minimum value of $k$ so that we can express $A$ as the sum of $k$ tensor products, and

2. Once we find the $k$ what is an efficient way of finding the orthonormal vectors $u_1, u_2, \cdots, u_k$ in $\mathbb{R}^m$, the orthonormal vectors $v_1, v_2, \cdots, v_k$ in $\mathbb{R}^n$ and the scalars $\alpha_1, \alpha_2, \cdots \alpha_n$ all $> 0$, to get the representation

$$A = \sum_{i=1}^{k} \alpha_i u_i \otimes v_i$$

One such decomposition of the matrix as a sum of a minimum number of tensor products is the Singular Value Decomposition. We shall see that this decomposition helps us to write down the pseudoinverse $A^\dagger$ of the matrix $A$ and hence facilitates solving the system $Ax = b$.

## 1.6 Fourth Problem

We shall next look at another important problem which is closely connected with all the three problems discussed above. In Figure 1.3 we have seen all the questions that arise in analysing a system of equations, $Ax = b$. We will have the following possible situations:

1. If $b$ satisfies the consistency condition then the following cases arise:

   (a) There is a UNIQUE solution for the system. Since the solution will depend on $b$ we shall denote this unique solution vector by $x_0(b)$

   (b) There are an infinite number of solutions, from which we want to select a unique representative solution. In this case we shall denote this unique representative solution by $x_0(b)$

2. If $b$ does not satisfy the consistency condition then we get only least square solution and the following cases arise:

   (a) There is a UNIQUE least square solution for the system. Since the least square solution will depend on $b$ we shall denote this unique solution vector by $x_0(b)$

   (b) There are an infinite number of least square solutions, from which we want to select a unique representative least square solution. In this case we shall denote this unique representative least square solution by $x_0(b)$

Thus in each case we want our final answer to be a single vector $x_0(b)$ which will depend on $b$. We can interpret this as follows:
We consider a linear system for which the inputs are all $n \times 1$ vectors, the transfer function is the $m \times n$ matrix $A$. Then for any input $x$ the output is $Ax$.

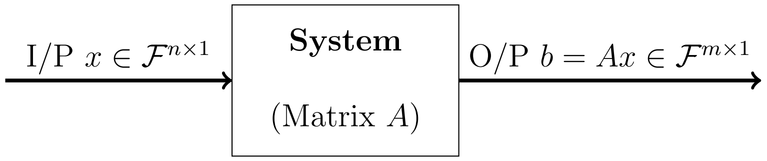
I/P $x \in \mathcal{F}^{n\times 1}$
**System**
(Matrix $A$)
O/P $b = Ax \in \mathcal{F}^{m\times 1}$

Figure 1.5

When we say that we want to solve the system $Ax = b$ what we mean is that we want a particular output $b$ and we want to determine the input $x$ that will produce this desired output $b$. Our final answer to this question is $x_0(b)$. What do we mean by this?

1. Whenever there is only one possible input that gives this desired output $b$, then our answer $x_0(b)$ will be precisely this input

2. Whenever there are infinite number of possible inputs that can give this desired output $b$, then our answer $x_0(b)$ will be the unique representative among these that will produce the output $b$. (The representative will be chosen based on some criterion)

3. Whenever there is no possible input that gives this desired output $b$, then our answer $x_0(b)$ will be precisely that input that can give the output which is "closest' to the desired output $b$.

4. Whenever there are infinite number of possible inputs that can give this output closest to the desired output $b$, then our answer $x_0(b)$ will be the unique representative among these that will produce the output closest to the desired output $b$. (The representative will be chosen based on some criterion)

Thus starting from the given desired output $b$ we want to construct $x_0(b)$. We therefore want to construct another system for which the inputs are $m \times 1$ vectors, and whose system matrix is an $n \times m$ matrix $A^{\dagger}$, such that when we input $b$ into this system the output $A^{\dagger}$ is precisely $x_0(b)$. Hence we are faced with the following fundamental problem:
**Given an $m \times n$ matrix $A$ construct an $n \times m$ matrix $A^{\dagger}$ such that $A^{\dagger}b = x_0(b)$ for every $m \times 1$ vector $b$**
We shall see that it is possible to construct such an $A^{\dagger}$. This matrix $A^{\dagger}$ is called the **Pseudoinverse** of the matrix $A$. It must be noted that we would like the $A^{\dagger}$ that we construct to be the actual inverse $A^{-1}$ of the matrix whenever $A$ is an invertible square matrix. However, the pseudoinverse will make sense even if $A$ is a singular square matrix or even if $A$ is a rectangular matrix.

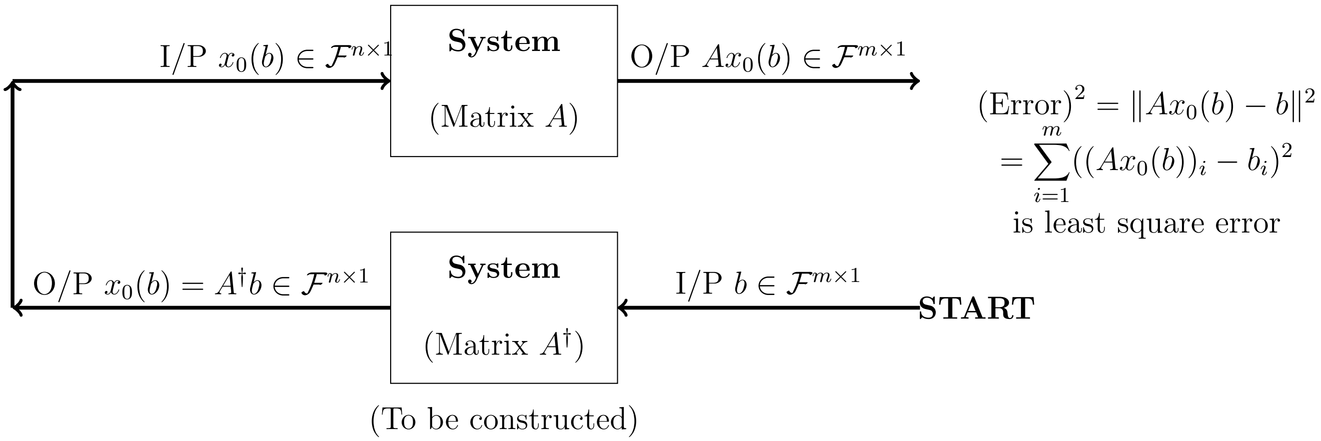
I/P $x_0(b) \in \mathcal{F}^{n\times 1}$
System
(Matrix $A$)
O/P $Ax_0(b) \in \mathcal{F}^{m\times 1}$
$(\text{Error})^2 = \|Ax_0(b) - b\|^2$
$= \sum_{i=1}^{m}((Ax_0(b))_i - b_i)^2$
is least square error
O/P $x_0(b) = A^\dagger b \in \mathcal{F}^{n\times 1}$
System
(Matrix $A^\dagger$)
I/P $b \in \mathcal{F}^{m\times 1}$
START
(To be constructed)

Figure 1.6

## 1.7 Summary

Summarising our discussion, we have highlighted four important problems, namely,

1. Solution of the system $Ax = b$
2. Diagonalization of a matrix and almost diagonalising a matrix
3. Decomposing a matrix as a finite sum of tensor products
4. Finding the pseudo inverse of a matrix

Finding the answers to these four problems will be the driving force for the course. It should be noted that the second, third and fourth problems arose out of the questions that we raised in Problem 1 on linear systems of equations. Thus all these problems are interrelated. In the analysis of Problems 2,3 and 4, (and hence obviously of Problem 1 also), a decisive role is played by the notions of eigenvalues and eigenvectors of a matrix.
In the context of Problem 3 of decomposing a matrix into the sum of a finite number of tensor products, two important special cases are,

1. Square Symmetric real matrices (or complex Hermitian matrices), and

2. square real or complex normal matrices

In these cases such a decomposition leads us to the "**Spectral Decomposition**".
The main goal of the course will be the following:

1. Develop the appropriate mathematical framework to analyse these problems

2. Find the answers to the various questions we have raised in these four problems

3. Look for generalisation of these ideas

## 1.8 Normal equation

We shall now look at a system related to the given system (1.1.6). Premultiplying both sides of the system (1.1.6) we get

$$A^T Ax = A^T b \tag{1.8.1}$$

We can write this as

$$Nx = y \tag{1.8.2}$$

where

$$N = A^T A \tag{1.8.3}$$
$$y = A^T b \tag{1.8.4}$$

The system (1.8.2) is called the "**Normal System**" or just Normal Equation. We shall see that it is directly connected with the solutions of the given system (1.1.6) as follows:

- The coefficient matrix $N$ of the Normal System is a square symmetric matrix

- The Normal System (1.8.2) is consistent irrespective of whether the given system (1.1.6) is consistent or not

- Whenever the given system (1.1.6) is consistent, the set of solutions of (1.1.6), is the same as the set of solutions of the Normal System (1.8.2)

- Whenever the given system (1.1.6) is not consistent, the set of least square solutions of (1.1.6), is the same as the set of solutions of the Normal System (1.8.2)

**Example 1.8.1** Consider the system

$$Ax = b \tag{1.8.5}$$

where

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 2 & 0 & 2 & 0 \end{pmatrix} \tag{1.8.6}$$

We then have

$$\begin{aligned} A^T A &= \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 2 & 0 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 0 & 6 & 0 \\ 0 & 2 & 0 & 2 \\ 6 & 0 & 6 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \end{aligned}$$

Hence the Normal System is given by

$$Nx = y \tag{1.8.7}$$

where

$$N = \begin{pmatrix} 6 & 0 & 6 & 0 \\ 0 & 2 & 0 & 2 \\ 6 & 0 & 6 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \tag{1.8.8}$$

$$y = A^T b$$

$$= \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$= \begin{pmatrix} b_1 + b_2 + 2b_3 \\ b_1 - b_2 \\ b_1 + b_2 + 2b_3 \\ b_1 - b_2 \end{pmatrix} \qquad (1.8.9)$$

Clearly with this $y$ the normal system (1.8.7) is consistent.
At first let us consider the vector

$$b = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \qquad (1.8.10)$$

Then the system (1.8.5) becomes

$$\left.\begin{array}{lcl} x_1 + x_2 + x_3 + x_4 & = & 1 \\ x_1 - x_2 + x_3 - x_4 & = & 1 \\ 2x_1 + 2x_3 & = & 2 \end{array}\right\} \qquad (1.8.11)$$

The third equation is the sum of the first two equations. The system is consistent. We have, from these,

$$\left.\begin{array}{lcl} x_1 + x_3 & = & 1 \\ x_2 + x_4 & = & 0 \end{array}\right\} \qquad (1.8.12)$$

Eliminating $x_3$ and $x_4$ we get

$$\left.\begin{array}{lcl} x_3 & = & 1 - x_1 \\ x_4 & = & -x_2 \end{array}\right\} \qquad (1.8.13)$$

where $x_1$ and $x_2$ can be chosen arbitrarily. Hence the general solution of the given system (1.8.5) with the rhs $b$ as given in (1.8.10) can be written as

$$x = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ -\alpha \\ -\beta \end{pmatrix} \qquad (1.8.14)$$

where $\alpha$ and $\beta$ can be chosen arbitrarily. Let us now look at the corresponding Normal equation with this $b$. we have, from (1.8.9) and (1.8.10),

$$y \;=\; A^T b = \begin{pmatrix} 6 \\ 0 \\ 6 \\ 0 \end{pmatrix} \tag{1.8.15}$$

Hence the normal equation $Nx = y$ beomes

$$\left.\begin{array}{rcl} 6x_1 + 6x_3 &=& 6 \\ 2x_2 + 2x_4 &=& 0 \\ 6x_1 + 6x_3 &=& 6 \\ 2x_2 + 2x_4 &=& 0 \end{array}\right\} \tag{1.8.16}$$

The third and fourth equations are the same as the first and second and hence we can write this system as

$$\left.\begin{array}{rcl} x_1 + x_3 &=& 1 \\ x_2 + x_4 &=& 0 \end{array}\right\} \tag{1.8.17}$$

which is the same as (1.8.12) that we obtained for the original system (1.8.5). Hence the normal system has the same set of solutions as the original system. Next let us consider the given system (1.8.5) with the vector $b$ now as

$$b \;=\; \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \tag{1.8.18}$$

Let us first look at the Normal system. We have from (1.8.9),

$$y \;=\; \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \end{pmatrix} \tag{1.8.19}$$

Hence the Normal system becomes

$$\left.\begin{array}{rcl} 6x_1 + 6x_3 &=& 2 \\ 2x_2 + 2x_4 &=& 0 \\ 6x_1 + 6x_3 &=& 2 \\ 2x_2 + 2x_4 &=& 0 \end{array}\right\} \tag{1.8.20}$$

which is consistent. On the other hand the given system (1.8.5) with this $b$ is inconsistent and hence we can only find least square solutions.. We have for any $x$ the square error as

$$\begin{aligned}\|Ax-b\|^2 &= ((Ax)_1-b_1)^2+((Ax)_2-b_2)^2+((Ax)_3-b_3)^2 \\ &= (x_1+x_2+x_3+x_4-1)^2 \\ &\quad +(x_1-x_2+x_3-x_4-1)^2 \\ &\quad +(2x_1+2x_3)^2\end{aligned}$$

Let us denote this error by $F(x_1,x_2,x_3,x_4)$, that is

$$F(x_1,x_2,x_3,x_4) = \left\{\begin{array}{l}(x_1+x_2+x_3+x_4-1)^2 \\ +(x_1-x_2+x_3-x_4-1)^2 \\ +(2x_1+2x_3)^2\end{array}\right. \tag{1.8.21}$$

Our aim is to find $x_1, x_2, x_3$ and $x_4$ such that the above is minimum. Hence from calculus we know that we have to choose these such that

$$\frac{\partial F}{\partial x_i} = 0 \text{ for } 1 \leq i \leq 4 \tag{1.8.22}$$

Hence we get, differentiating and simplifyng,

$$\left.\begin{aligned}6x_1+6x_3 &= 2 \\ 2x_2+2x_4 &= 0 \\ 6x_1+6x_3 &= 2 \\ 2x_2+2x_4 &= 0\end{aligned}\right\}$$

which is the same as the normal equation (1.8.20) we obtained above. Hence the set of least square solutions for the given system is the same as the set of solutions of the Normal system. From (1.8.20) we, therefore, get the least square solutions as,

$$x = \begin{pmatrix}0\\0\\\frac{1}{3}\\0\end{pmatrix} + \begin{pmatrix}\alpha\\\beta\\-\alpha\\-\beta\end{pmatrix} \tag{1.8.23}$$

where $\alpha$, $\beta$ can be chosen arbitrarily.