Linear Algebra: Lecture Notes

Dr Rachel Quinlan
School of Mathematics, Statistics and Applied Mathematics
NUI Galway

November 6, 2013

Contents

Chapter 1

Systems of Linear Equations

1.1 Introduction

Consider the equation

$$2x + y = 3.$$

This is an example of a *linear equation* in the variables x and y. As it stands, the statement "$2x + y = 3$" is neither true nor untrue : it is just a statement involving the abstract symbols x and y. However if we replace x and y with some particular pair of real numbers, the statement will become either true or false. For example

$$\begin{aligned} \text{Putting } & x = 1,\ y = 1 \text{ gives } 2x + y = 2(1) + (1) = 3: \text{ True} \\ & x = 1,\ y = 2 \text{ gives } 2x + y = 2(1) + (2) \neq 3: \text{ False} \\ & x = 0,\ y = 3 \text{ gives } 2x + y = 2(0) + (3) = 3: \text{ True} \end{aligned}$$

Definition 1.1.1 *A pair (x_0, y_0) of real numbers is a* solution *to the equation $2x+y = 3$ if setting $x = x_0$ and $y = y_0$ makes the equation true; i,.e. if $2x_0 + y_0 = 3$.*

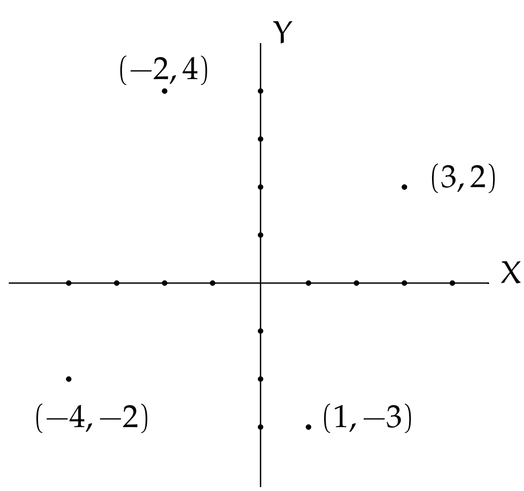
For example $(1, 1)$ and $(0, 3)$ are solutions - so are $(2, -1)$, $(3, -3)$, $(-1, 5)$ and $(-1/2, 4)$ (check these).

However $(1, 4)$ is not a solution since setting $x = 1$, $y = 4$ gives $2x + y = 2(1) + 4 \neq 3$.

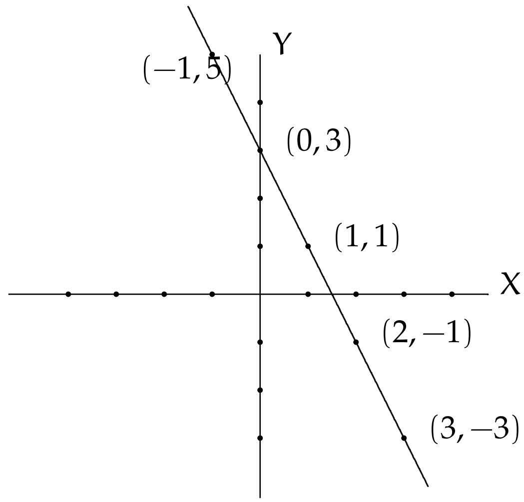
The set of all solutions to the equation is called its *solution set*.

GEOMETRIC INTERPRETATION

RECALL: The Cartesian Coordinate System. The 2-dimensional plane is described by a pair of perpendicular axes, labelled X and Y. A point is described by a pair of real numbers, its X and Y-coordinates.



We plot in the plane those points which correspond to the pairs of numbers which we found to be solutions to the equation $2x + y = 3$. These points form a *line* in the plane.



Now consider the equation $4x + 3y = 4$. Solutions to this equation include

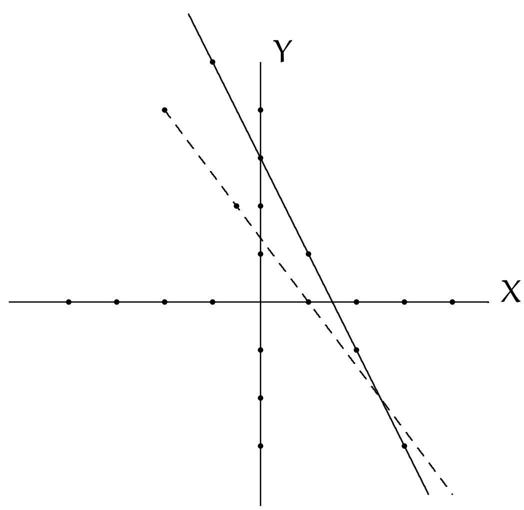
$$(1,0),\ (4,-4),\ (-2,4),\ (-1/2,2).$$

Again the full solution set forms a line.
Question: Consider the equations

$$2x + y = 3, \quad 4x + 3y = 4$$

together. Can we find simultaneous solutions of these equations? This means - can we find pairs of numbers (x_0, y_0) such that setting $x = x_0$ and $y = y_0$ makes *both* equations true?

Equivalently - can we find a point of intersection of the two lines? From the picture it looks as if there is exactly one such point, at $(5/2, -2)$.



We can solve the problem algebraically as follows :

$$\left.\begin{array}{rcrcrl} 2x & + & y & = & 3 & (A) \\ 4x & + & 3y & = & 4 & (B) \end{array}\right\} \text{A system of linear equations.}$$

Step 1: Multiply Equation (A) by 2 : $4x + 2y = 6$ (A2).
Any solution of (A2) is a solution of (A).

Step 2: Multiply Equation (B) by -1 : $-4x - 3y = -4$ (B2)
Any solution of (B2) is a solution of (B).

Step 3: Now add equations (A2) and (B2).

$$\begin{array}{rcrcr} 4x & + & 2y & = & 6 \\ -4x & - & 3y & = & -4 \\ \hline \\ & & -y & = & 2 \end{array}$$

Step 4: So $y = -2$ and the value of y in any simultaneous solution of (A) and (B) is -2 : Now we can use (A) to find the value of x.

$$\begin{aligned} 2x + y = 3 \text{ and } y = -2 \Longrightarrow 2x + (-2) &= 3 \\ \Longrightarrow 2x &= 5 \\ \Longrightarrow x &= \frac{5}{2} \end{aligned}$$

So $x = 5/2,\ y = -2$ is the *unique* solution to this system of linear equations.

This kind of "ad hoc" approach may not always work if we have a more complicated system, involving a greater number of variables, or more equations. We will devise a general strategy for solving complicated systems of linear equations.

1.2 Elementary Row Operations

Example 1.2.1 *Find all solutions of the following system :*

$$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array}$$

In other (perhaps simpler) examples we were able to find solutions by simplifying the system (perhaps by eliminating certain variables) through operations of the following types :

1. We could multiply one equation by a non-zero constant.
2. We could add one equation to another (for example in the hope of eliminating a variable from the result).

A similar approach will work for Example 1.2.1 but with this and other harder examples it may not always be clear how to proceed. We now develop a new technique both for describing our system and for applying operations of the above types more systematically and with greater clarity.

Back to Example 1.2.1: We associate a *matrix* to our system of equations (a matrix is a rectangular array of numbers).

$$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array}$$

$$\begin{array}{cl} \left(\begin{array}{rrrr} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ -1 & 4 & 2 & 0 \end{array}\right) & \begin{array}{l} \text{Eqn 1} \\ \text{Eqn 2} \\ \text{Eqn 3} \end{array} \end{array}$$

Note that the first *row* of this matrix contains as its four entries the coefficients of the variables x, y, z, and the number appearing on the right-hand-side of Equation 1 of the system. Rows 2 and 3 correspond similarly to Equations 2 and 3. The *columns* of the matrix correspond (from left to right) to the variables x, y, z and the right hand side of our system of equations.

Definition 1.2.2 *The above matrix is called the* augmented matrix *of the system of equations in Example 1.2.1.*

In solving systems of equations we are allowed to perform operations of the following types:

1. Multiply an equation by a non-zero constant.
2. Add one equation (or a non-zero constant multiple of one equation) to another equation.

These correspond to the following operations on the augmented matrix :

1. Multiply a *row* by a non-zero constant.
2. Add a multiple of one row to another row.
3. We also allow operations of the following type : Interchange two rows in the matrix (this only amounts to writing down the equations of the system in a different order).

Definition 1.2.3 *Operations of these three types are called* Elementary Row Operations (ERO's) *on a matrix.*

We now describe how ERO's on the augmented matrix can be used to solve the system of Example 1.2.1. The following table describes how an ERO is performed at each step to produce a new augmented matrix corresponding to a new (hopefully simpler) system.

ERO	Matrix	System
	$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ -1 & 4 & 2 & 0 \end{pmatrix}$	$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array}$
1. $R3 \to R3 + R1$	$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ 0 & 6 & 1 & 5 \end{pmatrix}$	$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ & & 6y & + & z & = & 5 \end{array}$
2. $R2 \to R2 - 3R1$	$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & -5 & 1 & -6 \\ 0 & 6 & 1 & 5 \end{pmatrix}$	$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & - & 5y & + & z & = & -6 \\ & & 6y & + & z & = & 5 \end{array}$
3. $R2 \to R2 + R3$	$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 6 & 1 & 5 \end{pmatrix}$	$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & & y & + & 2z & = & -1 \\ & & 6y & + & z & = & 5 \end{array}$
4. $R3 \to R3 - 6R2$	$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -11 & 11 \end{pmatrix}$	$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & & y & + & 2z & = & -1 \\ & & & & -11z & = & 11 \end{array}$
5. $R3 \times \left(-\frac{1}{11}\right)$	$\begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$	$\begin{array}{rcrcrcrl} x & + & 2y & - & z & = & 5 & (A) \\ & & y & + & 2z & = & -1 & (B) \\ & & & & z & = & -1 & (C) \end{array}$

We have produced a new system of equations. This is easily solved :

$$\underline{\text{Backsubstitution}} \quad \begin{cases} (C) & z = -1 \\ (B) & y = -1 - 2z \quad \Longrightarrow \quad y = -1 - 2(-1) = 1 \\ (A) & x = 5 - 2y + z \quad \Longrightarrow \quad x = 5 - 2(1) + (-1) = 2 \end{cases}$$

$\boxed{\textbf{Solution : } x = 2,\ y = 1, z = -1}$

You should check that this is a solution of the original system. It is the only solution both of the final system and of the original one (and every intermediate one).

NOTE : The matrix obtained in Step 5 above is in *Row-Echelon Form*. This means :

1. The first non-zero entry in each row is a 1 (called a *Leading 1*).
2. If a column contains a leading 1, then every entry of the column below the leading 1 is a zero.
3. As we move downwards through the rows of the matrix, the leading 1's move from left to right.
4. Any rows consisting entirely of zeroes are grouped together at the bottom of the matrix.

NOTE : The process by which the augmented matrix of a system of equations is reduced to row-echelon form is called *Gaussian Elimination*. In Example 1.2.1 the solution of the system was found by Gaussian elimination with *Backsubstitution*.

General Strategy to Obtain a Row-Echelon Form

1. Get a 1 as the top left entry of the matrix.
2. Use this first leading 1 to "clear out" the rest of the first column, by adding suitable multiples of Row 1 to subsequent rows.
3. If column 2 contains non-zero entries (other than in the first row), use ERO's to get a 1 as the second entry of Row 2. After this step the matrix will look like the following (where the entries represented by stars may be anything):

$$\begin{pmatrix} 1 & * & * & \dots & \dots \\ 0 & 1 & \dots & \dots & \dots \\ 0 & * & \dots & \dots & \dots \\ 0 & * & \dots & \dots & \dots \\ \vdots & \vdots & & & \vdots \\ 0 & * & \dots & \dots & \dots \end{pmatrix}$$

4. Now use this second leading 1 to "clear out" the rest of column 2 (below Row 2) by adding suitable multiples of Row 2 to subsequent rows. After this step the matrix will look like the following :

$$\begin{pmatrix} 1 & * & * & \dots & \dots \\ 0 & 1 & * & \dots & \dots \\ 0 & 0 & * & \dots & \dots \\ 0 & 0 & * & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & * & \dots & \dots \end{pmatrix}$$

5. Now go to column 3. If it has non-zero entries (other than in the first two rows) get a 1 as the third entry of Row 3. Use this third leading 1 to clear out the rest of Column 3, then proceed to column 4. Continue until a row-echelon form is obtained.

Example 1.2.4 *Let A be the matrix*

$$\begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 2 & 1 & -1 & 2 & 8 \\ 1 & -3 & 2 & 7 & 2 \end{pmatrix}$$

Reduce A to row-echelon form.

Solution:

1. Get a 1 as the first entry of Row 1. Done.

2. Use this first leading 1 to clear out column 1 as follows :

$$\begin{array}{lcl} & & \\ R2 & \to & R2 - 2R1 \\ R3 & \to & R3 - R1 \end{array} \quad \begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 0 & 3 & 1 & -2 & 8 \\ 0 & -2 & 3 & 5 & 2 \end{pmatrix}$$

3. Get a leading 1 as the second entry of Row 2, for example as follows :

$$R2 \quad \to \quad R2 + R3 \qquad \begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 0 & 1 & 4 & 3 & 10 \\ 0 & -2 & 3 & 5 & 2 \end{pmatrix}$$

4. Use this leading 1 to clear out whatever appears below it in Column 2 :

$$R3 \quad \to \quad R3 + 2R2 \qquad \begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 0 & 1 & 4 & 3 & 10 \\ 0 & 0 & 11 & 11 & 22 \end{pmatrix}$$

5. Get a leading 1 in Row 3 :

$$R3 \quad \times \quad \frac{1}{11} \qquad \begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 0 & 1 & 4 & 3 & 10 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix}$$

This matrix is now in row-echelon form.

Definition 1.2.5 *Let* A *be a matrix. The* rank *of* A*, denoted rank*(A) *is the number of leading 1's in a row-echelon form obtained from* A *by Gaussian elimination as above.*

REMARKS :

1. The rank of the matrix A in Example 1.2.4 is 3, since the row-echelon form obtained had 3 leading 1's (one in each row).

2. The rank of any matrix can be at most equal to the number of rows, since each row in a REF (row-echelon form) can contain at most one leading 1. If a REF obtained from some matrix contains rows full of zeroes, the rank of this matrix will be less than the number of rows.

3. Starting with a particular matrix, different sequences of ERO's can lead to different row-echelon forms. However, all have the same rank.

1.3 The Reduced Row-Echelon Form (RREF)

Definition 1.3.1 *A matrix is in* reduced row-echelon form (RREF) *if*

1. *It is in row-echelon form, and*
2. *If a particular column contains a leading 1, then* all *other entries of that column are zeroes.*

If we have a row-echelon form, we can use ERO's to obtain a reduced row-echelon form (using ERO's to obtain a RREF is called *Gauss-Jordan elimination*).

Example 1.3.2 *In Example 1.2.4, we obtained the following row-echelon form :*

$$\begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 0 & 1 & 4 & 3 & 10 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix} \quad \text{(REF, not \textit{reduced} REF)}$$

To get a RREF from this REF :

1. Look for the leading 1 in Row 2 - it is in column 2. Eliminate the non-zero entry *above* this leading 1 by adding a suitable multiple of Row 2 to Row 1.

$$\text{R1} \quad \to \quad \text{R1} + \text{R2} \qquad \begin{pmatrix} 1 & 0 & 3 & 5 & 10 \\ 0 & 1 & 4 & 3 & 10 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix}$$

2. Look for the leading 1 in Row 3 - it is in column 3. Eliminate the non-zero entries *above* this leading 1 by adding suitable multiples of Row 3 to Rows 1 and 2.

$$\begin{array}{lcl} \text{R1} & \to & \text{R1} - 3\text{R3} \\ \text{R2} & \to & \text{R2} - 4\text{R3} \end{array} \qquad \begin{pmatrix} 1 & 0 & 0 & 2 & 4 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix}$$

This matrix is in *reduced* row-echelon form.

The technique outlined in this example will work in general to obtain a RREF from a REF : you should practise with similar examples!

REMARK: Different sequences of ERO's on a matrix can lead to different row-echelon forms. However, only one *reduced* row-echelon form can be found from any matrix.

1.4 Leading Variables and Free Variables

Example 1.4.1 *Find the general solution of the following system :*

$$\begin{array}{rcrcrcrcrl} x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & 0 & \quad \mathrm{I} \\ 2x_1 & + & x_2 & - & x_3 & + & 2x_4 & = & 8 & \quad \mathrm{II} \\ x_1 & - & 3x_2 & + & 2x_3 & + & 7x_4 & = & 2 & \quad \mathrm{III} \end{array}$$

SOLUTION :

1. Write down the augmented matrix of the system :

$$\begin{array}{l} \text{Eqn I} \\ \text{Eqn II} \\ \text{Eqn III} \end{array} \quad \begin{array}{c} \left(\begin{array}{rrrrr} 1 & -1 & -1 & 2 & 0 \\ 2 & 1 & -1 & 2 & 8 \\ 1 & -3 & 2 & 7 & 2 \end{array}\right) \\ \begin{array}{rrrr} x_1 & x_2 & x_3 & x_4 \end{array} \end{array}$$

   Note : This is the matrix of Example 1.2.4

2. Use Gauss-Jordan elimination to find a reduced row-echelon form from this augmented matrix. We have already done this in Examples 1.2.4 and 1.3.2 :-

$$\text{RREF :} \quad \begin{array}{c} \left(\begin{array}{rrrrr} 1 & 0 & 0 & 2 & 4 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right) \\ \begin{array}{rrrr} x_1 & x_2 & x_3 & x_4 \end{array} \end{array}$$

   This matrix corresponds to a new system of equations:

$$\begin{array}{rcrcrl} x_1 & + & 2x_4 & = & 4 & \quad (A) \\ x_2 & - & x_4 & = & 2 & \quad (B) \\ x_3 & + & x_4 & = & 2 & \quad (C) \end{array}$$

   Remark : The RREF involves 3 leading 1's, one in each of the columns corresponding to the variables x_1, x_2 and x_3. The column corresponding to x_4 contains no leading 1.

   Definition 1.4.2 *The variables whose columns in the RREF contain leading 1's are called* leading variables*. A variable whose column in the RREF does not contain a leading 1 is called a* free variable*.*

   So in this example the leading variables are x_1, x_2 and x_3, and the variable x_4 is free. What does this distinction mean in terms of solutions of the system? The system corresponding to the RREF can be rewritten as follows :

$$\begin{array}{rcrcrl} x_1 & = & 4 & - & 2x_4 & \quad (A) \\ x_2 & = & 2 & + & x_4 & \quad (B) \\ x_3 & = & 2 & - & x_4 & \quad (C) \end{array}$$

   i.e. this RREF tells us how the values of the leading variables x_1, x_2 and x_3 *depend* on that of the free variable x_4 in a solution of the system. In a solution, the free variable x_4 may assume the value of *any* real number. However, once a value for x_4 is chosen, values are immediately assigned to x_1, x_2 and x_3 by equations A, B and C above. For example

   (a) Choosing $x_4 = 0$ gives $x_1 = 4 - 2(0) = 4,\ x_2 = 2 + (0) = 2,\ x_3 = 2 - (0) = 2$. Check that $x_1 = 4,\ x_2 = 2,\ x_3 = 2,\ x_4 = 0$ is a solution of the (original) system.

   (b) Choosing $x_4 = 3$ gives $x_1 = 4 - 2(3) = -2,\ x_2 = 2 + (3) = 5,\ x_3 = 2 - (3) = -1$. Check that $x_1 = -2,\ x_2 = 5,\ x_3 = -1,\ x_4 = 3$ is a solution of the (original) system.

Different choices of value for x_4 will give different solutions of the system. The number of solutions is *infinite*.

The *general solution* is usually described by the following type of notation. We assign the *parameter* name t to the value of the variable x_4 in a solution (so t may assume any real number as its value). We then have

$$x_1 = 4 - 2t,\ x_2 = 2 + t,\ x_3 = 2 - t,\ x_4 = t;\ t \in \mathbb{R}$$

or

$$\textbf{General Solution :}\ (x_1, x_2, x_3, x_4) = (4 - 2t, 2 + t, 2 - t, t);\ t \in \mathbb{R}$$

This general solution describes the infinitely many solutions of the system : we get a *particular* solution by choosing a specific numerical value for t : this determines specific values for x_1, x_2, x_3 and x_4.

Example 1.4.3 *Solve the following system of linear equations :*

$$\begin{array}{rcrcrcrcrr} x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & 0 & \text{I} \\ 2x_1 & + & x_2 & - & x_3 & + & 2x_4 & = & 8 & \text{II} \\ x_1 & - & 3x_2 & + & 2x_3 & + & 7x_4 & = & 2 & \text{III} \\ x_1 & - & x_2 & + & x_3 & - & x_4 & = & -6 & \text{IV} \end{array}$$

Remark : The first three equations of this system comprise the system of equations of Example 1.4.1. The problem becomes : Can we find a solution of the system of Example 1.4.1 which is in addition a solution of the equation $x_1 - x_2 + x_3 - x_4 = -6$?

SOLUTION We know that every simultaneous solution of the first three equations has the form

$$x_1 = 4 - 2t,\ x_2 = 2 + t,\ x_3 = 2 - t,\ x_4 = t,$$

where t can be any real number . Is there some choice of t for which the solution of the first three equations is also a solution of the fourth? i.e. for which

$$x_1 - x_2 + x_3 - x_4 = -6 \text{ i.e. } (4 - 2t) - (2 + t) + (2 - t) - t = -6$$

Solving for t gives

$$\begin{aligned} 4 - 5t &= -6 \\ \Longrightarrow -5t &= 10 \\ \Longrightarrow t &= 2 \end{aligned}$$

$t = 2\ :\ x_1 = 4 - 2t = 4 - 2(2) = 0;\ x_2 = 2 + t = 2 + 2 = 4;\ x_3 = 2 - t = 2 - 2 = 0;\ x_4 = t = 2$

SOLUTION : $x_1 = 0,\ x_2 = 4,\ x_3 = 0,\ x_4 = 2$ (or $(x_1, x_2, x_3, x_4) = (0, 4, 0, 2)$).
This is the *unique* solution to the system in Example 1.4.3.

REMARKS**:**

1. To solve the system of Example 1.4.3 directly (without 1.4.1) we would write down the augmented matrix :

$$\begin{pmatrix} 1 & -1 & -1 & 2 & 0 \\ 2 & 1 & -1 & 2 & 8 \\ 1 & -3 & 2 & 7 & 2 \\ 1 & -1 & 1 & -1 & -6 \end{pmatrix}$$

Check: Gauss-Jordan elimination gives the reduced row-echelon form :

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

which corresponds to the system

$$x_1 = 0;\ x_2 = 4;\ x_3 = 0;\ x_4 = 2$$

i.e. the unique solution is given exactly by the RREF. In this system, all four variables are leading variables. This is always the case for a system which has a unique solution : that each variable is a leading variable, i.e. corresponds in the RREF of the augmented matrix to a column which contains a leading 1.

2. The system of Example 1.4.1, consisting of Equations 1,2 and 3 of that in Example 1.4.3, had an infinite number of solutions. Adding the fourth equation in Example 1.4.3 pinpointed exactly one of these infinitely many solutions.

1.5 Consistent and Inconsistent Systems

Example 1.5.1 *Consider the following system :*

$$\begin{array}{rcrcrcr} 3x & + & 2y & - & 5z & = & 4 \\ x & + & y & - & 2z & = & 1 \\ 5x & + & 3y & - & 8z & = & 6 \end{array}$$

To find solutions, obtain a row-echelon form from the augmented matrix :

$$\begin{pmatrix} 3 & 2 & -5 & 4 \\ 1 & 1 & -2 & 1 \\ 5 & 3 & -8 & 6 \end{pmatrix} \quad \xrightarrow{R1 \leftrightarrow R2} \quad \begin{pmatrix} 1 & 1 & -2 & 1 \\ 3 & 2 & -5 & 4 \\ 5 & 3 & -8 & 6 \end{pmatrix}$$

$$\begin{array}{c} R2 \to R2 - 3R1 \\ \longrightarrow \\ R3 \to R3 - 5R1 \end{array} \begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & -2 & 2 & 1 \end{pmatrix} \quad \xrightarrow{R2 \times (-1)} \quad \begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & -2 & 2 & 1 \end{pmatrix}$$

$$\xrightarrow{R3 \to R3 + 2R2} \begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \xrightarrow{R3 \times (-1)} \quad \begin{pmatrix} 1 & 1 & -2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(Row-Echelon Form)

The system of equations corresponding to this REF has as its third equation

$$0x + 0y + 0z = 1 \quad \text{i.e. } 0 = 1$$

This equation clearly has no solutions - no assignment of numerical values to x, y and z will make the value of the expression $0x + 0y + 0z$ equal to anything but zero. Hence the system has no solutions.

Definition 1.5.2 *A system of linear equations is called* inconsistent *if it has no solutions. A system which has a solution is called* consistent*.*

If a system is inconsistent, a REF obtained from its augmented matrix will include a row of the form 0 0 0 ... 0 1, i.e. will have a leading 1 in its rightmost column. Such a row corresponds to an equation of the form $0x_1 + 0x_2 + \cdots + 0x_n = 1$, which certainly has no solution.

Example 1.5.3 *(MA203 Summer 2005, Q1)*

(a) Find the unique value of t for which the following system has a solution.

$$\begin{array}{rcrcrcrcr} -x_1 & & & + & x_3 & - & x_4 & = & 3 \\ 2x_1 & + & 2x_2 & - & x_3 & - & 7x_4 & = & 1 \\ 4x_1 & - & x_2 & - & 9x_3 & - & 5x_4 & = & t \\ 3x_1 & - & x_2 & - & 8x_3 & - & 6x_4 & = & 1 \end{array}$$

SOLUTION: First write down the augmented matrix and begin Gauss-Jordan elimination.

$$\begin{pmatrix} -1 & 0 & 1 & -1 & 3 \\ 2 & 2 & -1 & -7 & 1 \\ 4 & -1 & -9 & -5 & t \\ 3 & -1 & -8 & -6 & 1 \end{pmatrix} \quad \xrightarrow{R1 \times (-1)} \quad \begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 2 & 2 & -1 & -7 & 1 \\ 4 & -1 & -9 & -5 & t \\ 3 & -1 & -8 & -6 & 1 \end{pmatrix}$$

$$\begin{array}{c} R2 \to R2 - 2R1 \\ R3 \to R3 - 4R1 \\ \longrightarrow \\ R4 \to R4 - 3R1 \end{array} \begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 0 & 2 & 1 & -9 & 7 \\ 0 & -1 & -5 & -9 & t+12 \\ 0 & -1 & -5 & -9 & 10 \end{pmatrix} \quad \xrightarrow{R3 \to R3 - R4} \quad \begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 0 & 2 & 1 & -9 & 7 \\ 0 & 0 & 0 & 0 & t+2 \\ 0 & -1 & -5 & -9 & 10 \end{pmatrix}$$

From the third row of this matrix we can see that the system can be consistent only if $t + 2 = 0$. i.e. only if $t = -2$.

(b) Find the general solution of this system for this value of t.
SOLUTION: Set $t = -2$ and continue with the Gaussian elimination. We omit the third row, which consists fully of zeroes and carries no information.

$$\begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 0 & 2 & 1 & -9 & 7 \\ 0 & -1 & -5 & -9 & 10 \end{pmatrix} \begin{array}{c} R4 \times (-1) \\ \longrightarrow \\ R3 \leftrightarrow R4 \end{array} \begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 0 & 1 & 5 & 9 & -10 \\ 0 & 2 & 1 & -9 & 7 \end{pmatrix}$$

$$\begin{array}{c} R3 \to R3 - 2R2 \\ \longrightarrow \end{array} \begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 0 & 1 & 5 & 9 & -10 \\ 0 & 0 & -9 & -27 & 27 \end{pmatrix} \begin{array}{c} R3 \times (-\frac{1}{9}) \\ \longrightarrow \end{array} \begin{pmatrix} 1 & 0 & -1 & 1 & -3 \\ 0 & 1 & 5 & 9 & -10 \\ 0 & 0 & 1 & 3 & -3 \end{pmatrix}$$

$$\begin{array}{c} R1 \to R1 + R3 \\ \longrightarrow \\ R2 \to R2 + 5R3 \end{array} \begin{pmatrix} 1 & 0 & 0 & 4 & -6 \\ 0 & 1 & 0 & -6 & 5 \\ 0 & 0 & 1 & 3 & -3 \end{pmatrix}$$

Having reached a reduced row-echelon form, we can see that the variables x_1, x_2 and x_3 are leading variables, and the variable x_4 is free. We have from the RREF

$$x_1 = -6 - 4x_4, \ x_2 = 5 + 6x_4, \ x_3 = -3 - 3x_4.$$

If we assign the parameter name s to the value of the free variable x_4 in a solution of the system, we can write the general solution as

$$(x_1, x_2, x_3, x_4) = (-6 - 4s, 5 + 6s, -3 - 3s, s), \ s \in \mathbb{R}.$$

Summary of Possible Outcomes when Solving a System of Linear Equations:

1. The system may be inconsistent. This happens if a REF obtained from the augmented matrix has a leading 1 in its rightmost column.

2. The system may be consistent. In this case one of the following occurs :

   (a) There may be a unique solution. This will happen if all variables are leading variables, i.e. every column except the rightmost one in a REF obtained from the augmented matrix has a leading 1. In the case the *reduced* row-echelon form obtained from the augmented matrix will have the following form :

   $$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & * \\ 0 & 1 & 0 & \dots & 0 & * \\ 0 & 0 & 1 & \dots & 0 & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & * \end{pmatrix}$$

   with possibly some additional rows full of zeroes at the bottom. The unique solution can be read from the right-hand column.

   NOTE: If a system of equations has a unique solution, the number of equations must be at least equal to the number of variables (since the augmented matrix must have enough rows to accommodate a leading 1 for every variable).

   (b) There may be infinitely many solutions. This happens if the system is consistent but at least one of the variables is free. In this case the rank of the augmented matrix will be less than the number of variables in the system.

Chapter 2

Gauss-Jordan Elimination and Matrix Algebra

2.1 Review of Matrix Algebra

A $m \times n$ ("m by n") matrix is a matrix having m rows and n columns.

EXAMPLE: $\begin{pmatrix} 2 & 3 & -1 \\ -3 & -4 & 0 \end{pmatrix}$ is a 2×3 matrix.

$\begin{pmatrix} 2 & 3 \\ 2 & 7 \\ 4 & 0 \end{pmatrix}$ is a 3×2 matrix.

Two matrices are said to have the same *size* if they have the same number of rows and the same number of columns. (So for example a 3×2 matrix and a 2×3 matrix are considered to be of different size).

NOTATION: If A is an $m \times n$ matrix, the entry appearing in the ith row and jth column of A (called the (i,j) position) is denoted $(A)_{ij}$.

EXAMPLE: Let $A = \begin{pmatrix} 2 & 3 & -1 \\ 4 & 0 & 5 \end{pmatrix}$.
Then $(A)_{11} = 2$, $(A)_{21} = 4$, $(A)_{13} = -1$, etc.

Like numbers, matrices have arithmetic associated to them. In particular, a pair of matrices can be added or multiplied (subject to certain compatibility conditions on their sizes) to produce a new matrix.

MATRIX ADDITION:
LetA and B be matrices of the same size ($m \times n$). We define their *sum* $A + B$ to be the $m \times n$ matrix whose entries are given by

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n$
Thus $A + B$ is obtained from A and B by adding entries in corresponding positions.
EXAMPLE: Let $A = \begin{pmatrix} 2 & 0 & -1 & -1 \\ 1 & 2 & 4 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 1 & 0 & -2 \\ 3 & -3 & 1 & 1 \end{pmatrix}$. Then

$$A + B = \begin{pmatrix} 2 + (-1) & 0 + 1 & -1 + 0 & -1 + (-2) \\ 1 + 3 & 2 + (-3) & 4 + 1 & 2 + 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 & -3 \\ 4 & -1 & 5 & 3 \end{pmatrix}$$

Subtraction of matrices is now defined in the obvious way - e.g., with A and B as above, we have

$$A - B = \begin{pmatrix} 2 - (-1) & 0 - 1 & -1 - 0 & -1 - (-2) \\ 1 - 3 & 2 - (-3) & 4 - 1 & 2 - 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 & -1 & 1 \\ -2 & 5 & 3 & 1 \end{pmatrix}$$

MULTIPLICATION OF A MATRIX BY A REAL NUMBER:
Let A be a $m \times n$ matrix and let c be a real number. Then cA is the $m \times n$ matrix with entries defined by

$$(cA)_{ij} = c(A)_{ij}$$

i.e. cA is obtained from A by multiplying every entry by c.
EXAMPLE: If $A = \begin{pmatrix} 2 & 1 \\ 3 & -4 \end{pmatrix}$, then

$$2A = \begin{pmatrix} 4 & 2 \\ 6 & -8 \end{pmatrix}, \; -3A = \begin{pmatrix} -6 & -3 \\ -9 & 12 \end{pmatrix}, \; 0A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The $m \times n$ matrix whose entries are all zero is called the *zero* $(m \times n)$ matrix.

Matrix Multiplication
Unlike addition, the manner in which matrices are multiplied does not appear completely natural at first glance.

Suppose that A is a $m \times p$ matrix and B is a $q \times n$ matrix. Then the product AB is defined if and only if $p = q$, i.e. if and only if

- The number of columns in A = the number of rows in B, or
- The number of entries in a row of A = the number of entries in a column of B.

In this case the size of AB is $m \times n$.
In general the following "cancellation law" holds for the size of matrix products:

$$\text{“}(m \times \cancel{p}) \times (\cancel{p} \times n) = m \times n\text{”}.$$

This means : if A is a $m \times p$ matrix and B is a $p \times n$ matrix, then the product AB is defined and is a $m \times n$ matrix in which the entry in the ith row and jth column is given by combining the entries of the ith row of A with those of the jth column of B according to the following rule :

$$\text{product of 1st entries} + \text{product of 2nd entries} + \cdots + \text{product of } p\text{th entries}$$

Example 2.1.1 *Let* $A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & -1 \end{pmatrix}$ *and let* $B = \begin{pmatrix} 3 & 1 \\ 1 & -1 \\ 0 & 2 \end{pmatrix}$.

Find AB *and* BA.

SOLUTION :

1. $A: \ 2 \times 3, \ B: \ 3 \times 2 \Longrightarrow AB$ will be a 2×2 matrix.

$$\begin{aligned} \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & -1 \\ 0 & 2 \end{pmatrix} &= \begin{pmatrix} 2(3) + (-1)(1) + 3(0) & 2(1) + (-1)(-1) + 3(2) \\ 1(3) + 0(1) + (-1)(0) & 1(1) + 0(-1) + (-1)(2) \end{pmatrix} \\ &= \begin{pmatrix} 5 & 9 \\ 3 & -1 \end{pmatrix} \end{aligned}$$

2. $B: \ 3 \times 2, \ A: \ 2 \times 3 \Longrightarrow BA$ will be a 3×3.

$$BA = \begin{pmatrix} 7 & -3 & 8 \\ 1 & -1 & 4 \\ 2 & 0 & -2 \end{pmatrix}$$

(Exercise)

NOTE: $BA \neq AB$: Matrix multiplication is *not* commutative. In this example AB and BA are both defined but do not even have the same size. It is also possible for only one of AB and BA to be defined, for example this will happen if A is 2×4 and B is 4×3. Even if AB and BA are both defined and have the same size (for example if both are 3×3), the two products are typically different.

The next example shows how the computations involved in matrix multiplication can arise in a natural setting.

Example 2.1.2 *A salesperson sells items of three types I, II, and III, costing €10, €20 and €30 respectively. The following table shows how many items of each type are sold on Monday morning and afternoon.*

	Type I	Type II	Type III
morning	3	4	1
afternoon	5	2	2

Let A denote the matrix

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

Let B denote the 3×1 matrix whose entries are the prices of items of Type I, II and III respectively.

$$B = \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}$$

Let C denote the 2×1 matrix whose entries are respectively the total income from morning sales and the total income from afternoon sales. Then

1st entry of C : $(3 \times 10) + (4 \times 20) + (1 \times 30) = 140$
1st entry of C : $(5 \times 10) + (2 \times 20) + (2 \times 30) = 150$

So $C = \begin{pmatrix} 140 \\ 150 \end{pmatrix}$

Now note that according to the definition of matrix multiplication we have

$$AB = C.$$

.

<u>1st entry of C</u>: comes from combining the first row of A with the column of B according to :

product of 1st entries + product of 2nd entries + product of 3rd entries

$$(3 \times 10) + (4 \times 20) + (1 \times 30)$$

<u>2nd entry of C</u>: comes from combining the second row of A with the column of B in the same way.

$$(5 \times 10) + (2 \times 20) + (2 \times 30)$$

2.2 The $n \times n$ Identity Matrix

Notation: The set of $n \times n$ matrices with real entries is denoted $M_n(\mathbb{R})$.

Example 2.2.1 $A = \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix}$ *and let* $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. *Find* AI *and* IA.

Solution:

$$AI = \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2(1)+3(0) & 2(0)+3(1) \\ -1(1)+2(0) & -1(0)+2(1) \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix} = A$$

$$IA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1(2)+0(-1) & 1(3)+0(2) \\ 0(2)+1(-1) & 0(3)+1(2) \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix} = A$$

Both AI and IA are equal to A : multiplying A by I (on the left or right) does not affect A.

In general, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is *any* 2×2 matrix, then

$$AI = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

and $IA = A$ also.

Definition 2.2.2 $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *is called the* 2×2 identity matrix *(sometimes denoted* I_2*).*

Remarks:

1. The matrix I behaves in $M_2(\mathbb{R})$ like the real number 1 behaves in $\mathbb{R}$ - multiplying a real number x by 1 has no effect on x.
2. Generally in algebra an *identity element* (sometimes called a *neutral element*) is one which has no effect with respect to a particular algebraic operation.
   For example 0 is the identity element for addition of numbers because adding zero to another number has no effect.
   Similarly 1 is the identity element for multiplication of numbers.
   I_2 is the identity element for multiplication of 2×2 matrices.
3. The 3×3 identity matrix is $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ Check that if A is any 3×3 matrix then $AI_3 = I_3A = A$.

Definition 2.2.3 *For any positive integer* n, *the* $n \times n$ identity matrix I_n *is defined by*

$$I_n = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix}$$

(I_n has 1's along the "main diagonal" and zeroes elsewhere). The entries of I_n *are given by :*

$$(I_n)_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Theorem 2.2.4

1. *If* A *is any matrix with* n *rows then* $I_nA = A$.
2. *If* A *is any matrix with* n *columns, then* $AI_n = A$.

(i.e. multiplying any matrix A (of admissible size) on the left or right by I_n leaves A unchanged).
<u>Proof</u> (of Statement 1 of the Theorem): Let A be a $n \times p$ matrix. Then certainly the product $I_n A$ is defined and its size is $n \times p$.

We need to show that for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant p$,the entry in the ith row and jth column of the product $I_n A$ is equal to the entry in the ith row and jth column of A.

$$\underset{I_n}{\begin{pmatrix} & & & 0 & \\ & & \ddots & \vdots & \\ 0 & \dots & 0 & 1 & 0 \\ & & & \vdots & \\ & & & 0 & \end{pmatrix}} \underset{A}{\begin{pmatrix} & A_{1j} & \\ & \vdots & \\ \dots & A_{ij} & \\ & \vdots & \\ & A_{nj} & \end{pmatrix}} = \underset{I_n A}{\begin{pmatrix} & & \vdots & \\ & & \vdots & \\ \dots & \dots & \bullet & \dots \\ & & \vdots & \end{pmatrix}}$$

$(I_n A)_{ij}$ comes from the ith row of I_n and the jth column of A.

$$\begin{aligned} (I_n A)_{ij} &= (0)(A)_{1j} + (0)(A)_{2j} + \cdots + (1)(A)_{ij} + \cdots + (0)(A)_{nj} \\ &= (1)(A)_{ij} \\ &= (A)_{ij} \end{aligned}$$

Thus $(I_n A)_{ij} = (A)_{ij}$ for all i and j - the matrices $I_n A$ and A have the same entries in each position. Then $I_n A = A$.
The proof of Statement 2 is similar. $\square$

2.3 The Inverse of a Matrix

Notation: For a positive integer n, we let $M_n(\mathbb{R})$ denote the set of $n \times n$ matrices with entries in $\mathbb{R}$.

Remark: When we work in the full set of matrices over $\mathbb{R}$, it is not always possible to add or multiply two matrices (these operations are subject to restrictions on the sizes of the matrices involved). However, if we restrict attention to $M_n(\mathbb{R})$ we can add any pair of matrices and multiply any pair of matrices, and we never move outside $M_n(\mathbb{R})$.

$M_n(\mathbb{R})$ is an example of the type of algebraic structure known as a *ring*.

In this section we will consider how we might define a version of "division" for matrices in $M_n(\mathbb{R})$.

In the set $\mathbb{R}$ of real numbers, dividing by a non-zero number x means multiplying by the *reciprocal* $1/x$ of x. For example if we divide a real number by 5 we are multiplying it by $\frac{1}{5}$: $\frac{1}{5}$ is the *reciprocal* or *multiplicative inverse* of 5 in $\mathbb{R}$. This means

$$\frac{1}{5} \times 5 = 1,$$

i.e., if you multiply 5 by $\frac{1}{5}$, you get 1; multiplying by $\frac{1}{5}$ "reverses" the work of multiplying by 5.

Definition 2.3.1 *Let* A *be a* $n \times n$ *matrix.If* B *is a* $n \times n$ *matrix for which*

$$AB = I_n \text{ and } BA = I_n$$

then B *is called an* inverse *for* A*.*

Example: Let $A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ and let $B = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$. Then

$$\begin{aligned} AB = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \\ BA = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \end{aligned}$$

So B is an inverse for A.

Remarks:

1. Suppose B and C are *both* inverses for a particular matrix A, i.e.

   $$BA = AB = I_n \text{ and } CA = AC = I_n$$

   Then

   $$\begin{aligned} (BA)C &= I_nC = C \\ \text{Also } (BA)C &= B(AC) = BI_n = B \end{aligned}$$

   Hence $B = C$, and if A has an inverse, its inverse is unique. Thus we can talk about *the* inverse of a matrix.

2. The inverse of a $n \times n$ matrix A, if it exists, is denoted A^{-1}.

3. Not every square matrix has an inverse. For example the 2×2 zero matrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ does not.

   In Example 1.5.1 we saw that the system

   $$\begin{array}{rcrcrcr} 3x & + & 2y & - & 5z & = & 4 \\ x & + & y & - & 2z & = & 1 \\ 5x & + & 3y & - & 8z & = & 6 \end{array}$$

is inconsistent. This system can be written in matrix form as follows

$$\begin{pmatrix} 3x & + & 2y & - & 5z \\ x & + & y & - & 2z \\ 5x & + & 3y & - & 8z \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 6 \end{pmatrix}.$$

The left hand side of this equation can be written as the matrix product of the 3×3 coefficient matrix of the system and the column containing the variable names to obtain the following version :

$$\begin{pmatrix} 3 & 2 & -5 \\ 1 & 1 & -2 \\ 5 & 3 & -8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 6 \end{pmatrix}$$

We let A denote the 3×3 matrix above.

If this matrix had an inverse, we could multiply both sides of the above equation on the left by A^{-1} to obtain

$$A^{-1}A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 4 \\ 1 \\ 6 \end{pmatrix} \Longrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 4 \\ 1 \\ 6 \end{pmatrix}.$$

This would mean that the system has a unique solution in which the values of x, y, z are the entries of the matrix $A^{-1} \begin{pmatrix} 4 \\ 1 \\ 6 \end{pmatrix}$.

Since we know from Example 1.5.1 that the system has no solution, we must conclude that the matrix A has no inverse in $M_3(\mathbb{R})$.

General Fact : *Suppose that a system of equations ha a square coefficient matrix. If this coefficient matrix has an inverse the system has a unique solution.*

4. A square matrix that has an inverse is called invertible or non-singular. A matrix that has no inverse is called singular or non-invertible.

5. *A Converse to Item 3 above*: Suppose now that A is a $n \times n$ matrix (say 3×3) and that there is a system of equations with A as coefficient matrix that has a unique solution. Then the RREF obtained from the augmented matrix of the system has the following form

$$\begin{pmatrix} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{pmatrix}.$$

Since the rightmost column does not contribute to the choice of elementary row operations, it follows that *every* system of linear equations having A as coefficient matrix has an augmented matrix with a RREF of the above form. Thus every system of equations having A as coefficient matrix has a unique solution.

In particular then the system described by

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

has a unique solution $x = a_1,\ y = a_2,\ z = a_3$.

Similarly the systems described by

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\quad A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

have unique solutions given respectively by $x = b_1,\ y = b_2,\ z = b_3$ and $x = c_1,\ y = c_2,\ z = c_3$.

Now define

$$B = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix},$$

and look at the product AB. This is the 3×3 identity matrix I_3. Thus B is an inverse for A and A is invertible. We conclude that if A is the coefficient matrix of a system having a unique solution, then A is invertible.

Putting this together with Item 3. above and the remarks at the end of Section 2.2, we obtain the following :

Theorem 2.3.2 *A $n \times n$ matrix A is invertible if and only if the following equivalent conditions hold.*

(a) Every system of linear equations with A as coefficient matrix has a unique solution.

(b) A can be reduced by elementary row operations to the $n \times n$ identity matrix.

2.4 A Method to Calculate the Inverse of a Matrix

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

Assume for now that A is invertible and suppose that

$$A^{-1} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Then $AA^{-1} = I_3$ and in particular the first column of AA^{-1} is $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. This first column comes from the entries of A combined with the first column of A^{-1}. Thus we have

$$A \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

This means $x = a_1,\ y = a_2,\ z = a_3$ is the unique solution of the system of linear equations given by

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Thus the entries a_1, a_2, a_3 of the first column of A^{-1} will be written in the rightmost column of the RREF obtained from the matrix

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

Similarly the second and third columns of A^{-1} are respectively the unique solutions of the systems

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

They are respectively written in the rightmost columns of the RREFs obtained by EROs from the augmented matrices

$$\begin{pmatrix} 3 & 4 & -1 & 0 \\ 1 & 0 & 3 & 1 \\ 2 & 5 & -4 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 3 & 4 & -1 & 0 \\ 1 & 0 & 3 & 0 \\ 2 & 5 & -4 & 1 \end{pmatrix}.$$

So to find A^{-1} we need to reduce these three augmented matrices to RREF. This can be done with a single series of EROs if we start with the 3×6 matrix

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

Method : Reduce A' to RREF. If the RREF has I_3 in its first three columns, then columns 4,5,6 contain A^{-1}.

We proceed as follows.

$$\left(\begin{array}{rrrrrr} 3 & 4 & -1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 5 & -4 & 0 & 0 & 1 \end{array}\right) \quad \overset{R1 \leftrightarrow R2}{\longrightarrow} \quad \left(\begin{array}{rrrrrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 3 & 4 & -1 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 0 & 1 \end{array}\right)$$

$$\begin{array}{c} R2 \to R2 - 3R1 \\ \longrightarrow \\ R3 \to R3 - 2R1 \end{array} \left(\begin{array}{rrrrrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 4 & -10 & 1 & -3 & 0 \\ 0 & 5 & -10 & 0 & -2 & 1 \end{array}\right) \quad \overset{R3 \to R3 - R2}{\longrightarrow} \quad \left(\begin{array}{rrrrrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 4 & -10 & 1 & -3 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \end{array}\right)$$

$$\overset{R3 \leftrightarrow R2}{\longrightarrow} \left(\begin{array}{rrrrrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 4 & -10 & 1 & -3 & 0 \end{array}\right) \quad \overset{R3 \leftrightarrow R3 - 4R2}{\longrightarrow} \quad \left(\begin{array}{rrrrrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & -10 & 5 & -7 & -4 \end{array}\right)$$

$$\overset{R3 \times (-\frac{1}{10})}{\longrightarrow} \left(\begin{array}{rrrrrr} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{array}\right) \quad \overset{R1 \to R1 - 3R3}{\longrightarrow} \quad \left(\begin{array}{rrrrrr} 1 & 0 & 0 & \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{array}\right)$$

The above matrix is in RREF and its first three columns comprise I_3. We conclude that the matrix A^{-1} is written in the last three columns, i.e.

$$A^{-1} = \left(\begin{array}{rrr} \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ -1 & 1 & 1 \\ -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{array}\right).$$

(It is easily checked that $AA^{-1} = I_3$).

Note The above procedure can be used to find the inverse of any $n \times n$ matrix A or to show that A is not invertible.

- Form the matrix $A' = (A|I_n)$.
- Apply elementary row operations to A' to reduce it to RREF.
- If a row having 0 in all of the first n positions appears, then A is not invertible.
- If the RREF has I_n in the first n columns, then the matrix formed by the last n columns is A^{-1}.

Example: If we apply this technique to the matrix $A = \left(\begin{array}{rrr} 3 & 2 & -5 \\ 1 & 1 & -2 \\ 5 & 3 & -8 \end{array}\right)$ of Example 1.5.1, we get

$$\left(\begin{array}{rrrrrr} 3 & 2 & -5 & 1 & 0 & 0 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ 5 & 3 & -8 & 0 & 0 & 1 \end{array}\right) \quad \overset{R1 \leftrightarrow R2}{\longrightarrow} \quad \left(\begin{array}{rrrrrr} 1 & 1 & -2 & 0 & 1 & 0 \\ 3 & 2 & -5 & 1 & 0 & 0 \\ 5 & 3 & -8 & 0 & 0 & 1 \end{array}\right)$$

$$\begin{array}{c} R2 \to R2 - 3R1 \\ \longrightarrow \\ R3 \to R3 - 5R1 \end{array} \left(\begin{array}{rrrrrr} 1 & 1 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & -3 & 0 \\ 0 & -2 & 2 & 0 & -5 & 1 \end{array}\right) \quad \overset{R3 \to R3 - 2R2}{\longrightarrow} \quad \left(\begin{array}{rrrrrr} 1 & 1 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & -3 & 0 \\ 0 & 0 & 0 & * & * & * \end{array}\right)$$

At this stage we can conclude that the matrix A is not invertible.

Example (MA203 Summer 2004 Q2 (a))
Find the last row of A^{-1} where

$$A = \left(\begin{array}{rrrr} 1 & 1 & 0 & 1 \\ 2 & 0 & 2 & 2 \\ -1 & 0 & 2 & 1 \\ 2 & 1 & 0 & 1 \end{array}\right).$$

Solution: Suppose that the last row of A^{-1} is $(x\ y\ z\ w)$. Then

$$(x\ y\ z\ w) \left(\begin{array}{rrrr} 1 & 1 & 0 & 1 \\ 2 & 0 & 2 & 2 \\ -1 & 0 & 2 & 1 \\ 2 & 1 & 0 & 1 \end{array}\right) = (0\ 0\ 0\ 1).$$

Thus

$$\begin{array}{ccccccccc} x & + & 2y & - & z & + & 2w & = & 0 \\ x & + & & & & & w & = & 0 \\ & & 2y & + & 2z & & & = & 0 \\ x & + & 2y & + & z & - & w & = & 1 \end{array}$$

So the entries of the fourth row of A^{-1} are the values of x, y, z, w in the unique solution of the system with augmented matrix

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

The coefficient matrix here is A^T, the transpose of A. The right-hand column contains the entries of the fourth row of I_4. Applying elementary row operations to the above matrix results in

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{3}{7} \\ 0 & 1 & 0 & 0 & \frac{1}{7} \\ 0 & 0 & 1 & 0 & -\frac{1}{7} \\ 0 & 0 & 0 & 1 & -\frac{3}{7} \end{pmatrix}.$$

We conclude that the final row of A^{-1} is

$$\begin{pmatrix} \frac{3}{7} & \frac{1}{7} & -\frac{1}{7} & -\frac{3}{7} \end{pmatrix}.$$

2.5 Elementary Row Operations and the Determinant

RECALL: Let A be a 2×2 matrtix : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The *determinant* of A, denoted by $\det(A)$ or $|A|$, is the number $ad - bc$. So for example if

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 5 \end{pmatrix}, \ \det(A) = 2(5) - 4(1) = 6.$$

The matrix A is invertible if and only if $\det(A) \neq 0$, and in this case the inverse of A is given by

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

The matrix $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ is called the *adjoint* or *adjugate* of A, denoted $\mathrm{adj}(A)$.

Determinants are defined for all square matrices. They have various interpretations and applications in algebra, analysis and geometry. For every square matrix A, we have that A is invertible if and only if $\det(A) \neq 0$.

If A is a 2×2 matrix, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $|\det(A)|$ is the volume of the parallelogram having the vectors $\vec{v}_1 = (a, b)$ and $\vec{v}_2 = (c, d)$ as edges. Similarly if

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

is a 3×3 matrix we have $|\det(A)| =$ Volume of P, where P is the parallelepiped in $\mathbb{R}^3$ having $\vec{v}_1 = (a, b, c)$, $\vec{v}_2 = (d, e, f)$ and $\vec{v}_3 = (g, h, i)$ as edges.

The determinant of an $n \times n$ matrix can be defined recursively in terms of determinants of $(n-1) \times (n-1)$ matrices (which in turn are defined in terms of $(n-2) \times (n-2)$ determinants, etc.).

Definition 2.5.1 *Let A be a $n \times n$ matrix. For each entry $(A)_{ij}$ of A, we define the* minor M_{ij} *of $(A)_{ij}$ to be the determinant of the $(n-1) \times (n-1)$ matrix which remains when the ith row and jth column (i.e. the row and column containing $(A)_{ij}$) are deleted from A.*

EXAMPLE: Let $A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & -2 & 1 \\ -4 & 1 & -1 \end{pmatrix}$.

$M_{11}: \ M_{11} = \det \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} = -2(-1) - (1)(1) = 1$

$M_{12}: \ M_{12} = \det \begin{pmatrix} 2 & 1 \\ -4 & -1 \end{pmatrix} = 2(-1) - (1)(-4) = 2$

$M_{22}: \ M_{22} = \det \begin{pmatrix} 1 & 0 \\ -4 & -1 \end{pmatrix} = 1(-1) - (0)(-4) = -1$

$M_{23}: \ M_{23} = \det \begin{pmatrix} 1 & 3 \\ -4 & 1 \end{pmatrix} = 1(1) - (3)(-4) = 13$

$M_{32}: \ M_{32} = \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = 1(1) - (0)(2) = 1$

Definition 2.5.2 *We define the* cofactor C_{ij} *of the entry $(A)_{ij}$ of A as follows:*

$$\begin{array}{rcll} C_{ij} & = & M_{ij} & \textit{if} \quad i+j \textit{ is even} \\ C_{ij} & = & -M_{ij} & \textit{if} \quad i+j \textit{ is odd} \end{array}$$

So the cofactor C_{ij} is either equal to $+M_{ij}$ or $-M_{ij}$, depending on the position (i,j).

We have the following pattern of signs : in the positions marked "$-$", $C_{ij} = -M_{ij}$, and in the positions marked "$+$", $C_{ij} = M_{ij}$:

$$\begin{pmatrix} + & - & + & + & \cdots \\ - & + & - & \cdots & \\ + & - & + & \cdots & \\ - & + & \cdots & & \\ \vdots & & & & \end{pmatrix} \text{ e.g. for } 3\times 3 \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

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

$C_{11}:\ C_11 = M_{11} = \det\begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} = -2(-1) - (1)(1) = 1$

$C_{12}:\ C_{12} = -M_{12} = -\det\begin{pmatrix} 2 & 1 \\ -4 & -1 \end{pmatrix} = -(2(-1) - (1)(-4)) = -2$

$C_{22}:\ C_{22} = M_{22} = \det\begin{pmatrix} 1 & 0 \\ -4 & -1 \end{pmatrix} = 1(-1) - (0)(-4) = -1$

$C_{23}:\ C_{23} = -M_{23} = -\det\begin{pmatrix} 1 & 3 \\ -4 & 1 \end{pmatrix} = -(1(1) - (3)(-4)) = -13$

$C_{32}:\ C_{32} = -M_{32} = -\det\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = -(1(1) - (0)(2)) = -1$

Definition 2.5.3 *The* determinant $\det(A)$ *of the* $n \times n$ *matrix* A *is calculated as follows :*

1. *Choose a row or column of* A.
2. *For every* A_{ij} *in the chosen row or column, calculate its cofactor.*
3. *Multiply each entry of the chosen row or column by its own cofactor.*
4. *The sum of these products is* $\det(A)$.

Examples:

1. Let $A = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix}$. Find $\det(A)$.

   SOLUTION: We can calculate the determinant using cofactor expansion along the first row. Find the cofactors of the entries in the 1st row of A:

$$\begin{aligned} C_{11} &= +\det\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} = 4 \\ C_{12} &= -\det\begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} = 1 \\ C_{13} &= +\det\begin{pmatrix} -1 & 2 \\ -2 & 2 \end{pmatrix} = 2 \end{aligned}$$

   Then

$$\begin{aligned} &= A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13} \\ &= 2(4) + 1(1) + 3(2) \\ &= 15 \end{aligned}$$

NOTE: We could also do the cofactor expansion along the 2nd row:

$$\det(A) = \underbrace{A_{21}C_{21} + A_{22}C_{22} + A_{23}C_{23}}_{\text{entries of 2nd row of } A \text{ multiplied by their cofactors}}$$

$$\begin{aligned}
C_{21} &= -\det\begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} = 3 \\
C_{22} &= +\det\begin{pmatrix} 2 & 3 \\ -2 & 3 \end{pmatrix} = 12 \\
C_{23} &= -\det\begin{pmatrix} 2 & 1 \\ -2 & 2 \end{pmatrix} = -6
\end{aligned}$$

$$\det(A) = -1(3) + 2(12) + 1(-6) = 15$$

2. Let $B = \begin{pmatrix} 3 & 1 & 5 & -24 \\ 0 & 4 & 1 & -6 \\ 0 & 0 & 25 & 4 \\ 0 & 0 & 0 & -1 \end{pmatrix}$. Calculate $\det(A)$.

   SOLUTION: Use cofactor expansion along the first column to obtain :

$$\det(B) = 3\det\begin{pmatrix} 4 & 1 & -6 \\ 0 & 25 & 4 \\ 0 & 0 & -1 \end{pmatrix}$$

   On this 3×3 determinant, use the first column again. Then

$$\det(B) = 3 \times 4 \det\begin{pmatrix} 25 & 4 \\ 0 & -1 \end{pmatrix}$$

   On this 2×2 determinant, use the first column again. Then

$$\det(B) = 3 \times 4 \times 25 \times -1 = -300.$$

NOTES

1. The matrix B above is an example of an *upper triangular matrix* (all of its non-zero entries are located on or above its main diagonal). Note that $\det(B)$ is just the product of the entries along the main diagonal of B.
2. If calculating a determinant using cofactor expansion, it is usually a good idea to choose a row or column containing as many zeroes as possible.

DEFINITION: A $n \times n$ matrix A is called *upper triangular* if all entries located below (and to the left of) its main diagonal are zeroes (i.e. if $A_{ij} = 0$ whenever $i > j$).
(In the following diagram the entries indicated by "$*$" may be any real number.)

$$\begin{pmatrix} * & * & \cdots & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \vdots & \vdots \\ \vdots & \vdots & 0 & \ddots & * \\ 0 & \cdots & \cdots & 0 & * \end{pmatrix}$$

Upper triangular matrix

Theorem 2.5.4 *If A is upper triangular, then $\det A$ is the product of the entries on the main diagonal of A.*

The idea of the proof of Theorem 2.5.4 is suggested by Example 2 above - just use cofactor expansion along the first column.
CONSEQUENCE OF THEOREM 2.5.4: An upper triangular matrix is invertible if and only if none of the entries along its main diagonal is zero.

So determinants of upper triangular matrices are particularly easy to calculate. This fact can be used to calculate the determinant of any square matrix, after using elementary row operations to reduce it to row echelon form.

The following table describes the effect on the determinant of a square matrix of ERO's of the three types.

	Type of ERO	Effect on Determinant
1.	Add a multiple of one row to another row	No effect
2.	Multiply a row by a constant c	Determinant is multiplied by c
3.	Interchange two rows	Determinant changes sign

We can use these facts to find the determinant of any $n \times n$ matrix A as follows :

1. Use elementary row operations (ERO's) to obtain an upper triangular matrix A' from A.
2. Find $\det A'$ (product of entries on main diagonal).
3. Make adjustments to reverse changes to the determinant caused by ERO's in Step 1.

Example 2.5.5 *Find the determinant of the matrix*

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

SOLUTION:

Step 1: Perform elementary row operations to reduce A to upper triangular form.

$$\begin{pmatrix} 2 & 4 & 2 & 1 \\ 4 & 3 & 0 & -1 \\ -6 & 0 & 2 & 0 \\ 0 & 1 & 1 & 2 \end{pmatrix} \xrightarrow[R_3+3R_1]{R_2-2R_1} \begin{pmatrix} 2 & 4 & 2 & 1 \\ 0 & -5 & -4 & -3 \\ 0 & 12 & 8 & 3 \\ 0 & 1 & 1 & 2 \end{pmatrix} \xrightarrow[(\det\times(-1))]{R_2 \leftrightarrow R_4}$$

$$\begin{pmatrix} 2 & 4 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 12 & 8 & 3 \\ 0 & -5 & -4 & -3 \end{pmatrix} \xrightarrow[R_4+5R_2]{R_3-12R_2} \begin{pmatrix} 2 & 4 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -4 & -21 \\ 0 & 0 & 1 & 7 \end{pmatrix} \xrightarrow[(\det\times(-1))]{R_3 \leftrightarrow R_4}$$

$$\begin{pmatrix} 2 & 4 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & -4 & -21 \end{pmatrix} \xrightarrow{R_4+4R_3} \begin{pmatrix} 2 & 4 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

Step 2: Call this upper triangular matrix A'. Then $\det A' = 2 \times 1 \times 1 \times 7 = 14$.

Step 3: $\det(A') = \det(A)$ since the determinant changed sign twice during the row reduction at Step 1 but was otherwise unchanged. Thus

$$\det(A) = \det(A') = 2 \times 1 \times 1 \times 7 = 14$$

Explanation of Effects of EROs on the Determinant

	Type of ERO	Effect on Determinant
1.	Multiply a row by a constant c	Determinant is multiplied by c
2.	Add a multiple of one row to another row	No effect
3.	Interchange two rows	Determinant changes sign

1. Suppose that a square matrix A' results from multiplying Row i of A by the non-zero constant c. Using cofactor expansion by Row i to calculate $\det(A)$ and $\det(A')$, and using C_{ij} and C'_{ij} to denote the cofactors of the entries in the (i,j)-positions of A and A' resepctively, we find

$$\begin{aligned}\det(A) &= A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in} \\ \det(A') &= A'_{i1}C'_{i1} + A'_{i2}C'_{i2} + \cdots + A'_{in}C'_{in} \\ &= cA_{i1}C'_{i1} + cA_{i2}C'_{i2} + \cdots + cA_{in}C'_{in} \\ &= c(A_{i1}C'_{i1} + A_{i2}C'_{i2} + \cdots + A_{in}C'_{in}).\end{aligned}$$

Since A and A' have the same entries outside Row i, the cofactors of entries in Row i of A and Row i of A' are the same. Thus

$$\det(A') = c(A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in}) = c\det(A).$$

2. First suppose that in a square matrix B, Row 2 is a multiple of Row 1. This means that there is a real number c for which

$$B_{21} = cB_{11},\ B_{22} = cB_{12}, \ldots, B_{2n} = cB_{in}.$$

If we subtract $c\times$Row i from Row k of B, we obtain a matrix having a full row of zeroes. Thus the RREF obtainable from B by EROs is not the $n \times n$ identity matrix, as it contains at least one row full of zeroes. Hence B is not invertible and $\det(B) = 0$.

Thus : any matrix in which one row is a multiple of another has determinant zero.

Now suppose that A' is obtained from A by adding $c\times$Row k to Row i. So the entries in Row i of A' are

$$cA_{k_1} + A_{i1},\ cA_{k2} + A_{i2},\ \ldots, cA_{kn} + A_{in}$$

Outside Row i, A and A' have the same entries. Hence the cofactors of the entries in Row i of A and A' are the same. We let C_{ij} denote the cofactor of the entry in the (i,j) position of either of these matrices. Now if we calculate $\det(A)$ and $\det(A')$ using Row i we obtain

$$\begin{aligned}\det(A) &= A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in} \\ \det(A') &= (cA_{k_1} + A_{i1})C_{i1} + (cA_{k2} + A_{i2})C_{i2} + \cdots + (cA_{kn} + A_{in})C_{in} \\ &= (cA_{k1}C_{i1} + cA_{k2}C_{i2} + \cdots + cA_{in}C_{in}) + (A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in})\end{aligned}$$

Now

$$cA_{k1}C_{i1} + cA_{k2}C_{i2} + \cdots + cA_{in}C_{in}$$

is the determinant of a matrix in which Row k is just $c\times$Row i, hence this number is zero by the remarks above. Thus

$$\det(A') = A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in} = \det(A).$$

3. We omit the full details, but consider the case where A' is obtained from A by swapping the first two rows. Use M_{ij} and M'_{ij} to denote minors of A and A', and C_{ij} and C'_{ij} respectively for cofactors of A and A'. Using the first row of A and the second row of A' we obtain

$$\begin{aligned}\det(A) &= A_{11}C_{11} + A_{12}C_{12} + \cdots + A_{1n}C_{1n} \\ &= A_{11}M_{11} - A_{12}M_{12} + \cdots \pm A_{1n}C_{1n}. \\ \det(A') &= A'_{21}C'_{21} + A'_{22}C'_{22} + \cdots + A'_{2n}C'_{2n} \\ &= A_{11}C'_{21} + A_{12}C'_{22} + \cdots + A_{1n}C'_{2n} \\ &= -A_{11}M'_{21} + A_{12}M'_{22} - \cdots \pm A_{1n}M'_{1n} \\ &= -A_{11}M_{11} + A_{12}M_{12} + \cdots \pm A_{1n}M_{1n} \\ &= -\det(A).\end{aligned}$$

The case of general row swaps is messier but basically similar.

Chapter 3

Eigenvalues and Eigenvectors

3.1 Powers of Matrices

Definition 3.1.1 *Let A be a square matrix ($n \times n$). If k is a positive integer, then A^k denotes the matrix*

$$\underbrace{A \times A \times \cdots \times A}_{k \text{ times}}.$$

Calculating matrix powers using the definition of matrix multiplication is computationally very laborious. One of the topics that we will discuss in this chapter is how powers of matrices may be calculated efficiently.

First we look at a reason for calculating such powers at all.

Example 3.1.2 *Suppose that two competing Broadband companies, A and B, each currently have 50% of the market share. Suppose that over each year, A captures 10% of B's share of the market, and B captures 20% of A's share. What is each company's market share after 5 years?*

SOLUTION: Let a_n and b_n denote the proportion of the market held by A and B respectively at the end of the nth year. We have $a_0 = b_0 = 0.5$ (beginning of Year 1 = end of Year 0).

Now a_{n+1} and b_{n+1} depend on a_n and b_n according to

$$\begin{array}{rcrcr} a_{n+1} & = & 0.8a_n & + & 0.1b_n \\ b_{n+1} & = & 0.2a_n & + & 0.9b_n \end{array}$$

We can write this in matrix form as follows

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

We define $A = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$. Then

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = A \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = A \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \quad \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = A \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = A^2 \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

In general

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = A^n \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}^n \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

So if we had an efficient way to calculate A^n, we could use it to calculate a_n and b_n.

3.2 The Characteristic Equation of a Matrix

Let A be a 2×2 matrix; for example

$$A = \begin{pmatrix} 2 & 8 \\ 3 & -3 \end{pmatrix}.$$

If $\vec{v}$ is a vector in $\mathbb{R}^2$, e.g. $\vec{v} = (2,3)$, then we can think of the components of $\vec{v}$ as the entries of a column vector (i.e. a 2×1 matrix). Thus

$$[2,3] \leftrightarrow \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

If we multiply this vector on the left by the matrix A, we get another column vector with two entries :

$$A\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 & 8 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2(2)+8(3) \\ 3(2)+(-3)(3) \end{pmatrix} = \begin{pmatrix} 28 \\ -3 \end{pmatrix}$$

So multiplication on the left by the 2×2 matrix A is a function sending the set of 2×1 column vectors to itself - or, if we wish, we can think of it as a function from the set of vectors in $\mathbb{R}^2$ to itself.

NOTE: In fact this function is an example of a *linear transformation* from $\mathbb{R}^2$ into itself. Linear transformations are functions which have certain interesting geometric properties. Basically they are functions which can be represented in this way by matrices.

In general, if v is a column vector with two entries, then Av is a another vector (with two entries), which typically does not resemble v at all. For example if $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ then

$$Av = \begin{pmatrix} 2 & 8 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 18 \\ -3 \end{pmatrix}$$

However, suppose $v = \begin{pmatrix} 8 \\ 3 \end{pmatrix}$. Then

$$Av = \begin{pmatrix} 2 & 8 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} 8 \\ 3 \end{pmatrix} = \begin{pmatrix} 40 \\ 15 \end{pmatrix} = 5\begin{pmatrix} 8 \\ 3 \end{pmatrix}$$

i.e. $A\begin{pmatrix} 8 \\ 3 \end{pmatrix} = 5\begin{pmatrix} 8 \\ 3 \end{pmatrix}$, or

Multiplying the vector $\begin{pmatrix} 8 \\ 3 \end{pmatrix}$ *(on the left) by the matrix* $\begin{pmatrix} 2 & 8 \\ 3 & -3 \end{pmatrix}$
is the same as multiplying it by 5.

TERMINOLOGY: $\begin{pmatrix} 8 \\ 3 \end{pmatrix}$ is called an *eigenvector* for the matrix $A = \begin{pmatrix} 2 & 8 \\ 3 & -3 \end{pmatrix}$ with corresponding *eigenvalue* 5.

Definition 3.2.1 *Let A be a $n \times n$ matrix, and let v be a non-zero column vector with n entries (so not all of the entries of v are zero). Then v is called an* eigenvector *for A if*

$$Av = \lambda v,$$

where λ is some real number.

In this situation λ is called an *eigenvalue* for A, and v is said to *correspond* to λ.
NOTE: "λ" is the symbol for the Greek letter *lambda*. It is conventional to use this symbol to denote an eigenvalue.

Example 3.2.2 *If* $A = \begin{pmatrix} -1 & 1 \\ -2 & -4 \end{pmatrix}$ *and* $v = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, *then*

$$Av = \begin{pmatrix} -1 & 1 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \end{pmatrix} = -3\begin{pmatrix} 1 \\ -2 \end{pmatrix} = -3v$$

Thus $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ *is an eigenvector for the matrix* $\begin{pmatrix} -1 & 1 \\ -2 & -4 \end{pmatrix}$ *corresponding to the eigenvalue* -3.

QUESTION: Given a $n \times n$ matrix A, how can we find its eigenvalues and eigenvectors?
ANSWER: We are looking for column vectors v and real numbers λ satisfying

$$\begin{aligned} Av &= \lambda v \\ \text{i.e. } \lambda v - Av &= \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ \Longrightarrow \lambda I_n v - Av &= \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ \Longrightarrow \underbrace{(\lambda I_n - A)}_{\text{a } n\times n \text{ matrix}} v &= \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \end{aligned}$$

This may be regarded as a system of linear equations in which the coefficient matrix is $\lambda I_n - A$ and the variables are the n entries of the column vector v, which we can denote by $x_1, \dots, x_n$. We are looking for solutions to

$$(\lambda I_n - A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

This system always has at least one solution : namely $x_1 = x_2 = \dots = x_n = 0$ - all entries of v are zero. However this solution does *not* give an eigenvector since eigenvectors must be non-zero.

The system can have additional solutions only if $\det(\lambda I_n - A) = 0$ (otherwise if the square matrix $\lambda I_n - A$ is invertible, the system will have $x_1 = x_2 = \dots = x_n = 0$ as its *unique* solution).
CONCLUSION: The *eigenvalues* of A are those values of λ for which $\det(\lambda I_n - A) = 0$.

Example 3.2.3 *Let* $A = \begin{pmatrix} 10 & -8 \\ 4 & -2 \end{pmatrix}$. *Find all eigenvalues of* A *and find an eigenvector corresponding to each eigenvalue.*

SOLUTION: We need to find all values of λ for which $\det(\lambda I_2 - A) = 0$.

$$\begin{aligned} \lambda I_2 - A &= \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 10 & -8 \\ 4 & -2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} - \begin{pmatrix} 10 & -8 \\ 4 & -2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda - 10 & 8 \\ -4 & \lambda + 2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \det(\lambda I_2 - A) &= (\lambda - 10)(\lambda + 2) - 8(-4) \\ &= \lambda^2 - 10\lambda + 2\lambda - 20 + 32 \\ &= \lambda^2 - 8\lambda + 12 \end{aligned}$$

So $\det(\lambda I_2 - A)$ is a polynomial of degree 2 in λ. The eigenvalues of A are those values of λ for which

$$\det(\lambda I_2 - A) = 0$$

i.e. $\lambda^2 - 8\lambda + 12 = 0 \Longrightarrow (\lambda - 6)(\lambda - 2) = 0,\ \ \lambda = 6$ or $\lambda = 2$
Eigenvalues of A : 6, 2.

To find an eigenvector of A corresponding to $\lambda = 6$, we need a vector $\begin{pmatrix} x \\ y \end{pmatrix}$ for which

$$\begin{aligned} A\begin{pmatrix} x \\ y \end{pmatrix} &= 6\begin{pmatrix} x \\ y \end{pmatrix} \\ \text{i.e.} \begin{pmatrix} 10 & -8 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= 6\begin{pmatrix} x \\ y \end{pmatrix} \\ \Longrightarrow \begin{pmatrix} 10x - 8y \\ 4x - 2y \end{pmatrix} &= \begin{pmatrix} 6x \\ 6y \end{pmatrix} \end{aligned}$$

$$\Longrightarrow 10x - 8y = 6x \quad \text{and} \quad 4x - 2y = 6y$$

Both of these equations say $x - 2y = 0$; hence *any* non-zero vector $\begin{pmatrix} x \\ y \end{pmatrix}$ in which $x = 2y$ is an eigenvector for A corresponding to the eigenvalue 6. For example we can take $y = 1$, $x = 2$ to obtain the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

EXERCISES:

1. Show that $\begin{pmatrix} 10 & -8 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 6\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.
2. Find an eigenvector for A corresponding to the other eigenvalue $\lambda = 2$.

Definition 3.2.4 *Let A be a square matrix $(n \times n)$. The* characteristic polynomial *of A is the determinant of the $n \times n$ matrix $\lambda I_n - A$. This is a polynomial of degree n in λ.*

Example 3.2.5

(a) Let $A = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$. Then

$$\lambda I_2 - A = \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} - \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \lambda - 4 & 1 \\ -2 & \lambda - 1 \end{pmatrix}$$

$\det(\lambda I_2 - A) = (\lambda - 4)(\lambda - 1) - 1(-2) = \lambda^2 - 5\lambda + 6$

<u>Characteristic Polynomial of A</u>: $\lambda^2 - 5\lambda + 6$.

(b) Let $B = \begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{pmatrix}$.

$$\lambda I_3 - B = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} - \begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} \lambda - 5 & -6 & -2 \\ 0 & \lambda + 1 & 8 \\ -1 & 0 & \lambda + 2 \end{pmatrix}$$

We can calculate $\det(\lambda I_3 - B)$ using cofactor expansion along the first row.

$$\begin{aligned} \det(\lambda I_3 - B) &= (\lambda - 5)[(\lambda + 1)(\lambda + 2) - (0)(8)] \\ &\quad -(-6)[0(\lambda + 2) - 8(-1)] + (-2)[0(0) - (-1)(\lambda + 1)] \\ &= (\lambda - 5)(\lambda^2 + 3\lambda + 2) + 6(8) - 2(\lambda + 1) \\ &= \lambda^3 - 2\lambda^2 - 13\lambda - 10 + 48 - 2\lambda - 2 \\ &= \lambda^3 - 2\lambda^2 - 15\lambda + 36. \end{aligned}$$

<u>Characteristic polynomial of B</u> : $\lambda^3 - 2\lambda^2 - 15\lambda + 36$.

As we saw, the eigenvalues of a matrix A are those values of λ for which $\det(\lambda I - A) = 0$; i.e., the eigenvalues of A are the *roots* of the characteristic polynomial.

Example 3.2.6 *Find the eigenvalues of the matrices* A *and* B *of Example 6.2.2.*

(a) $A = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$

Characteristic Equation : $\lambda^2 - 5\lambda + 6 = 0 \Longrightarrow (\lambda - 3)(\lambda - 2) = 0$

Eigenvalues of A: $\lambda = 3,\ \lambda = 2$.

(b) $B = \begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & 2 \end{pmatrix}$

Characteristic Equation: $\lambda^3 - 2\lambda^2 - 15\lambda + 36 = 0$

To find solutions to this equation we need to factor the characteristic polynomial, which is cubic in λ (in general solving a cubic equation like this is not an easy task unless we can factorize). First we try to find an integer root.

Fact: The only possible integer roots of a polynomial are factors of its constant term.

So in this example the only possible candidates for an integer root of the characteristic polynomial $p(\lambda) = \lambda^3 - 2\lambda^2 - 15\lambda + 36$ are the integer factors of 36 : i.e.

$$\pm 1,\ \pm 2,\ \pm 3,\ \pm 4,\ \pm 6,\ \pm 9,\ \pm 12,\ \pm 18,\ \pm 36$$

Try some of these :

$$\begin{aligned} p(1) &= 1^3 - 2(1)^2 - 15(1) + 36 \neq 0 \\ p(2) &= 2^3 - 2(2)^2 - 15(2) + 36 \neq 0 \\ p(3) &= 3^3 - 2(3)^2 - 15(3) + 36 = 0 \end{aligned}$$

$\Longrightarrow$ 3 is a root of $p(\lambda)$, and $(\lambda - 3)$ is a factor of $p(\lambda)$. Then

$$p(\lambda) = \lambda^3 - 2\lambda^2 - 15\lambda + 36 = (\lambda - 3)(\lambda^2 + a\lambda - 12)$$

To find a, look at the coefficients of λ^2 (or λ) on the left and right

$$\lambda^2 :\ -2 = -3 + a \Longrightarrow a = 1$$

$$\begin{aligned} \lambda^3 - 2\lambda^2 - 15\lambda + 36 &= (\lambda - 3)(\lambda^2 + \lambda - 12) \\ &= (\lambda - 3)(\lambda - 3)(\lambda + 4) \\ &= (\lambda - 3)^2(\lambda + 4) \end{aligned}$$

Eigenvalues of B: $\lambda = 3$ (occurring twice), $\lambda = -4$.

We conclude this section by calculating eigenvectors of B corresponding to these eigenvalues.

Example 3.2.7 *Let* $B = \begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{pmatrix}$

From Example 3.2.5, the eigenvalues of B *are* $\lambda = 3$ *(occurring twice),* $\lambda = -4$.
Find an eigenvector of B *corresponding to the eigenvalue* $\lambda = -4$.

SOLUTION: We need a column vector $v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, with entries not all zero, for which

$$\begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = -4 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\text{i.e.}\quad \begin{pmatrix} 5x_1 & + & 6x_2 & + & 2x_3 \\ & - & x_2 & - & 8x_3 \\ x_1 & & & - & 2x_3 \end{pmatrix} = \begin{pmatrix} -4x_1 \\ -4x_2 \\ -4x_3 \end{pmatrix}$$

$$\Longrightarrow \begin{array}{ccccccc} 5x_1 & + & 6x_2 & + & 2x_3 & = & -4x_1 \\ & - & x_2 & - & 8x_3 & = & -4x_2 \\ x_1 & & & - & 2x_3 & = & -4x_3 \end{array} \Longrightarrow \underbrace{\begin{array}{ccccccc} 9x_1 & + & 6x_2 & + & 2x_3 & = & 0 \\ & & 3x_2 & - & 8x_3 & = & 0 \\ x_1 & & & + & 2x_3 & = & 0 \end{array}}_{\text{system of 3 equations in } x_1, x_2, x_3}$$

So we need to solve the system of linear equations with augmented matrix

$$\begin{pmatrix} 9 & 6 & 2 & 0 \\ 0 & 3 & -8 & 0 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

NOTE: The coefficient matrix here is just $B - (-4)I_3$ i.e.

$$\begin{pmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{pmatrix} - \begin{pmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

To find solutions to the system :

$$\begin{pmatrix} 9 & 6 & 2 & 0 \\ 0 & 3 & -8 & 0 \\ 1 & 0 & 2 & 0 \end{pmatrix} \xrightarrow{R3 \leftrightarrow R1} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & -8 & 0 \\ 9 & 6 & 2 & 0 \end{pmatrix} \xrightarrow{R3 - 9 \times R1}$$

$$\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & -8 & 0 \\ 0 & 6 & -16 & 0 \end{pmatrix} \xrightarrow{R3 - 2 \times R2} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R2 \times \frac{1}{3}}$$

$$\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -\frac{8}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad : \text{RREF}$$

The variable x_3 is free : let $x_3 = t$. Then

$$\begin{array}{rcr} x_1 + 2x_3 = 0 & \Longrightarrow & x_1 = -2t \\ x_2 - \frac{8}{3}x_3 = 0 & \Longrightarrow & x_2 = \frac{8}{3}t \end{array}$$

For example if we take $t = 3$ we find $x_1 = -6$ and $x_2 = 8$. Hence $v = \begin{pmatrix} -6 \\ 8 \\ 3 \end{pmatrix}$ is an eigenvector for B corresponding to $\lambda = -4$

EXERCISE: Check that $Bv = -4v$.

NOTES:

1. To find an eigenvector v of a $n \times n$ matrix A corresponding to the eigenvalue λ : solve the system

$$(A - \lambda I_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

   i.e. the system whose coefficient matrix is $A - \lambda I_n$ and in which the constant term (on the right in each equation) is 0.

2. If v is an eigenvector of a square matrix A, corresponding to the eigenvalue λ, and if $k \neq 0$ is a real number, then kv is also an eigenvector of A corresponding to λ, since

$$A(kv) = k(Av) = k(\lambda v) = \lambda(kv)$$

In the above example any (non-zero) scalar multiple of $\begin{pmatrix} -6 \\ 8 \\ 3 \end{pmatrix}$ is an eigenvector of A corresponding to $\lambda = -4$ (these arise from different choices of value for the free variable t in the solution of the relevant system of equations).

Example 3.2.8 *Find an eigenvector of* B *corresponding to the eigenvalue* $\lambda = 3$.

SOLUTION: We need to solve the system whose augmented matrix consists of $B - 3I_3$ and a fourth column all of whose entries are zero.

$$B - 3I_3 = \begin{pmatrix} 2 & 6 & 2 \\ 0 & -4 & -8 \\ 1 & 0 & -5 \end{pmatrix}$$

(obtained by subtracting 3 from each of the entries on the main diagonal of B and leaving the other entries unchanged).

We apply elementary row operations to the augmented matrix of the system :

$$\begin{pmatrix} 2 & 6 & 2 & 0 \\ 0 & -4 & -8 & 0 \\ 1 & 0 & -5 & 0 \end{pmatrix} \begin{array}{c} R1 \times \frac{1}{2} \\ \longrightarrow \\ R2 \times (-\frac{1}{4}) \end{array} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 0 & -5 & 0 \end{pmatrix} \begin{array}{c} R3 - R1 \\ \longrightarrow \end{array}$$

$$\begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -3 & -6 & 0 \end{pmatrix} \begin{array}{c} R3 + 3 \times R2 \\ \longrightarrow \end{array} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{c} R1 - 3 \times R2 \\ \longrightarrow \end{array}$$

$$\begin{pmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad : \mathrm{RREF}$$

Let $x_3 = t$. Then

$$\begin{aligned} x_1 - 5x_3 = 0 &\implies x_1 = 5t \\ x_2 + 2x_3 = 0 &\implies x_2 = -2t \end{aligned}$$

Eigenvectors are given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5t \\ -2t \\ t \end{pmatrix}$$

for $t \in \mathbb{R},\ t \neq 0$. For example of we choose $t = 1$ we find that $v = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$ is an eigenvector for B corresponding to $\lambda = 3$. (Exercise: Check this).

3.3 Diagonalization

Let $A = \begin{pmatrix} -4 & 1 \\ 4 & -4 \end{pmatrix}$. Then $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ are eigenvectors of A, with corresponding eigenvalues -2 and -6 respectively (check). This means

$$\begin{pmatrix} -4 & 1 \\ 4 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} -4 & 1 \\ 4 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = -6\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Thus

$$\begin{pmatrix} -4 & 1 \\ 4 & -4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} = \left(-2\begin{pmatrix} 1 \\ -2 \end{pmatrix} \; -6\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right) = \begin{pmatrix} -2 & -6 \\ -4 & 12 \end{pmatrix}$$

We have

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

(Think about this). Thus $AE = ED$ where $E = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$ has the eigenvectors of A as columns and $D = \begin{pmatrix} -2 & 0 \\ 0 & -6 \end{pmatrix}$ is the *diagonal* matrix having the eigenvalues of A on the main diagonal, in the order in which their corresponding eigenvectors appear as columns of E.

Definition 3.3.1 *A* $n \times n$ *matrix is* A diagonal *if all of its non-zero entries are located on its main diagonal, i.e. if* $A_{ij} = 0$ *whenever* $i \neq j$.

Diagonal matrices are particularly easy to handle computationally. If A and B are diagonal $n \times n$ matrices then the product AB is obtained from A and B by simply multiplying entries in corresponding positions along the diagonal, and $AB = BA$.
If A is a diagonal matrix and k is a positive integer, then A^k is obtained from A by replacing each entry on the main diagonal with its kth power.
Back to our Example : We have $AE = ED$. Note that $\det(E) \neq 0$ so E is invertible. Thus

$$\begin{aligned} AE &= ED \\ \Longrightarrow AEE^{-1} &= EDE^{-1} \\ \Longrightarrow A &= EDE^{-1}. \end{aligned}$$

It is convenient to write A in this form if for some reason we need to calculate powers of A. Note for example that

$$\begin{aligned} A^3 &= (EDE^{-1})(EDE^{-1})(EDE^{-1}) \\ &= EDI_2DI_2DE^{-1} \\ &= ED^3E^{-1} \\ &= E\begin{pmatrix} (-2)^3 & 0 \\ 0 & (-6)^3 \end{pmatrix}E^{-1}. \end{aligned}$$

In general $A^n = E\begin{pmatrix} (-2)^n & 0 \\ 0 & (-6)^n \end{pmatrix}E^{-1}$, for any positive integer n. (In fact this is true for negative integers too if we interpret A^{-n} to mean the nth power of the inverse A^{-1} of A).

Example 3.3.2 *Solve the recurrence relation*

$$\begin{aligned} x_{n+1} &= -4x_n + 1y_n \\ y_{n+1} &= 4x_n - 4y_n \end{aligned}$$

given that $x_0 = 1$, $y_0 = 1$.

NOTE: this means we have sequences $x_0, x_1, \ldots$ and $y_0, y_1, \ldots$ defined by the above relations. If for some n we know x_n and y_n, the relations tell us how to calculate x_{n+1} and y_{n+1}.
For example

$$\begin{aligned} x_1 &= -4x_0 + y_0 = -4(1) + 1 = -3 \\ y_1 &= 4x_0 - 4y_0 = 4(1) - 4(1) = 0 \\ \\ x_2 &= -4x_1 + y_1 = -4(-3) + 0 = 12 \\ y_2 &= 4x_1 - 4y_1 = 4(-3) - 4(0) = -12. \end{aligned}$$

What it means to *solve* the recurrence relation is to give a pair of formulae stating exactly how x_n and y_n depend on n, so that for example x_{100} could be calculated without knowledge of x_{99} and y_{99}. Of course it is not obvious in advance that such formulae exist.

SOLUTION OF THE PROBLEM:
The relations can be written in matrix form as

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} -4x_n + 1y_n \\ 4x_n - 4y_n \end{pmatrix} = \begin{pmatrix} -4 & 1 \\ 4 & -4 \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = A \begin{pmatrix} x_n \\ y_n \end{pmatrix},$$

where A is the matrix $\begin{pmatrix} -4 & 1 \\ 4 & -4 \end{pmatrix}$. Thus

$$\begin{aligned} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} &= A \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} &= A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = A \left(A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = A^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} &= A \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = A \left(A^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = A^3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ etc.} \end{aligned}$$

In general $\begin{pmatrix} x_n \\ y_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
To obtain general formulae for x_n and y_n we need a general formula for A^n. We have

$$A^n = (EDE^{-1})^n = ED^nE^{-1}$$

where $E = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$ and $D = \begin{pmatrix} -2 & 0 \\ 0 & -6 \end{pmatrix}$.
Note

$$E^{-1} = -\frac{1}{4} \begin{pmatrix} -2 & -1 \\ -2 & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 2 & 1 \\ 2 & -1 \end{pmatrix}.$$

Thus

$$\begin{aligned} A^n &= \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} (-2)^n & 0 \\ 0 & (-6)^n \end{pmatrix} \frac{1}{4} \begin{pmatrix} 2 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} (-2)^n & (-6)^n \\ 2(-2)^n & -2(-6)^n \end{pmatrix} \frac{1}{4} \begin{pmatrix} 2 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} (-2)^n(2) + (-6)^n(2) & (-2)^n - (-6)^n \\ 4(-2)^n - 4(-6)^n & 2(-2)^n + 2(-6)^n \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \frac{1}{4} \begin{pmatrix} (-2)^n(2) + (-6)^n(2) & (-2)^n - (-6)^n \\ 4(-2)^n - 4(-6)^n & 2(-2)^n + 2(-6)^n \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 3(-2)^n + (-6)^n \\ 6(-2)^n - 2(-6)^n \end{pmatrix} \end{aligned}$$

We conclude that

$$\begin{aligned} x_n &= \frac{3}{4}(-2)^n + \frac{1}{4}(-6)^n \\ y_n &= \frac{3}{2}(-2)^n - \frac{1}{2}(-6)^n \end{aligned}$$

for $n \geqslant 0$.
(This is easily verified for small values of n using the recurrence relations). See Problem Sheet 3 for more problems of this type.

Definition 3.3.3 *The $n \times n$ matrix A is* diagonalizable *(or diagonable) if there exists an invertible matrix E for which*

$$E^{-1}AE$$

is diagonal.

We have already seen that if E is a matrix whose columns are eigenvectors of A, then $AE = ED$, where D is the diagonal matrix whose entry in the (i, i) position is the eigenvalue of A to which the ith column of E corresponds as an eigenvector of A. If E is invertible then $E^{-1}AE = D$ and A is diagonalizable. Hence we have the following statement

1. If there exists an invertible matrix whose columns are eigenvectors of A, then A is diagonalizable.

On the other hand, suppose that A is diagonalizable. Then there exists an invertible $n \times n$ matrix E and a diagonal matrix D whose entry in the (i, i) position can be denoted d_i, for which

$$D = E^{-1}AE.$$

This means $ED = AE$, so

$$E \begin{pmatrix} d_1 & \dots & & \\ \vdots & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} = AE$$

. Looking at the jth column of each of these products shows that

$$\begin{pmatrix} E_{1j} \\ E_{2j} \\ \vdots \\ E_{nj} \end{pmatrix} d_j = A \begin{pmatrix} E_{1j} \\ E_{2j} \\ \vdots \\ E_{nj} \end{pmatrix}.$$

Thus the jth column of E is an eigenvector of A (with corresponding eigenvalue d_j). So

2. If the $n \times n$ matrix A is diagonalizable, then there exists an invertible matrix whose columns are eigenvectors of A.

Putting this together with 1. above gives

Theorem 3.3.4 *The square matrix A is diagonalizable if and only if there exists an invertible matrix having eigenvectors of A as columns.*

It is not true that every square matrix is diagonalizable.

Example 3.3.5 *Let* $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$.

Then

$$\det(\lambda I - A) = \lambda^2 - 6\lambda + 9 = (\lambda - 3)^2.$$

So $\lambda = 3$ is the only eigenvalue of A and it occurs twice.
Eigenvectors : Suppose $A\binom{x}{y} = 3\binom{x}{y}$. Then

$$\begin{array}{rcrcl} 2x & - & y & = & 3x \\ x & + & 4y & = & 3y \end{array} \implies x + y = 0, \ \ x = -y.$$

So every eigenvector of A has the form $\begin{pmatrix} -y \\ y \end{pmatrix}$ for some non-zero real number y. Thus every 2×2 matrix having eigenvectors of A as columns as of the form $\begin{pmatrix} -a & -b \\ a & b \end{pmatrix}$ for some non-zero real numbers a and b. The determinant of such a matrix is $-ab - (-ab) = 0$. Thus no matrix having eigenvectors of A as columns is invertible, and A is not diagonalizable.

Although the above example shows that not all square matrices are diagonalizable, we do have the following fact.

Theorem 3.3.6 *Suppose that the $n \times n$ matrix A has n* distinct *eigenvalues $\lambda_1, \dots, \lambda_n$. If E is a matrix whose columns are eigenvectors of A corresponding to the different eigenvalues, then E is invertible. Thus A is diagonalizable.*

3.4 Further Properties of Eigenvectors and Eigenvalues

1. Suppose that v is an eigenvector of the square matrix A, corresponding to the eigenvalue λ. Then so is kv for any non-zero real number k. To see this note that

$$A(kv) = k\,A(v) = k(\lambda v) = (\lambda k)v = \lambda(kv).$$

2. If v is an eigenvector of A corresponding to the eigenvalue λ, then v is also an eigenvector of A^2 and the eigenvalue to which it corresponds is λ^2. To see this note

$$A^2(v) = A(Av) = A(\lambda v) = \lambda(Av) = \lambda(\lambda v) = \lambda^2 v.$$

   Similarly v is an eigenvector of A^n for any positive integer n, corresponding to the eigenvalue λ^n.

3. For any square matrix A, let A^T denote the transpose of A. Then $\det(A) = \det(A^T)$. It follows that

$$\det(\lambda I - A) = \det(\lambda I - A)^T = \det(\lambda I - A^T).$$

   Thus A and A^T have the same characteristic equation, and they have the same eigenvalues. (However there is no general connection between the eigenvectors of A^T and those of A).

4. Suppose that A has the property that for each of its rows, the sum of the entries in that row is the same number s. For example if

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

   the row sums of A are all equal to 10.

   Then

$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} s \\ s \\ \vdots \\ s \end{pmatrix} = s \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

   Thus the vector whose entries are all equal to 1 is an eigenvector of A corresponding to the eigenvalue s. In particular the common row sum s is an eigenvalue of A.

5. On the other hand suppose that the sum of the entries of every *column* of A is the same number k. Then by 4 above k is an eigenvalue of A^T and hence by 3 above k is an eigenvalue of A. In particular if the sum of the entries in every column of A is 1, then 1 is an eigenvalue of A.

Chapter 4

Markov Processes

4.1 Markov Processes and Markov Chains

Recall the following example from Section 3.1.
Two competing Broadband companies, A and B, each currently have 50% of the market share. Suppose that over each year, A captures 10% of B's share of the market, and B captures 20% of A's share.

This situation can be modelled as follows. Let a_n and b_n denote the proportion of the market held by A and N respectively at the end of the nth year. We have $a_0 = b_0 = 0.5$ (beginning of Year 1 = end of Year 0).

Now a_{n+1} and b_{n+1} depend on a_n and b_n according to

$$\begin{array}{rcrcr} a_{n+1} & = & 0.8a_n & + & 0.1b_n \\ b_{n+1} & = & 0.2a_n & + & 0.9b_n \end{array}$$

We can write this in matrix form as follows

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}.$$

We define $M = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$. Note that every entry in M is non-negative and that the sum of the entries in each column is 1. This is no accident since the entries in the first column of M are the respective proportions of A's market share that are retained and lost respectively by A from one year to the next. Column 2 contains similar data for B.

Definition 4.1.1 *A* stochastic matrix *is a square matrix with the following properties :*

(i) All entries are non-negative.

(ii) The sum of the entries in each column is 1.

So the matrix M of our example is stochastic.

Returning to the example, if we let v_n denote the vector $\begin{pmatrix} a_n \\ b_n \end{pmatrix}$ describing the position at the end of year n, we have

$$v_0 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix},\ v_1 = Mv_0,\ v_{n+1} = Mv_n.$$

Note that the sum of the entries in each v_i is 1.

Definition 4.1.2 *A column vector with non-negative entries whose sum is 1 is called a* probability vector.

It is not difficult to see that if v is a probability vector and A is a stochastic matrix, then Av is a probability vector. In our example, the sequence $v_0, v_1, v_2, \dots$ of probability vectors is an example of a *Markov Chain*. In algebraic terms a Markov chain is determined by a probability vector v and a stochastic matrix A (called the transition matrix of the process or chain). The chain itself is the sequence

$$v_0,\ v_1 = Av_0,\ v_2 = Av_1, \ldots$$

More generally a Markov process is a process in which the probability of observing a particular state at a given observation period depends only on the state observed at the preceding observation period.

Remark: Suppose that A is a stochastic matrix. Then from Item 5 in Section 3.4 it follows that 1 is an eigenvalue of A (all the columns of A sum to 1). The transition matrix in our example is

$$M = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}.$$

Eigenvectors of M corresponding to the eigenvalue 1 are non-zero vectors $\begin{pmatrix} x \\ y \end{pmatrix}$ for which

$$\begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Thus

$$\left.\begin{array}{rcrcl} 0.8x & + & 0.1y & = & x \\ 0.2x & + & 0.9y & = & y \end{array}\right\} \Longrightarrow y = 2x.$$

So any non-zero vector of the form $\begin{pmatrix} x \\ 2x \end{pmatrix}$ is an eigenvector of M corresponding to the eigenvalue 1. Amongst all these vectors exactly one is a probability vector, namely the one with $x + 2x = 1$, i.e. $x = \frac{1}{3}$. This eigenvector is $\begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$

The Markov process in our example is $v_0, v_1, v_2, \ldots$, where $v_0 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$ and $v_{i+1} = Mv_i$. We can observe

$$\begin{array}{rcccc} v_5 & = & M^5 v_0 & \approx & \begin{pmatrix} 0.3613 \\ 0.6887 \end{pmatrix} \\ \\ v_{10} & = & M^{10} v_0 & \approx & \begin{pmatrix} 0.3380 \\ 0.6620 \end{pmatrix} \\ \\ v_{20} & = & M^{20} v_0 & \approx & \begin{pmatrix} 0.3335 \\ 0.6665 \end{pmatrix} \\ \\ v_{30} & = & M^{30} v_0 & \approx & \begin{pmatrix} 0.3333 \\ 0.6667 \end{pmatrix} \end{array}$$

So it appears that the vectors in the Markov chain approach the eigenvector $\begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$ of M as the process develops. This vector is called the *steady state* of the process.

This example is indicative of a general principle.

Definition 4.1.3 *A stochastic $n \times n$ matrix M is called* regular *if M itself or some power of M has all entries positive (i.e. no zero entries).*

Example

- $M = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$ is a regular stochastic matrix.
- $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is a stochastic matrix but it is not regular :

$$A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ A^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A, \text{ etc}$$

  The positive powers of A just alternate between I_2 and A itself. So no positive integer power of A is without zero entries.

Theorem 4.1.4 *Suppose that A is a regular stochastic $n \times n$ matrix. Then*

- *There is a unique probability vector v for which $Av = v$.*
- *If u_0 is any probability vector then the Markov chain $u_0, u_1, \dots$ defined for $i \geqslant 1$ by $u_i = Au_{i-1}$ converges to v.*
  (This means that for $1 \leqslant i \leqslant n$, the sequence of the ith entries of $u_0, u_1, u_2, \dots$ converges to the ith entry of v).

Notes

1. Theorem 4.1.4 says that if a Markov process has a regular transition matrix, the process will converge to the steady state v regardless of the initial position.
2. Theorem 4.1.4 does not apply when the transition matrix is not regular. For example if $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $u_0 = \begin{pmatrix} a \\ b \end{pmatrix}$ $(a \neq b)$ is a probability vector, consider the Markov chain with initial state u_0 that has A as a transition matrix.
$$u_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a \end{pmatrix},\ u_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$
This Markov chain will switch between $\begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} b \\ a \end{pmatrix}$ and not converge to a steady state.

Example 4.1.5 *(Summer 2004 Q4) An airline has planes based in Knock, Cork and Shannon. Each week $\frac{1}{4}$ of the planes originally based in Galway end up in Knock and $\frac{1}{3}$ end up in Shannon - the rest return to Galway.*

Of the planes starting the week in Knock, $\frac{1}{5}$ end up in Galway and $\frac{1}{10}$ in Shannon. The rest return to Knock.

Finally, of the planes starting the week in Shannon, $\frac{1}{5}$ end up in Galway and $\frac{1}{5}$ in Knock, the rest returning to Shannon.

Find the steady state of this Markov process.

Solution: The Markov process is a sequence $v_1, v_2, \dots$ of column vectors of length 3. The entries of the vector v_i are the proportions of the airline's fleet that are located at Galway, Knock and Shannon at the end of Week i. They are related by

$$v_{i+1} = Mv_i,$$

where M is the transition matrix of the process.
Step 1: Write down the transition matrix. If we let $g_i,\ k_i,\ s_i$ denote the proportion of the airline's fleet at Galway, Knock and Shannon after Week i, we have

$$\begin{aligned} g_{i+1} &= \tfrac{5}{12}g_i + \tfrac{1}{5}k_i + \tfrac{1}{5}s_i \\ k_{i+1} &= \tfrac{1}{4}g_i + \tfrac{7}{10}k_i + \tfrac{1}{5}s_i \\ s_{i+1} &= \tfrac{1}{3}g_i + \tfrac{1}{10}k_i + \tfrac{3}{5}s_i \end{aligned}$$

Thus

$$v_{i+1} = \begin{pmatrix} g_{i+1} \\ k_{i+1} \\ s_{i+1} \end{pmatrix} = \begin{pmatrix} \frac{5}{12} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{4} & \frac{7}{10} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{10} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} g_i \\ k_i \\ s_i \end{pmatrix} = Mv_i.$$

M is the transition matrix of the process.
Note: If the rows and columns of M are labelled G, K, S for Galway, Knock and Shannon, then the entry in the (i, j) position is the proportion of those planes that start the week in the airport

labelling Column j which finish the week in the airport labelling Row i. Note that M is a regular stochastic matrix.

Step 2: The steady state of the process is the unique eigenvector of m with eigenvalue 1 that is a probability vector. To calculate this we need to solve the system of equations whose coefficient matrix is $M - 1I_3$ (and which has zeroes on the right). The coefficient matrix is

$$M - I_3 = \begin{pmatrix} -\frac{7}{12} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{4} & -\frac{3}{10} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{10} & -\frac{2}{5} \end{pmatrix}$$

Remark: If A is a stochastic matrix (transition matrix), then the sum of the entries in each column of A is 1. It follows that the sum of the entries in each column of $A - I$ is 0, since $A - I$ is obtained from A by subtracting 1 from exactly one entry of each column. So the sum of the rows of $A - I$ is the row full of zeroes. This means that in reducing $A - I$ to reduced row echelon form, we can begin by simply eliminating one of the rows (by adding the sum of the remaining rows to it).

We proceed as follows with elementary row operations on the matrix $M - I$.

$$\begin{pmatrix} -\frac{7}{12} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{4} & -\frac{3}{10} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{10} & -\frac{2}{5} \end{pmatrix} \quad \begin{array}{c} R1 \leftrightarrow R3 \\ \longrightarrow \end{array} \quad \begin{pmatrix} \frac{1}{3} & \frac{1}{10} & -\frac{2}{5} \\ \frac{1}{4} & -\frac{3}{10} & \frac{1}{5} \\ -\frac{7}{12} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

$$\begin{array}{c} R3 \to R3 + (R1 + R2) \\ \longrightarrow \end{array} \quad \begin{pmatrix} \frac{1}{3} & \frac{1}{10} & -\frac{2}{5} \\ \frac{1}{4} & -\frac{3}{10} & \frac{1}{5} \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{array}{c} R1 \times 3 \\ \longrightarrow \\ R2 \times 4 \end{array} \quad \begin{pmatrix} 1 & \frac{3}{10} & -\frac{6}{5} \\ 1 & -\frac{12}{10} & \frac{4}{5} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{array}{c} R2 \to R2 - R1 \\ \longrightarrow \end{array} \quad \begin{pmatrix} 1 & \frac{3}{10} & -\frac{6}{5} \\ 0 & -\frac{15}{10} & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{array}{c} R2 \times (-2/3) \\ \longrightarrow \end{array} \quad \begin{pmatrix} 1 & \frac{3}{10} & -\frac{6}{5} \\ 0 & 1 & -\frac{4}{3} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{array}{c} R1 \to R1 - (3/10)R2 \\ \longrightarrow \end{array} \quad \begin{pmatrix} 1 & 0 & -\frac{4}{5} \\ 0 & 1 & -\frac{4}{3} \\ 0 & 0 & 0 \end{pmatrix}$$

Thus any vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ satisfying $x = \frac{4}{5}z$ and $y = \frac{4}{3}z$ is an eigenvector of M corresponding to the eigenvalue $\lambda = 1$. We need the unique such eigenvector in which the sum of the entries is 1, i.e.

$$\frac{4}{5}z + \frac{4}{3}z + z = 1 \Longrightarrow \frac{47}{15}z = 1.$$

Thus $z = \frac{15}{47}$, and the steady state vector is

$$\begin{pmatrix} \frac{12}{47} \\ \frac{20}{47} \\ \frac{15}{47} \end{pmatrix}.$$