1 BASICS OF LINEAR ALGEBRA

1.1 Basic Definitions

Definitions

- A linear equation in n variables \( x_1, x_2, \ldots, x_n \) is one that can be expressed in the form
  \[
  a_1x_1 + a_2x_2 + \cdots + a_nx_n = b
  \]
  where \( a_1, a_2, \ldots, a_n, \) and \( b \) are real constants. The variables in a linear equation are sometimes called unknowns.

- An arbitrary System of m linear equations in n unknowns is one that can be written as
  \[
  a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1
  
  a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2
  
  \vdots
  
  a_{1m}x_1 + a_{1m}x_2 + \cdots + a_{mn}x_n = b_m
  \]
  where once again \( x_1, x_2, \ldots, x_n \) are the unknowns and the subscripted \( a \)'s and \( b \)'s denote constants.

- A system of linear equations is said to be homogeneous if the system has the form
  \[
  a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0
  
  a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0
  
  \vdots
  
  a_{1m}x_1 + a_{1m}x_2 + \cdots + a_{mn}x_n = 0
  \]
Example

\[4x_1 - x_2 + 3x_3 = -1\]
\[3x_1 + x_2 + 9x_3 = -4\]

is a system of 2 linear equations in 3 unknowns. The system has the solution \(x_1 = 1, x_2 = 2, x_3 = -1\) since these values satisfy both equations. The set of values \(x_1 = 1, x_2 = 8, x_3 = 1\) is not a solution since these values satisfy only the first of the two equations.

1.2 Solving a System of Linear Equations

1.2.1 The Augmented Matrix

An arbitrary system of \(m\) linear equations in \(n\) unknowns

\[a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1\]
\[a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2\]
\[\vdots\]
\[a_{1m}x_1 + a_{1m}x_2 + \cdots + a_{mn}x_n = b_m\]

can be written in the form of an **Augmented Matrix:**

\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
  a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn} & b_m
\end{bmatrix}
\]

Remark

When constructing an augmented matrix, the unknowns must appear in the same order in each equation and the constants must be on the right.
1.2.2 Elementary Row Operations

A system of linear equations can be solved by using certain operations to simplify the augmented matrix. These operations are called **Elementary Row Operations**.

- The three **Elementary Row Operations** are:
  
  1. Multiplying row $i$ by a non-zero number $t$:

     $$ R_i \rightarrow tR_i $$

  2. Interchange rows $i$ and $j$:

     $$ R_i \leftrightarrow R_j $$

  3. Adding $t$ times row $i$ to row $j$:

     $$ R_j \rightarrow R_j + tR_i $$

- Matrix $A$ is said to be **row-equivalent** to matrix $B$ if $B$ can be obtained from $A$ by a sequence of elementary row operations. Clearly if $B$ is row-equivalent to $A$ then $A$ is row-equivalent to $B$.

**Remark**

Elementary row operations can be used to convert an augmented matrix into a special form called reduced row-echelon form without changing the solution set. Once the conversion is complete the solution set is easily obtained (as we will see below).

1.2.3 Reduced Row-Echelon Form

**Definitions**

- A matrix is in **reduced row-echelon form** (rref) when the following four conditions are satisfied:
1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. (This is called a **leading 1**.)

2. If there are any rows that consists entirely of zeros, then they are grouped together at the bottom of the matrix.

3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the higher row occurs closer to the left than the leading 1 in the lower row.

4. Each column that contains a leading 1 has zeros everywhere else.

- A matrix having properties 1, 2 and 3 (but not necessarily 4) is said to be in **row echelon form** or **ref**.

**Examples**

- The following matrices are in reduced row-echelon form:

  \[
  \begin{bmatrix}
  1 & 0 & 0 & 4 \\
  0 & 1 & 0 & 7 \\
  0 & 0 & 1 & -1 \\
  \end{bmatrix},
  \begin{bmatrix}
  1 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 1 \\
  \end{bmatrix},
  \begin{bmatrix}
  0 & 1 & -2 & 0 & 1 \\
  0 & 0 & 0 & 1 & 3 \\
  0 & 0 & 0 & 0 & 0 \\
  \end{bmatrix}
  \]

- The following matrices are in row-echelon form:

  \[
  \begin{bmatrix}
  1 & 4 & 3 & 7 \\
  0 & 1 & 6 & 2 \\
  0 & 0 & 1 & 5 \\
  \end{bmatrix},
  \begin{bmatrix}
  1 & 1 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 0 \\
  \end{bmatrix},
  \begin{bmatrix}
  0 & 1 & 2 & 6 & 0 \\
  0 & 0 & 1 & -1 & 0 \\
  0 & 0 & 0 & 0 & 1 \\
  \end{bmatrix}
  \]
1.2.4 Gauss Jordan Algorithm

Let $A$ be the augmented matrix that completely describes a given system of $m$ linear
equations in $n$ unknowns, and let the rref of $A$ be $B$. Given the matrix $A$, how is $B$
obtained?

The Gauss Jordan Algorithm converts a matrix $A$ to its rref. The procedure is
outlined below by the use of an example. We wish to reduce the following matrix to
reduced row-echelon form.

$$A = \begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

$$R_1 \leftrightarrow R_2 \begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

$$R_1 \rightarrow \frac{1}{2}R_1 \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

$$R_3 \rightarrow R_3 + -2R_1 \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

$$R_2 \rightarrow \frac{-1}{2}R_2 \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & \frac{-7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$
\[
R_3 \rightarrow R_3 + -5R_2 \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & \frac{-7}{2} & -6 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 1 \end{bmatrix}
\]

\[
R_3 \rightarrow 2R_3 \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & \frac{-7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}
\]

The entire matrix is now in row-echelon form. To find the reduced row-echelon form we continue using the elementary row operations:

\[
R_2 \rightarrow R_2 + \frac{7}{2}R_3 \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}
\]

\[
R_1 \rightarrow R_1 + -6R_3 \begin{bmatrix} 1 & 2 & -5 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}
\]

\[
R_1 \rightarrow R_1 + 5R_2 \begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} = B
\]

Remark

It can be shown that every matrix has a unique reduced row-echelon form.

1.3 A Sample Solution using Gauss-Jordan

We now apply the row reducing method above to solve a system of linear equations.
Solve by Gauss-Jordan elimination:

\[ \begin{align*}
x_1 + 3x_2 - 2x_3 &+ 2x_5 = 0 \\
2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\
5x_3 + 10x_4 + 15x_6 &= 5 \\
2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 &= 6
\end{align*} \]

1. Form the corresponding Augmented Matrix for the system:

\[
\begin{bmatrix}
1 & 3 & -2 & 0 & 2 & 0 & 0 \\
2 & 6 & -5 & -2 & 4 & -3 & -1 \\
0 & 0 & 5 & 10 & 0 & 15 & 5 \\
2 & 6 & 0 & 8 & 4 & 18 & 6
\end{bmatrix}
\]

2. Convert the Augmented Matrix to rref. Here’s one possible procedure:

\[
\begin{align*}
R_2 &\rightarrow R_2 + -2R_1 \\
R_4 &\rightarrow R_4 + -2R_1 \\
R_2 &\rightarrow -R_2 \\
R_3 &\rightarrow R_3 + -5R_2 \\
R_4 &\rightarrow R_4 + -4R_2 \\
R_3 &\leftrightarrow R_4 \\
R_3 &\rightarrow \frac{1}{6}R_3 \\
R_2 &\rightarrow R_2 + -3R_3 \\
R_1 &\rightarrow R_1 + 2R_2
\end{align*}
\]

\[
\begin{bmatrix}
1 & 3 & 0 & 4 & 2 & 0 & 0 \\
0 & 0 & 1 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[ \text{check!} \]
3. Write down the corresponding system of equations:

\[
\begin{align*}
x_1 + 3x_2 + 4x_4 + 2x_5 &= 0 \\
x_3 + 2x_4 &= 0 \\
x_6 &= \frac{1}{3}
\end{align*}
\]

4. Solve each equation for the leading variables. In this case they are \(x_1, x_3\) and \(x_6\):

\[
\begin{align*}
x_1 &= -3x_2 - 4x_4 - 2x_5 \\
x_3 &= -2x_4 \\
x_6 &= \frac{1}{3}
\end{align*}
\]

5. Assign the independent variables arbitrary values. In this case the independent variables are \(x_2, x_4\) and \(x_5\). Assign the arbitrary values \(r, s\) and \(t\) respectively. The general solution is given by the formulas.

\[
x_1 = -3r - 4s - 2t, x_2 = r, x_3 = -2s, x_4 = s, x_5 = t, x_6 = \frac{1}{3}
\]

1.4 Consistent and Inconsistent Systems

Definitions

- A system of linear equations that has no solution is said to be **inconsistent**.

- A system of linear equations with at least one solution is said to be **consistent**.

1.4.1 Example in the \(x - y\) Plane

To illustrate the possibilities that can occur in solving systems of linear equations, consider a general system of two linear equations in the unknowns \(x\) and \(y\):

\[
\begin{align*}
l_1 : & \quad a_1x + b_1y = c_1 \\
l_2 : & \quad a_2x + b_2y = c_2
\end{align*}
\]
(NB $a_1$ and $b_1$ not both zero and $a_2$ and $b_2$ not both zero). The graphs of these equations are lines: call them $l_1$ and $l_2$. The solutions to the system of equations correspond to points of intersection of $l_1$ and $l_2$. There are three possibilities:

- The lines $l_1$ and $l_2$ may be parallel, in which case there is no intersection and consequently no solution to the system. Thus the system of equations is said to be **inconsistent**.

- The lines $l_1$ and $l_2$ may intersect at only one point, in which case the system has exactly one solution. Thus the system of equations is said to be **consistent**.

- The lines $l_1$ and $l_2$ may coincide, in which case there are infinitely many points of intersection and consequently infinitely many solutions to the system. Thus this system is also said to be **consistent**.

Although we have only demonstrated this for a special case, it can be shown in general that the following theorem holds.

**Theorem 1.1** Every system of linear equations has either no solutions, exactly one solution, or infinitely many solutions.

### 1.5 Solving Homogenous Systems of Linear Equations

Solve the homogenous system of linear equations:

\[
\begin{align*}
2x_1 + 2x_2 - x_3 + x_5 &= 0 \\
-x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\
x_1 + x_2 - 2x_3 - x_5 &= 0 \\
x_3 + x_4 + x_5 &= 0
\end{align*}
\]
The augmented matrix for the system can be reduced to the following rref:

\[
\begin{bmatrix}
1 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

The corresponding system of equations is

\[
\begin{align*}
x_1 + x_2 + x_5 &= 0 \\
x_3 + x_5 &= 0 \\
x_4 &= 0
\end{align*}
\]

Solve each equation for the leading variables

\[
\begin{align*}
x_1 &= -x_2 - x_5 \\
x_3 &= -x_5 \\
x_4 &= 0
\end{align*}
\]

Then the general solution is

\[
x_1 = -s - t, x_2 = s, x_3 = -t, x_4 = 0, x_5 = t
\]

Note that the trivial solution is obtained when \( s = t = 0 \).

This example illustrates two important points about solving a homogenous system of linear equations:

- None of the three elementary row operations alter the final column of zero’s.
  Therefore, the corresponding system of equations to the rref is also a homogenous system of linear equations.

- The number of equations in the reduced system is less than or equal to the number of equations in the original system.

These two observations are instrumental in proving the following theorem.
**Theorem 1.2** A Homogenous System of linear equations with more unknowns than equations has infinitely many solutions.

**Proof:** If a given homogenous system has $m$ equations in $n$ unknowns such that $m < n$, and if there are $r$ non-zero rows in the rref of the augmented matrix, it follows that $r < n$. The system of equations corresponding to the rref of the augmented matrix will have the following form:

\[
\begin{align*}
\cdots x_{k_1} &+ \sum() = 0 \\
\cdots x_{k_2} &+ \sum() = 0 \\
\vdots \\
\cdots x_{k_r} &+ \sum() = 0
\end{align*}
\]

where $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ are the leading variables and $\sum()$ denotes sums (possibly all different) that involve the $n - r$ free variables. Solving for the leading variables gives

\[
\begin{align*}
x_{k_1} &= -\sum() \\
x_{k_2} &= -\sum() \\
\vdots \\
x_{k_r} &= -\sum()
\end{align*}
\]

As in the above example, arbitrary values can be assigned to the free variables on the right hand side and thus infinitely many solutions are obtained for the system. ☐

### 1.6 Important Matrices

There are two matrices which are of special significance in linear algebra:

**Definitions**

- Square matrices taking the form of 1’s on the main diagonal and 0’s everywhere else, are called **Identity Matrices**. $I_n$ denotes the $n \times n$ identity matrix.
Example:

\[
I_3 = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

If \( A \) is an \( m \times n \) matrix, then

\[ AI_n = A \text{ and } I_mA = A \]

Thus the role of the identity matrix in matrix multiplication is analogous to multiplying a number by 1: \( a \cdot 1 = 1 \cdot a = 1 \).

- If \( A \) is an \( n \times n \) matrix, and if an \( n \times n \) matrix \( B \) can be found such that \( AB = BA = I \), then \( A \) is said to be invertible or non-singular and \( B \) is called the inverse of \( A \). If \( A \) is not invertible it is said to be singular.

- A square matrix in which all of the entries off the main diagonal are zero is called a diagonal matrix. Some examples are:

\[
\begin{bmatrix}
2 & 0 \\
0 & -5
\end{bmatrix}, \quad \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}, \quad \begin{bmatrix}
6 & 0 & 0 \\
0 & -4 & 0 \\
0 & 0 & 8
\end{bmatrix}
\]

A general \( n \times n \) diagonal matrix \( D \) can be written as

\[
D = \begin{bmatrix}
d_1 & 0 & \ldots & 0 \\
0 & d_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & d_n
\end{bmatrix}
\]

1.6.1 Properties of Diagonal Matrices

Diagonal matrices enjoy the following properties:
1. A diagonal matrix is invertible if and only if all of its diagonal entries are nonzero; in this case the inverse of the general form is

\[
D^{-1} = \begin{bmatrix}
\frac{1}{d_1} & 0 & \cdots & 0 \\
0 & \frac{1}{d_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{d_n}
\end{bmatrix}
\]

Exercise: verify that \(DD^{-1} = D^{-1}D = I\).

2. Powers of diagonal matrices: it can be easily shown that if \(D\) is the general diagonal matrix and \(k\) is any positive integer, then

\[
D^k = \begin{bmatrix}
d_1^k & 0 & \cdots & 0 \\
0 & d_2^k & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & d_n^k
\end{bmatrix}
\]

3. A diagonal matrix is its own transpose: \(D = D^T\).

4. Multiplying a matrix \(A\) on the left by a diagonal matrix \(D\) is the same as multiplying each row of \(A\) by the corresponding diagonal entry of \(D\):

\[
\begin{bmatrix}
d_1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_3
\end{bmatrix}
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
a_{31} & a_{32} & a_{33} & a_{34}
\end{bmatrix}
= 
\begin{bmatrix}
d_{1}a_{11} & d_{1}a_{12} & d_{1}a_{13} & d_{1}a_{14} \\
d_{2}a_{21} & d_{2}a_{22} & d_{2}a_{23} & d_{2}a_{24} \\
d_{3}a_{31} & d_{3}a_{32} & d_{3}a_{33} & d_{3}a_{34}
\end{bmatrix}
\]

5. Multiplying a matrix \(A\) on the right by a diagonal matrix \(D\) is the same as multiplying each row of \(A\) by the corresponding diagonal entries of \(D\):

\[
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33} \\
a_{41} & a_{42} & a_{43}
\end{bmatrix}
\begin{bmatrix}
d_1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_3
\end{bmatrix}
= 
\begin{bmatrix}
d_{1}a_{11} & d_{2}a_{12} & d_{3}a_{13} \\
d_{1}a_{21} & d_{2}a_{22} & d_{3}a_{23} \\
d_{1}a_{31} & d_{2}a_{32} & d_{3}a_{33} \\
d_{1}a_{41} & d_{2}a_{42} & d_{3}a_{43}
\end{bmatrix}
\]