INSTITIÚID TEICNEOLAÍOCHTA CHEATHARLACH

INSTITUTE OF TECHNOLOGY CARLOW

MATRICES AND LINEAR EQUATIONS

Contents

1 Matrices and Linear Equations

1.1 Introduction

Linear Algebra is a fundamental topic of mathematics for students of computer science. An extensive subject, it covers topics such as vectors and matrices, systems of linear equations, determinants and other topics. Linear Algebra has important geometric applications and hence is central to subjects like computer graphics. Matrices are of fundamental importance in graphics programming. We will see how simple rotations, reflections, projections etc. of images in two and three dimensions can be achieved by matrix multiplication.

Consider the following problem:

Example

Solve the system of linear equations:

$$x + 3y = 5 \tag{1}$$
$$2x + y = 2 \tag{2}$$

Method A.

Multiply (1) by 2 to obtain the equations

$$\begin{aligned} 2x + 6y &= 10 \\ 2x + y &= 2 \end{aligned}$$

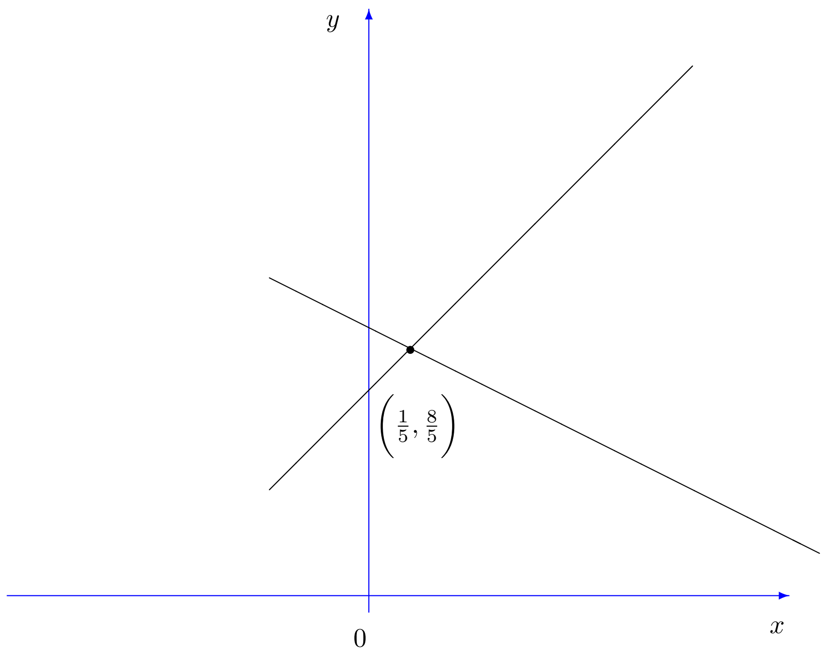
Subtract the second equation from the first to get the equation

$$\begin{aligned} 5y &= 8 \\ &= \frac{8}{5} \end{aligned}$$

Back substituting into (1) yields

$$\begin{aligned} x &= 5 - 3y \\ &= \frac{1}{5} \end{aligned}$$

The solution is $(\frac{1}{5}, \frac{8}{5})$.



Method B.

Rewrite the equations, in an obvious way, as an array of numbers:

$$\begin{array}{cc|c} 1 & 3 & 5 \\ 2 & 1 & 2 \end{array}$$

Subtract twice row 1. from row 2 to obtain the array

$$\begin{array}{cc|c} 1 & 3 & 5 \\ 0 & -5 & -8 \end{array}$$

Row 2 of this array now represents the equation

$$\begin{aligned} 0x - 5y &= -8 \\ y &= \frac{8}{5} \end{aligned}$$

From row 1, we find that

$$\begin{aligned} x &= 5 - 3y \\ &= \frac{1}{5} \end{aligned}$$

The solution is $(\frac{1}{5}, \frac{8}{5})$.

Remark

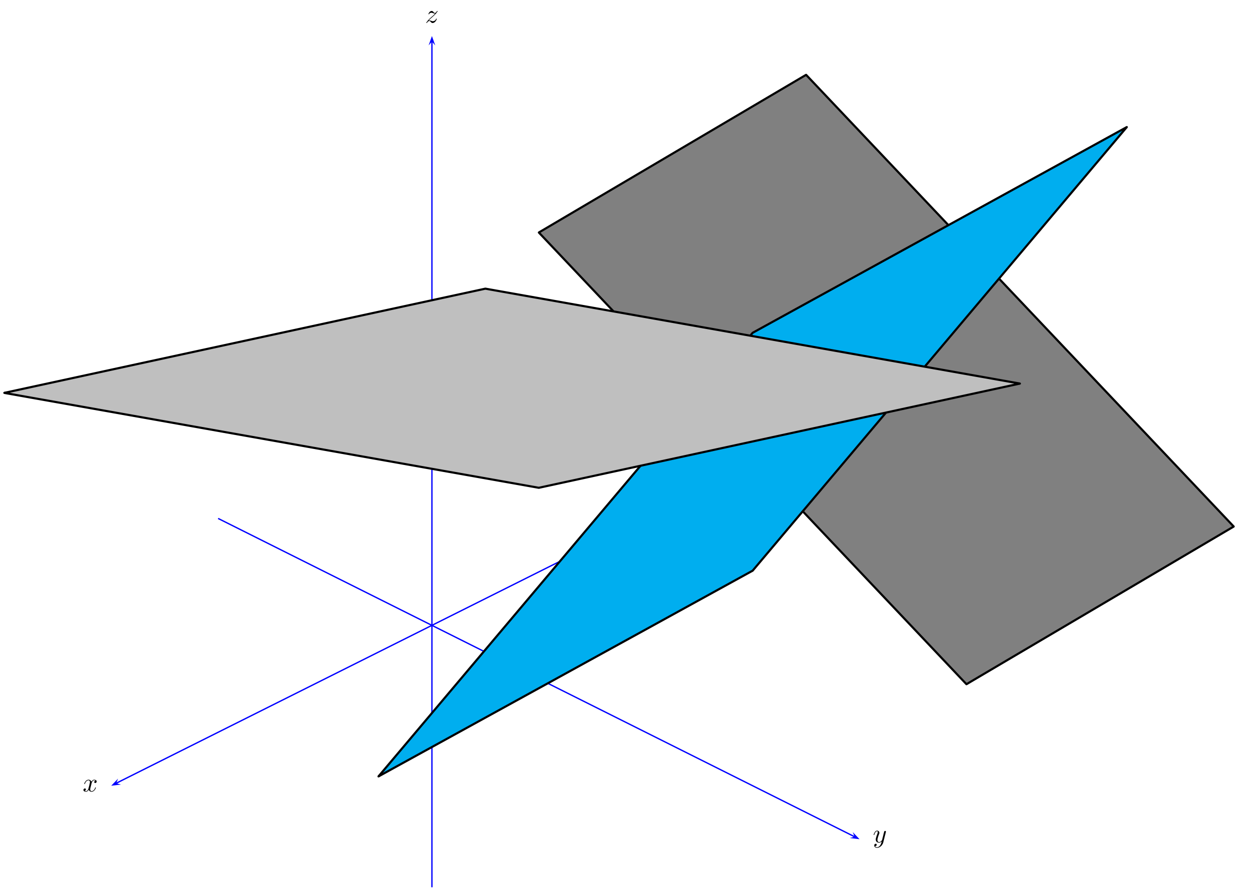
In solving a system like this, consisting of two linear equations in two unknowns, it makes little difference which method we use, though the second method has one obvious advantage: we do not have to keep writing the symbols x and y for the unknowns. It is a greater advantage in more complicated systems to work with the coefficients alone; instead of writing the unknowns in order to 'label' the coefficients, we keep tract of the coefficients by maintaining their positions in an array. Of course, the main advantage of the second method is that we can write computer programs which use this method of manipulating arrays of numbers to solve linear equations.

Let us consider a second problem to illustrate the use of an array in solving a system of linear equations.

Example

Solve the system of linear equations:

$$\begin{aligned} x + 2y - 3z &= 4 \\ x + 3y + z &= 11 \\ 2x + 5y - 4z &= 13 \end{aligned}$$



We write the equations as an array, and perform on the array certain operations which we denote as follows: $R_i + kR_j$ indicates that we are forming a new array in which each row is the same as the corresponding row preceding array, except for row i, which is the sum of row i and k times row j in the preceding array. (This corresponds to the operation of subtraction some multiple of the j^{th} equation from the i^{th} equation to obtain an equation in which the coefficients of some unknown is zero.)

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 1 & 3 & 1 & 11 \\ 2 & 5 & -4 & 13 \end{array}\right)$$

$$\frac{R_2 - R_1}{R_3 - 2R_1} \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 1 & 2 & 5 \end{array}\right)$$

$$\underline{R_3 - R_2} \left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & 4 & 7 \\ 0 & 0 & -2 & -2 \end{array}\right)$$

This array corresponds to the system of equations

$$\begin{aligned} x + 2y - 3z &= 4 \\ y + 4z &= 7 \\ -2z &= -2 \end{aligned}$$

The system has the same solution as the original system, and is easier to solve: we can find the solution directly from the array, by what is called '*back substitution*'. From row 3, we have $z = 1$; from row 2, we have $y = 7 - 4z = 3$; and from row 1, we have $x = 4 - 2y + 3z = 1$. The solution is $x = 1, y = 3, z = 1$.

This method of solving linear equations is known as *row reduction of a matrix to echelon form.* I what follows we explain what is meant by the term *matrix*, and define some operations, such as addition and multiplication, that can be performed on matrices. We will consider row reduction to echelon form later.

1.2 Matrix Definition

Definition 1 *A matrix is a rectangular array of numbers of the form*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{pmatrix}$$

This is written $A = (a_{ij})$ for short.

The matrix A has m rows.

$$\begin{aligned} row\ 1 &= (a_{11}\ a_{12}\ \ldots\ldots\ldots\ldots\ a_{1n}) \\ row\ 2 &= (a_{21}\ a_{22}\ \ldots\ldots\ldots\ldots\ a_{2n}) \\ \vdots & \quad \vdots \quad\quad \vdots \quad\quad \vdots \\ row\ m &= (a_{m1}\ a_{m2}\ \ldots\ldots\ldots\ldots\ a_{mn}) \end{aligned}$$

The matrix A has n columns.

$$column\ 1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad column\ 2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \ldots\ldots \quad column\ n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

The number $a_{i,j}$ is called the i, j - entry of A; it occurs in row i, column j.

Definition 2 *Let $A = (a_{i,j})$. The dimension or size of A is denoted by $m \times n$ where m is the number of rows and n is the number of columns.*

Example The matrix

$$A = \begin{pmatrix} 1 & 2 & -3 \\ -7 & 3 & 0 \end{pmatrix}$$

is a 2×3 matrix; it has 2 rows and 3 columns. The entry at position a_{21} is -7. The entry at position a_{13} is -3.

Example The matrix

$$A = \begin{pmatrix} -9 & 5 & -2 \\ 3 & 1 & 5 \\ 7 & 6 & -3 \end{pmatrix}$$

is a 3×3 matrix; it has 3 rows and 3 columns. The entry at position a_{23} is 5. The entry at position a_{13} is -2.

Remark Any vector (a_1, a_2, a_3) in $\mathbb{R}^3$ is a 1×3 matrix. For this reason, a $1 \times n$ matrix or *n–tuple*

$$(a_1 \ a_2 \ .... \ a_n)$$

is often simply called a *row vector*. Similarly, an $m \times 1$ matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

is called a column vector. Many of the properties of vectors in $\mathbb{R}^3$ hold also for these more general row and column vectors. For example, we could define, in the obvious way, the dot product of two row vectors, or of two column vectors, or indeed of a row vector and a column vector, provided that in each case both vectors have the same number of components. Thus, if

$$A = (a_1 \ a_2 \ .... \ a_n)$$

and

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

then the dot product of A and B is the scalar $A.B = a_1b_1 + a_2b_2 + ... + a_nb_n$.

We could then proceed to consider what is meant by the "angle" between such vectors, when they are orthogonal, and so on. This is presented in the section on vectors. The definitions that will be given for addition and multiplication of matrices apply in particular to vectors with more that three components. Properties of matrix addition and multiplication are the same as those of such vectors.

Definition 3 *The $m \times n$ zero matrix is the $m \times n$ matrix with every entry 0. It is denoted by 0.*

Definition 4 *The $n \times n$ identity matrix is the $n \times n$ matrix*

$$I_n = (\delta_{i,j}), \qquad where \qquad \delta_{i,j} = \begin{cases} 1 & if\ i = j \\ 0 & if\ i \neq j \end{cases}$$

So, for example,

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The identity matrix is written simply as I.

Definition 5 *The square matrix is one with the same number of rows as columns.*

Definition 6 *The diagonal of a square matrix A is the set of all elements $a_{11}, a_{22}, ..., a_{nn}$.*

Definition 7 *A diagonal matrix is a square matrix with every non-diagonal entry zero.*

Thus, for a square matrix $A = (a_{ij})$, A is diagonal if and only if $a_{ij} = 0$ when $i \neq j$. In particular, the $n \times n$ zero matrix is a diagonal matrix; so also is I.

1.3 Operations on Matrices

Let A and B be matrices with $A = (a_{ij})$ and $B = (b_{ij})$.

We will specify the dimensions of A and B when it is necessary to do so.

Definition 8 (Equality)

The matrices A and B are equal when they are of the same dimension and their corresponding entries are equal; thus

$$A = B \Leftrightarrow \begin{cases} \#\ \text{rows of } A = \#\ \text{rows of } B \\ \#\ \text{columns of } A = \#\ \text{columns of } B \\ a_{ij} = b_{ij}, \quad \forall i, j. \end{cases}$$

Definition 9 (Addition)

The addition of A and B, denoted by $A+B$, is defined, only when A and B have the same dimension, by the equation

$$A + B = (a_{ij} + b_{ij})$$

Example If

$$A = \begin{pmatrix} 1 & 2 & -3 \\ -7 & 3 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -3 & -2 & 9 \\ 5 & -3 & 4 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} -2 & 0 & 6 \\ -2 & 0 & 4 \end{pmatrix}$$

Remark The following properties hold – for matrices A, B, C all of the same size,

$$\begin{aligned} (A + B) + C &= A + (B + C) \\ A + B &= B + A \\ A + 0 &= A \ = \ 0 + A \end{aligned}$$

These are proved by establishing equality of the ij - entry. For example, to prove the second property, that matrix addition is commutative, we note that the ij^{th} entry of $A + B$ is $a_{ij} + b_{ij}$, the ij^{th} entry of $B + A$ is $b_{ij} + a_{ij}$ and since addition of real numbers is commutative

$$a_{ij} + b_{ij} = b_{ij} + a_{ij}$$

Definition 10 (Scalar Multiplication)

For any scalar k (that is, any real or complex number k), the scalar multiple of k and A is the matrix kA defined by

$$kA = (ka_{ij})$$

Example Let

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 5 \\ 3 & 1 \end{pmatrix}$$

Then

$$2A = \begin{pmatrix} 4 & 0 \\ 2 & 10 \\ 6 & 2 \end{pmatrix} \quad , \quad \frac{1}{2}A = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{5}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix} \quad , \quad (-1)A = \begin{pmatrix} -2 & 0 \\ -1 & -5 \\ -3 & -1 \end{pmatrix}$$

Note We will write $(-1)A$ as $-A$ and $A + (-B)$ as $A - B$; thus

$$\begin{aligned} -A &= (-a_{ij}) \\ A - B &= (a_{ij} - b_{ij}) \end{aligned}$$

Example Let

$$A = \begin{pmatrix} 1 & 2 \\ -3 & -5 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 2 & -7 \\ -4 & 3 \end{pmatrix}$$

Then

$$-B = \begin{pmatrix} -2 & 7 \\ 4 & -3 \end{pmatrix} \quad , \quad A - B = \begin{pmatrix} -1 & 9 \\ 1 & -8 \end{pmatrix}$$

Remark Properties similar to those of 'ordinary' multiplication hold also for scalar multiplication: for matrices A and B of the same size, and scalars k, k_1, k_2:

$$\begin{aligned} k(A + B) &= kA + kB \\ (k_1 k_2)A &= k_1(k_2 A) \\ (k_1 + k_2)A &= k_1 A + k_2 A \\ A + (-A) &= 0 = (-A) + A \\ 1A &= A \\ 0A &= 0 \end{aligned}$$

In the last equality the zero on the left-hand side denotes the real number zero whereas on the the right-hand side the zero is the zero matrix. The properties listed above are the same properties that hold when A and B are vectors in $\mathbb{R}^3$.

Definition 11 (Transposition)

The transpose of an $m \times n$ matrix A is the $n \times m$ matrix A^t whose ij^{th} entry is the ji^{th} entry of A.

Example Let

$$A = \begin{pmatrix} 1 & -3 & 7 \\ 5 & 7 & -4 \end{pmatrix}$$

Then

$$A^t = \begin{pmatrix} 1 & 5 \\ -3 & 7 \\ 7 & -4 \end{pmatrix}$$

Remark The matrix A is transposed, and we obtain the matrix A^t, simply by writing the rows of A as columns. Transposition has the following properties

$$\begin{aligned} (A^t)^t &= A \\ (A+B)^t &= A^t + B^t \\ (AB)^t &= B^t A^t \end{aligned}$$

The definitions of addition and scalar multiplication of matrices are the 'obvious' ones, particularly in view of the corresponding definitions for vectors. The definition of matrix multiplication is not an obvious one.

The product AB of two matrices is defined only when the number of columns of A is equal to the number of rows of B.

Definition 12 (Matrix Multiplication)

Let A be of dimension $m \times n$, let B be of dimension $n \times p$. Then AB is the $m \times p$ matrix whose ij^{th} entry is

$$\sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \dots\dots\dots + a_{in} b_{nj}$$

that is, the ij^{th} entry of AB is the dot product of row i of A and column j of B.

Example Let

$$A = \begin{pmatrix} 1 & -3 \\ 5 & 7 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 2 & 4 \\ -6 & 8 \end{pmatrix}$$

Then

$$\begin{aligned} AB &= \begin{pmatrix} 1 & -3 \\ 5 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 & 4 \\ -6 & 8 \end{pmatrix} \\ &= \begin{pmatrix} 2+18 & 4-24 \\ 10-42 & 20+56 \end{pmatrix} \\ &= \begin{pmatrix} 20 & -20 \\ -32 & 76 \end{pmatrix} \end{aligned}$$

Also

$$\begin{aligned} BA &= \begin{pmatrix} 2 & 4 \\ -6 & 8 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 5 & 7 \end{pmatrix} \\ &= \begin{pmatrix} 2+20 & -6+28 \\ -6+40 & 18+56 \end{pmatrix} \\ &= \begin{pmatrix} 22 & 22 \\ 34 & 74 \end{pmatrix} \end{aligned}$$

In this example, $AB \neq BA$. It follows that matrix multiplication is **not** commutative. Thus multiplication of matrices differs in a very fundamental way from multiplication of real or complex numbers, in that the order in which we write two matrices which are to be multiplied together, as AB or as BA, **may** determine different products. There are some matrices that do commute with each other – for example

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

In general, however, we cannot assume that $AB = BA$, and we must take care to distinguish AB and BA.

Example Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -1 & -2 \\ -3 & -4 \end{pmatrix}$$

Then

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} -1 & -2 \\ -3 & -4 \end{pmatrix} \\ &= \begin{pmatrix} -1-6 & -2-8 \\ -3-12 & -6-16 \\ -5-18 & -10-24 \end{pmatrix} \\ &= \begin{pmatrix} -7 & -10 \\ -15 & -22 \\ -23 & -34 \end{pmatrix} \end{aligned}$$

Notice that the matrix BA is not defined.

Remark Matrix multiplication, though not commutative, has many of the properties of 'ordinary' multiplication of real numbers. For matrices A, B, C and identity matrix I of the appropriate size for the product to be defined, and scalar k we have the following:

$$\begin{aligned}(AB)C &= A(BC)\\ A(B+C) &= AB+AC\\ (B+C)A &= BA+CA\\ k(AB) &= (kA)B = A(kB)\\ AI &= A = IA\end{aligned}$$

Exercise Let

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 4 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 4 & 0 & -3 \\ -1 & -2 & 3 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 2 & -3 & 0 & 1 \\ 5 & -1 & -4 & 2 \\ -1 & 0 & 0 & 3 \end{pmatrix}$$

Determine each of the following (where defined)

$$A+B \qquad A+C \qquad 2A.C \qquad A^tB \qquad AB^t$$

Exercise Let

$$A = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \\ 1 & -1 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 3 \\ 4 & 0 & -1 \end{pmatrix}$$

Determine each of the following (where defined)

$$AB \qquad BA \qquad B^tA \qquad A^tB \qquad 2B^t+C \qquad C^t-2C$$

1.4 Solving Systems of Linear Equations

In general the following system of m linear equations in n unknowns $x_1, x_2, ..., x_n$

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

may be represented in matrix form as

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & ..... & a_{1m} & b_1 \\ a_{21} & a_{22} & ..... & a_{2m} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & ..... & a_{mn} & b_m \end{array}\right)$$

Example The system of linear equations:

$$\begin{aligned} x + 3y &= 5 \\ 2x + y &= 2 \end{aligned}$$

may be represented as

$$\left(\begin{array}{cc|c} 1 & 3 & 5 \\ 2 & 1 & 2 \end{array}\right)$$

Example The system of linear equations:

$$\begin{aligned} x + 2y - 3z &= 4 \\ x + 3y + z &= 11 \\ 2x + 5y - 4z &= 13 \end{aligned}$$

may be represented as

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 1 & 3 & 1 & 11 \\ 2 & 5 & -4 & 13 \end{array}\right)$$

Note

To solve such systems of linear equations we will use the technique of *row reduction to echelon form.* We begin by defining a row echelon matrix and row operations.

Definition 13

A matrix $A = (a_{ij})$ *is in row echelon form (REF), or is an row echelon matrix, if the number of zeros preceding the first non-zero entry of a row increases row by row until only zero rows remain: that is, there are non-zero entries*

$$a_{1j_1}, a_{2j_2}, ..., a_{rj_r},$$

such that

$$j_1 < j_2 < .... < j_r;$$

when $i = 1, 2, ..., r, a_{ij} = 0$ *for all* $j < j_i$*; when* $i > r,\ a_{ij} = 0$ *for all* j.

Note The entries $a_{1j_1}, a_{2j_2}, ..., a_{rj_r}$, are called the *distinguished entries.*

The following illustrations will help in our understanding:

A

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

This matrix is in *row echelon form* (REF). It is a *row echelon matrix.*
The distinguished entries are $a_{11} = 2, a_{22} = 1$.

B

$$\begin{pmatrix} 0 & 2 & 3 & 0 & -1 & 5 \\ 0 & 0 & 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix.*
The distinguished entries are $a_{12} = 2, a_{25} = 5$.

C

$$\begin{pmatrix} 0 & 3 & -1 & 4 & 0 \\ 0 & 0 & 7 & 3 & 1 \\ 0 & 0 & 0 & 1 & -5 \end{pmatrix}$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix.*
The distinguished entries are $a_{12} = 3, a_{23} = 7, a_{34} = 1$.

D

$$\begin{pmatrix} 1 & 2 & 3 & -6 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

This matrix is **not** in *row echelon form* (REF). It is **not** a *row echelon matrix.*

E

$$\begin{pmatrix} 1 & 2 & 3 & -6 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

This matrix is **not** in *row echelon form* (REF). It is **not** a *row echelon matrix.*

Definition 14

A matrix $A = (a_{ij})$ is in row–reduced echelon form (RREF), or a reduced echelon matrix, if

i A is an echelon matrix;

ii the distinguished entries are all 1;

iii the distinguished entries are the only non–zero entries in their columns.

Consider the following illustrations:

A

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -5 \end{pmatrix}$$

This matrix is in *row–reduced echelon form* (RREF). It is a *reduced echelon matrix.*

B

$$\begin{pmatrix} 0 & 1 & 3 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This matrix is in *row–reduced echelon form* (RREF). It is a *reduced echelon matrix.*

C

$$\begin{pmatrix} 0 & 1 & 3 & 6 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This matrix is **not** in *row–reduced echelon form* (RREF). It is **not** a *reduced echelon matrix.*

D

$$\begin{pmatrix} 0 & 2 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This matrix is **not** in *row–reduced echelon form* (RREF). It is **not** a *reduced echelon matrix.*

1.5 Row Operations

There are certain operations we perform on a matrix A in order to bring the matrix A to *row echelon form* (or to *row-reduced echelon form*). These operations are called *elementary row operations*.

Firstly we form the $n \times 2n$ matrix $[A|I_n]$ to obtain the row equivalent *row-reduced echelon matrix* $[I_n|B]$. We use the notation $R_1, R_2, ..., R_n$ to label each row.

Definition 15

Let k be a non–zero scalar. The following are the valid row operations:

i Add k times R_j to R_i: i.e., $R_i + kR_j$

ii Interchange R_i and R_j: i.e., $R_i \leftrightarrow R_j$

iii Multiply R_i by a non–zero scalar k: i.e., kR_i.

Note For the first and third of the elementary row operations, $R_i{+}kR_j$ and kR_i, the only change occurs in row i of the matrix – all the other rows remain unchanged. Under the second operation, $R_i \leftrightarrow kR_j$, changes occur in row i and in row j, and in no other row. One further definition is required. A matrix B is *row equivalent* to a matrix A if B can be obtained from A by a finite sequence of elementary row operations. Each row operation is reversible:

i $R_i + kR_j$ is reversed by $R_i - kR_j$

ii $R_i \leftrightarrow R_j$ is reversed by $R_j \leftrightarrow R_i$

iii kR_i is reversed by $\frac{1}{k}R_i$

If B is row equivalent to A, then A is row equivalent to B, so we can say A and B are row equivalent.

Theorem 1

Any matrix is row equivalent to an echelon matrix, and to a reduced echelon matrix.

Example Consider the following matrix A

$$A = \begin{pmatrix} 1 & 2 & -3 & 5 \\ 3 & -1 & 5 & 1 \end{pmatrix}$$

Using a sequence of valid row operations we can reduce this matrix to an equivalent *row echelon form* (REF) and *row-reduced echelon form* (RREF).

$$\underline{R_2 - 3R_1} \begin{pmatrix} 1 & 2 & -3 & 5 \\ 0 & -7 & 14 & -14 \end{pmatrix}$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix* row equivalent to A. Also

$$\underline{R_2 \times -\tfrac{1}{7}} \begin{pmatrix} 1 & 2 & -3 & 5 \\ 0 & 1 & -2 & 2 \end{pmatrix}$$

$$\underline{R_1 - 2R_2} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -2 & 2 \end{pmatrix}$$

This matrix is in *row–reduced echelon form* (RREF). It is a *reduced echelon matrix* row equivalent to A.

Example To solve the following system of linear equations:

$$\begin{aligned} x - y &= 3 \\ 2x + y &= 0 \end{aligned}$$

we form the following matrix A

$$A = \left(\begin{array}{cc|c} 1 & -1 & 3 \\ 2 & 1 & 0 \end{array}\right)$$

Using a sequence of valid row operations we can reduce this matrix to an equivalent *row echelon form* (REF) and *row-reduced echelon form* (RREF).

$$\underline{R_2 - 2R_1} \left(\begin{array}{cc|c} 1 & -1 & 3 \\ 0 & 3 & -6 \end{array}\right)$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix* row equivalent to A.

We could apply *back substitution* at this stage to yield the solution.

Alternatively, we can proceed to row-reduced echelon form to read off the solution.

$$\xrightarrow{R_2 \times \frac{1}{3}} \left(\begin{array}{cc|c} 1 & -1 & 3 \\ 0 & 1 & -2 \end{array} \right)$$

$$\xrightarrow{R_1 + R_2} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -2 \end{array} \right)$$

This matrix is in *row–reduced echelon form* (RREF).

It is a *reduced echelon matrix* row equivalent to A.

The required solution to the system of linear equations is $x = 1, y = -2$.

Example To solve the following system of linear equations:

$$\begin{aligned} 3x + y &= 1750 \\ 2x + 4y &= 4000 \end{aligned}$$

we form the following matrix A

$$A = \left(\begin{array}{cc|c} 3 & 1 & 1750 \\ 2 & 4 & 4000 \end{array} \right)$$

Using a sequence of valid row operations we can reduce this matrix to an equivalent *row echelon form* (REF) and *row-reduced echelon form* (RREF).

$$\xrightarrow{R_1 - R_2} \left(\begin{array}{cc|c} 1 & -3 & -2250 \\ 2 & 4 & 4000 \end{array} \right)$$

$$\xrightarrow{R_2 - 2R_1} \left(\begin{array}{cc|c} 1 & -3 & -2250 \\ 0 & 10 & 8500 \end{array} \right)$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix* row equivalent to A.

We could apply *back substitution* at this stage to yield the solution.

Alternatively, we can proceed to row-reduced echelon form to read off the solution.

$$\xrightarrow{R_2 \times \frac{1}{10}} \left(\begin{array}{cc|c} 1 & -3 & -2250 \\ 0 & 1 & 850 \end{array} \right)$$

$$\xrightarrow{R_1+3R_2} \left(\begin{array}{cc|c} 1 & 0 & 300 \\ 0 & 1 & 850 \end{array}\right)$$

This matrix is in *row–reduced echelon form* (RREF).

It is a *reduced echelon matrix* row equivalent to A.

The required solution to the system of linear equations is $x = 300, y = 850$.

Example To solve the following system of linear equations:

$$\begin{aligned} 2x - 3y - 4z &= -3 \\ 3x + 5y + 2z &= 2 \\ -4x + 7y + 7z &= 3 \end{aligned}$$

we form the following matrix

$$A = \left(\begin{array}{ccc|c} 2 & -3 & -4 & -3 \\ 3 & 5 & 2 & 2 \\ -4 & 7 & 7 & 3 \end{array}\right)$$

Using a sequence of valid row operations we can reduce this matrix to an equivalent *row echelon form* (REF) and *row-reduced echelon form* (RREF).

$$\xrightarrow{2R_2} \left(\begin{array}{ccc|c} 2 & -3 & -4 & -3 \\ 6 & 10 & 4 & 4 \\ -4 & 7 & 7 & 3 \end{array}\right)$$

$$\xrightarrow[R_3+2R_1]{R_2-3R_1} \left(\begin{array}{ccc|c} 2 & -3 & -4 & -3 \\ 0 & 19 & 16 & 13 \\ 0 & 1 & -1 & -3 \end{array}\right)$$

$$\xrightarrow{R_2 \leftrightarrow R_1} \left(\begin{array}{ccc|c} 2 & -3 & -4 & -3 \\ 0 & 1 & -1 & -3 \\ 0 & 19 & 16 & 13 \end{array}\right)$$

$$\xrightarrow{R_3-19R_2} \left(\begin{array}{cccc} 2 & -3 & -4 & -3 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 35 & 70 \end{array}\right)$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix* row equivalent to A.

We could apply *back substitution* at this stage to yield the solution.

Alternatively, we can proceed to row-reduced echelon form to read off the solution.

$$\underline{R_3 \times \frac{1}{35}} \begin{pmatrix} 2 & -3 & -4 & -3 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\underline{\begin{matrix} R_1 + 4R_3 \\ R_2 + R_3 \end{matrix}} \begin{pmatrix} 2 & -3 & 0 & 5 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\underline{R_1 + 3R_2} \begin{pmatrix} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\underline{R_1 \times \frac{1}{2}} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

This matrix is in *row–reduced echelon form* (RREF).

It is a *reduced echelon matrix* row equivalent to A.

The required solution to the system of linear equations is $x = 1, y = -1, z = 2$.

Remark

There is a great degree of **flexibility** in this method of row reduction. We have a choice of possible row operations to use and hence there are many different ways in which a matrix may be reduced to echelon form. As a consequence there are infinitely many echelon matrices all row equivalent to any given matrix, however, there is a unique reduced echelon matrix row equivalent to any given matrix. There are some observations that can be made at this stage which may be used as general guidelines.

- Using appropriate row operation(s) bring the distinguished entry in R_1 to 1.
- Entries below this distinguished entry can easily be brought to 0 since the distinguished entry in R_1 is 1.
- Using appropriate row operation(s) bring the distinguished entry in R_2 to 1.
- Entries below this distinguished entry can easily be brought to 0 since the distinguished entry in R_2 is 1.

These guidelines were loosely adhered to in the above example – however, where possible, we will try and follow this routine. A good general rule in row reduction is to avoid fractions since they are very tedious to work with and the chance of making simple errors increase. Whether we are row reducing on paper or on a computer the aim is to increase the number of zeros in the matrix in an economical and systematic way. According to the method outlined we first work from left to right in the matrix to produce zeros beneath a distinguished entry in each non–zero column and so obtain an echelon matrix and if a reduced echelon matrix is required we then work back from right to left to produce zeros above each distinguished entry.

Row reduction to echelon form is a basic computational technique in linear algebra. A good exercise would be to write a computer program, in any language, to row reduce a matrix to echelon form. This will test your understanding of the method and if your program works you will be able to use it in many contexts to do calculations which, though straightforward, are tedious to do on paper. We will see how this technique can be used to invert matrices.

Example To solve the following system of linear equations

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

we form the following matrix

$$A = \left(\begin{array}{ccc|c} 3 & 5 & -1 & 3 \\ 1 & 0 & 1 & 5 \\ -1 & -1 & 2 & 4 \end{array}\right)$$

Using a sequence of valid row operations we can reduce this matrix to an equivalent *row echelon form* (REF) and *row-reduced echelon form* (RREF).

$$\underline{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 3 & 5 & -1 & 3 \\ -1 & -1 & 2 & 4 \end{array}\right)$$

$$\frac{R_2 - 3R_1}{R_3 + R_1} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 5 & -4 & -12 \\ 0 & -1 & 3 & 9 \end{array}\right)$$

$$\underline{R_2 \leftrightarrow R_3} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & -1 & 3 & 9 \\ 0 & 5 & -4 & -12 \end{array}\right)$$

$$\xrightarrow{R_3 + 5R_2} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & -1 & 3 & 9 \\ 0 & 0 & 11 & 33 \end{array}\right)$$

This matrix is in *row echelon form* (REF). It is a *row echelon matrix* row equivalent to A.

We could apply *back substitution* at this stage to yield the solution.

Alternatively, we can proceed to row-reduced echelon form to read off the solution.

$$\xrightarrow{R_3 \times \frac{1}{11}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & -1 & 3 & 9 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\xrightarrow[R_1 - R_3]{R_2 - 3R_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\xrightarrow{R_2 \times -1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

This matrix is in *row–reduced echelon form* (RREF).

It is a *reduced echelon matrix* row equivalent to A.

The required solution to the system of linear equations is $x = 2, y = 0, z = 3$.

Exercise

Solve, by *row reduction*, the system of linear equations.

$$\begin{aligned} x &= -1 \\ 2x + 2y - z &= 5 \\ x - y + z &= 3. \end{aligned}$$

Exercise

Solve, by *row reduction*, the system of linear equations.

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

1.6 The Inverse of a Matrix

Definition 16

A square matrix A is invertible if there is a matrix B such that

$$AB = I = BA$$

Remark

i Such a matrix B if it exists, is uniquely determined by A. The unique matrix B such that $AB = I = BA$ is called the *inverse of* A, and it is denoted by A^{-1}. Thus

$$AA^{-1} = I = A^{-1}A$$

ii A matrix may not necessarily have an inverse. For example, there is no matrix B such that

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} B = I = B \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

iii We will first consider using valid *row operations* to determine the inverse of a square matrix A. We simply apply row operations to the $n \times 2n$ matrix $[A|I_n]$ to obtain the row equivalent *row-reduced echelon matrix* $[I_n|B]$. If this is successful then A is invertible, and $A^{-1} = B$.

$$\left[A \middle| I_n\right] \longrightarrow \left[I_n \middle| A^{-1}\right]$$

Row Operations

Example To find the inverse of

$$A = \begin{pmatrix} 1 & -2 \\ 5 & -7 \end{pmatrix}$$

Firstly, we form the following matrix

$$\left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 5 & -7 & 0 & 1 \end{array}\right)$$

We now apply row operations to this 2×4 matrix $[A|I_n]$ to obtain the row equivalent *row-reduced echelon matrix* $[I_n|A^{-1}]$

$$\xrightarrow{R_2 - 5R_1} \left(\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 3 & -5 & 1 \end{array}\right)$$

$$\xrightarrow{3R_1} \left(\begin{array}{cc|cc} 3 & -6 & 3 & 0 \\ 0 & 3 & -5 & 1 \end{array}\right)$$

$$\xrightarrow{R_1 + 2R_2} \left(\begin{array}{cc|cc} 3 & 0 & -7 & 2 \\ 0 & 3 & -5 & 1 \end{array}\right)$$

$$\xrightarrow[(\frac{1}{3})R_2]{(\frac{1}{3})R_1} \left(\begin{array}{cc|cc} 1 & 0 & -\frac{7}{3} & \frac{2}{3} \\ 0 & 1 & -\frac{5}{3} & \frac{1}{3} \end{array}\right)$$

Therefore, we have

$$A^{-1} = \begin{pmatrix} -\frac{7}{3} & \frac{2}{3} \\ -\frac{5}{3} & \frac{1}{3} \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -7 & 2 \\ -5 & 1 \end{pmatrix}$$

Exercise Find the inverse of the matrix

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix}$$

Exercise Find the inverse of the matrix

$$A = \begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}$$

Example Find the inverse of

$$A = \begin{pmatrix} -5 & 1 & 4 \\ -1 & 1 & 1 \\ -4 & 1 & 3 \end{pmatrix}$$

We form the following matrix

$$\left(\begin{array}{ccc|ccc} -5 & 1 & 4 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 1 & 0 \\ -4 & 1 & 3 & 0 & 0 & 1 \end{array}\right)$$

We now apply row operations to this 3×6 matrix $[A|I_n]$ to obtain the row equivalent *row-reduced echelon matrix* $[I_n|A^{-1}]$

$$\frac{R_1 \leftrightarrow R_2}{(-1)R_1} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 0 & -1 & 0 \\ -5 & 1 & 4 & 1 & 0 & 0 \\ -4 & 1 & 3 & 0 & 0 & 1 \end{array}\right)$$

$$\frac{R_2 + 5R_1}{R_3 + 4R_1} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 0 & -1 & 0 \\ 0 & -4 & -1 & 1 & -5 & 0 \\ 0 & -3 & -1 & 0 & -4 & 1 \end{array}\right)$$

$$\underline{R_2 - R_3} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 & -1 & -1 \\ 0 & -3 & -1 & 0 & -4 & 1 \end{array}\right)$$

$$\underline{R_3 - 3R_2} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 & -1 & -1 \\ 0 & 0 & -1 & -3 & -1 & 4 \end{array}\right)$$

$$\frac{(-1)R_2}{(-1)R_3} \left(\begin{array}{ccc|ccc} 1 & -1 & -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 3 & 1 & -4 \end{array}\right)$$

$$\underline{R_1 + R_3} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 3 & 0 & -4 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 3 & 1 & -4 \end{array}\right)$$

$$\underline{R_1 + R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & -3 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 3 & 1 & -4 \end{array}\right)$$

Therefore, we have

$$A^{-1} = \begin{pmatrix} 2 & 1 & -3 \\ -1 & 1 & 1 \\ 3 & 1 & -4 \end{pmatrix}$$

We can check this answer by matrix multiplication:

$$A.A^{-1} = \begin{pmatrix} -5 & 1 & 4 \\ -1 & 1 & 1 \\ -4 & 1 & 3 \end{pmatrix} . \begin{pmatrix} 2 & 1 & -3 \\ -1 & 1 & 1 \\ 3 & 1 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Example Find the inverse of

$$A = \begin{pmatrix} 6 & -4 & -7 \\ 3 & -2 & -3 \\ 1 & 0 & -2 \end{pmatrix}$$

We form the following matrix

$$\left(\begin{array}{ccc|ccc} 6 & -4 & -7 & 1 & 0 & 0 \\ 3 & -2 & -3 & 0 & 1 & 0 \\ 1 & 0 & -2 & 0 & 0 & 1 \end{array}\right)$$

We now apply row operations to this 3×6 matrix $[A|I_n]$ to obtain the row equivalent *row-reduced echelon matrix* $[I_n|A^{-1}]$

$$\underline{R_1 \leftrightarrow R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & 0 & 1 \\ 3 & -2 & -3 & 0 & 1 & 0 \\ 6 & -4 & -7 & 1 & 0 & 0 \end{array}\right)$$

$$\frac{R_2 - 3R_1}{R_3 - 6R_1} \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & 0 & 1 \\ 0 & -2 & 3 & 0 & 1 & -3 \\ 0 & -4 & 5 & 0 & 0 & -6 \end{array}\right)$$

$$\underline{R_3 - 2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & 0 & 1 \\ 0 & -2 & 3 & 0 & 1 & -3 \\ 0 & 0 & -1 & 1 & -2 & 0 \end{array}\right)$$

$$\frac{R_2 + 3R_3}{R_1 - 2R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 4 & 1 \\ 0 & -2 & 0 & 3 & -5 & -3 \\ 0 & 0 & -1 & 1 & -2 & 0 \end{array}\right)$$

$$\frac{(-\frac{1}{2})R_2}{(-1)R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 4 & 1 \\ 0 & 1 & 0 & -\frac{3}{2} & \frac{5}{2} & \frac{3}{2} \\ 0 & 0 & 1 & -1 & 2 & 0 \end{array}\right)$$

Therefore, we have

$$A^{-1} = \frac{1}{2} \begin{pmatrix} -4 & 8 & 2 \\ -3 & 5 & 3 \\ -2 & 4 & 0 \end{pmatrix}$$

We can check this answer by matrix multiplication:

$$A^{-1}.A = \frac{1}{2}\begin{pmatrix} -4 & 8 & 2 \\ -3 & 5 & 3 \\ -2 & 4 & 0 \end{pmatrix} . \begin{pmatrix} 6 & -4 & -7 \\ 3 & -2 & -3 \\ 1 & 0 & -2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Exercise Find the inverse of

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & 6 & 3 \\ 2 & -4 & 0 \end{pmatrix}$$

Remark

For a system of n linear equations in n unknowns represented as $Ax = B$, we can now use the inverse A^{-1} to solve the system using matrix multiplication as follows.

$$\begin{aligned} A^{-1}Ax &= A^{-1}B \\ \Rightarrow \quad Ix &= A^{-1}B \\ \therefore \quad x &= A^{-1}B \end{aligned}$$

Exercise Let

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

By a sequence of *row operations* find A^{-1}, the *inverse* of A. Use this inverse to find the solution of the following system of linear equations.

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

1.7 Determinants

A 'determinant' is a certain kind of function that associates a real number with a square matrix. The 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible if $ad - bc \neq 0$. The expression $ad - bc$ occurs so frequently in mathematics that it is called the **determinant of the matrix** A. It is denoted by the the symbol $det(A)$ or $|A|$. With this notation, the formula for the inverse of A, i.e., A^{-1} is given as

$$A^{-1} = \frac{1}{det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

where $det(A) = ad - bc$.

Remark

This is a formula for the inverse of a 2×2 matrix. It arises naturally from another method of finding inverses, by using what is called *determinants*. We remark that there is little to choose between the methods for "small" matrices – we will find the unique inverse of a given 3×3 or 4×4 matrix in more or less the same amount of time, whichever method we follow. For larger matrices, however, the method which uses row reduction is much quicker than that which uses determinants.