Fundamentals of Linear Algebra

Marcel B. Finan
Arkansas Tech University

October 9, 2001

PREFACE

Linear algebra has evolved as a branch of mathematics with wide range of applications to the natural sciences, to engineering, to computer sciences, to management and social sciences, and more.

This book is addressed primarely to second and third your college students who have already had a course in calculus and analytic geometry. It is the result of lecture notes given by the author at The University of North Texas and the University of Texas at Austin. It has been designed for use either as a supplement of standard textbooks or as a textbook for a formal course in linear algebra.

This book is not a "traditional" book in the sense that it does not include any applications to the material discussed. Its aim is solely to learn the basic theory of linear algebra within a semester period. Instructors may wish to incorporate material from various fields of applications into a course.

I have included as many problems as possible of varying degrees of difficulty. Most of the exercises are computational, others are routine and seek to fix some ideas in the reader's mind; yet others are of theoretical nature and have the intention to enhance the reader's mathematical reasoning. After all doing mathematics is the way to learn mathematics.

Marcecl B. Finan
Austin, Texas
March, 2001.

Contents

Chapter 1

Linear Systems

In this chapter we shall develop the theory of general systems of linear equations. The tool we will use to find the solutions is the row-echelon form of a matrix. In fact, the solutions can be read off from the row- echelon form of the augmented matrix of the system. The solution technique, known as **elimination** method, is developed in Section 1.4.

1.1 Systems of Linear Equations

Many practical problems can be reduced to solving systems of linear equations. The main purpose of linear algebra is to find systematic methods for solving these systems. So it is natural to start our discussion of linear algebra by studying linear equations.

A **linear equation** in n variables is an equation of the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \tag{1.1}$$

where $x_1, x_2, ..., x_n$ are the **unknowns** (i.e. quantities to be found) and $a_1, \cdots, a_n$ are the **coefficients** (i.e. given numbers). Also given the number b known as the **constant term.** Observe that a linear equation does not involve any products, inverses, or roots of variables. All variables occur only to the first power and do not appear as arguments for trigonometric, logarithmic, or exponential functions.

Exercise 1
Determine whether the given equations are linear or not:
(a) $3x_1 - 4x_2 + 5x_3 = 6.$
(b) $4x_1 - 5x_2 = x_1x_2.$
(c) $x_2 = 2\sqrt{x_1} - 6.$
(d) $x_1 + \sin x_2 + x_3 = 1.$
(e) $x_1 - x_2 + x_3 = \sin 3.$

Solution
(a) The given equation is in the form given by (1.1) and therefore is linear.
(b) The equation is not linear because the term on the right side of the equation involves a product of the variables x_1 and x_2.
(c) A nonlinear equation because the term $2\sqrt{x_1}$ involves a square root of the variable x_1.
(d) Since x_2 is an argument of a trigonometric function then the given equation is not linear.
(e) The equation is linear according to (1.1) ■

A **solution** of a linear equation (1.1) in n unknowns is a finite ordered collection of numbers $s_1, s_2, ..., s_n$ which make (1.1) a true equality when $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$ are substituted in (1.1). The collection of all solutions of a linear equation is called the **solution set** or the **general solution**.

Exercise 2
Show that $(5 + 4s - 7t, s, t)$, *where* $s, t \in \mathbb{R}$, *is a solution to the equation*

$$x_1 - 4x_2 + 7x_3 = 5.$$

Solution
$x_1 = 5 + 4s - 7t, x_2 = s$, and $x_3 = t$ is a solution to the given equation because

$$x_1 - 4x_2 + 7x_3 = (5 + 4s - 7t) - 4s + 7t = 5. ■$$

A linear equation can have infinitely many solutions, exactly one solution or no solutions at all (See Theorem 5 in Section 1.7).

Exercise 3
Determine the number of solutions of each of the following equations:
(a) $0x_1 + 0x_2 = 5$.
(b) $2x_1 = 4$.
(c) $x_1 - 4x_2 + 7x_3 = 5$.

Solution.
(a) Since the left-hand side of the equation is 0 and the right-hand side is 5 then the given equation has no solution.
(b) By dividing both sides of the equation by 2 we find that the given equation has the unique solution $x_1 = 2$.
(c) To find the solution set of the given equation we assign arbitrary values s and t to x_2 and x_3 , respectively, and solve for x_1, we obtain

$$\left\{ \begin{array}{ccc} x_1 & = & 5 + 4s - 7t \\ x_2 & = & s \\ x_3 & = & t \end{array} \right.$$

Thus, the given equation has infinitely many solutions ■

s and t of the previous exercise are referred to as **parameters.** The solution in this case is said to be given in **parametric form.**

Many problems in the sciences lead to solving more than one linear equation. The general situation can be described by a linear system.
A **system of linear equations** or simply a **linear system** is any finite collection of linear equations. A **particular solution** of a linear system is any common solution of these equations. A system is called **consistent** if it has a solution. Otherwise, it is called **inconsistent.** A **general solution** of a system is a formula which gives all the solutions for different values of parameters (See Exercise 3 (c)).

A linear system of m equations in n variables has the form

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

As in the case of a single linear equation, a linear system can have infinitely many solutions, exactly one solution or no solutions at all. We will provide a proof of this statement in Section 1.7 (See Theorem 5). An alternative proof of the fact that when a system has more than one solution then it must have an infinite number of solutions will be given in Exercise 14.

Exercise 4
Find the general solution of the linear system

$$\left\{\begin{array}{ccccc} x_1 & + & x_2 & = & 7 \\ 2x_1 & + & 4x_2 & = & 18. \end{array}\right.$$

Solution.
Multiply the first equation of the system by -2 and then add the resulting equation to the second equation to find $2x_2 = 4$. Solving for x_2 we find $x_2 = 2$. Plugging this value in one of the equations of the given system and then solving for x_1 one finds $x_1 = 5$ ■

Exercise 5
By letting $x_3 = t$, find the general solution of the linear system

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 7 \\ 2x_1 & + & 4x_2 & + & x_3 & = & 18. \end{array}\right.$$

Solution.
By letting $x_3 = t$ the given system can be rewritten in the form

$$\left\{\begin{array}{ccccc} x_1 & + & x_2 & = & 7-t \\ 2x_1 & + & 4x_2 & = & 18-t. \end{array}\right.$$

By multiplying the first equation by -2 and adding to the second equation one finds $x_2 = \frac{4+t}{2}$. Substituting this expression in one of the individual equations of the system and then solving for x_1 one finds $x_1 = \frac{10-3t}{2}$ ■

1.2 Geometric Meaning of Linear Systems

In the previous section we stated that a linear system can have exactly one solution, infinitely many solutions or no solutions at all. In this section, we support our claim using geometry. More precisely, we consider the plane since a linear equation in the plane is represented by a straight line.

Consider the $x_1x_2 - plane$ and the set of points satisfying $ax_1 + bx_2 = c$. If $a = b = 0$ but $c \neq 0$ then the set of points satisfying the above equation is empty. If $a = b = c = 0$ then the set of points is the whole plane since the equation is satisfied for all $(x_1, x_2) \in \mathbb{R}^2$.

Exercise 6
Show that if $a \neq 0$ or $b \neq 0$ then the set of points satisfying $ax_1 + bx_2 = c$ is a straight line.

Solution.
If $a \neq 0$ but $b = 0$ then the equation $x_1 = \frac{c}{a}$ is a vertical line in the x_1x_2-plane. If $a = 0$ but $b \neq 0$ then $x_2 = \frac{c}{b}$ is a horizontal line in the plane. Finally, suppose that $a \neq 0$ and $b \neq 0$. Since x_2 can be assigned arbitrary values then the given equation possesses infinitely many solutions. Let $A(a_1, a_2), B(b_1, b_2)$, and $C(c_1, c_2)$ be any three points in the plane with components satisfying the given equation. The slope of the line AB is given by the expression $m_{AB} = \frac{b_2 - a_2}{b_1 - a_1}$ whereas that of AC is given by $m_{AC} = \frac{c_2 - a_2}{c_1 - a_1}$. From the equations $aa_1 + ba_2 = c$ and $ab_1 + bb_2 = c$ one finds $\frac{b_2 - a_2}{b_1 - a_1} = -\frac{a}{b}$. Similarly, $\frac{c_2 - a_2}{c_1 - a_1} = -\frac{a}{b}$. This shows that the lines AB and AC are parallel. Since these lines have the point A in common then A, B, and C are on the same straight line ■

The set of solutions of the system

$$\left\{ \begin{array}{lllll} ax_1 & + & bx_2 & = & c \\ a'x_1 & + & b'x_2 & = & c' \end{array} \right.$$

is the intersection of the set of solutions of the individual equations. Thus, if the system has exactly one solution then this solution is the point of intersection of two lines. If the system has infinitely many solutions then the two lines coincide. If the system has no solutions then the two lines are parallel.

Exercise 7
Find the point of intersection of the lines $x_1 - 5x_2 = 1$ and $2x_1 - 3x_2 = 3$.

Solution.
To find the point of intersection we have to solve the system

$$\left\{ \begin{array}{lllll} x_1 & - & 5x_2 & = & 1 \\ 2x_1 & - & 3x_2 & = & 3. \end{array} \right.$$

Using either elimination of unknowns or substitution one finds the solution $x_1 = \frac{12}{7}, x_2 = \frac{1}{7}$. ■

Exercise 8
Do the three lines $2x_1 + 3x_2 = -1$, $6x_1 + 5x_2 = 0$, *and* $2x_1 - 5x_2 = 7$ *have a common point of intersection?*

Solution.
Solving the system

$$\begin{cases} 2x_1 & + & 3x_2 & = & -1 \\ 6x_1 & + & 5x_2 & = & 0 \end{cases}$$

we find the solution $x_1 = \frac{5}{8}, x_2 = -\frac{3}{4}$. Since $2x_1 - 5x_2 = \frac{5}{4} + \frac{15}{4} = 5 \neq 7$ then the three lines do not have a point in common ■

A similar geometrical interpretation holds for systems of equations in three unknowns where in this case an equation is represented by a plane in $\mathbb{R}^3$. Since there is no physical image of the graphs for linear equations in more than three unknowns we will prove later by means of an algebraic argument(See Theorem 5 of Section 1.7) that our statement concerning the number of solutions of a linear system is still valid.

Exercise 9
Consider the system of equations

$$\begin{cases} a_1x_1 & + & b_1x_2 & = & c_1 \\ a_2x_1 & + & b_2x_2 & = & c_2 \\ a_3x_1 & + & b_3x_2 & = & c_3. \end{cases}$$

Discuss the relative positions of the above three lines when

(a) the system has no solutions,
(b) the system has exactly one solution,
(c) the system has infinitely many solutions.

Solution.
(a) The lines have no point of intersection.
(b) The lines intersect in exactly one point.
(c) The three lines coincide ■

Exercise 10
In the previous exercise, show that if $c_1 = c_2 = c_3 = 0$ *then the system has always a solution.*

Solution.
If $c_1 = c_2 = c_3$ then the system has at least one solution, namely the trivial solution $x_1 = x_2 = 0$ ■

1.3 Matrix Notation

Our next goal is to discuss some means for solving linear systems of equations. In Section 1.4 we will develop an algebraic method of solution to linear systems. But before we proceed any further with our discusion, we introduce a concept that simplifies the computations involved in the method.
The essential information of a linear system can be recorded compactly in a rectangular array called a matrix. A **matrix of size** $m \times n$ is a rectangular array of the form

$$\left(\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right)$$

where the a_{ij}'s are the **entries** of the matrix, m is the number of rows, and n is the number of columns. If $n = m$ the matrix is called **square** matrix.
We shall often use the notation $A = (a_{ij})$ for the matrix A, indicating that a_{ij} is the (i, j) entry in the matrix A.
An entry of the form a_{ii} is said to be on the **main diagonal**. An $m \times n$ matrix A with entries a_{ij} is called **upper triangular** (resp. **lower triangular**) if the entries below (resp. above) the main diagonal are all 0. That is, $a_{ij} = 0$ if $i > j$ (resp. $i < j$). A is called a **diagonal matrix** if $a_{ij} = 0$ whenever $i \neq j$. By a triangular matrix we mean either an upper triangular, a lower triangular, or a diagonal matrix.
Further definitions of matrices and related properties will be introduced in the next chapter.

Now, let A be a matrix of size $m \times n$ and entries a_{ij}; B is a matrix of size $n \times p$ and entries b_{ij}. Then the **product** matrix is a matrix of size $m \times p$ and entries

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$$

that is c_{ij} is obtained by multiplying componentwise the entries of the ith row of A by the entries of the jth column of B. It is very important to keep in mind that the number of columns of the first matrix must be equal to the number of rows of the second matrix; otherwise the product is undefined.

Exercise 11
Consider the matrices

$$A = \left(\begin{array}{ccc} 1 & 2 & 4 \\ 2 & 6 & 0 \end{array}\right), B = \left(\begin{array}{cccc} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{array}\right)$$

Compute, if possible, AB and BA.

Solution.
We have

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{pmatrix} \begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 4+8 & 1-2+28 & 4+6+20 & 3+2+8 \\ 8 & 2-6 & 8+18 & 6+6 \end{pmatrix} \\ &= \begin{pmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{pmatrix}. \end{aligned}$$

BA is not defined since the number of columns of B is not equal to the number of rows of A ■

Next, consider a system of linear equations

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

Then the matrix of the coefficients of the x_i's is called the **coefficient matrix:**

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

The matrix of the coefficients of the x_i's and the right hand side coefficients is called the **augmented** matrix:

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & b_2 \\ ... & ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} & b_m \end{pmatrix}$$

Finally, if we let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and

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

then the above system can be represented in **matrix notation** as

$$Ax = b.$$

Exercise 12
Consider the linear system

$$\left\{\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9. \end{array}\right.$$

(a) Find the coefficient and augmented matrices of the linear system.
(b) Find the matrix notation.

Solution.
(a) The coefficient matrix of this system is

$$\begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{pmatrix}$$

and the augmented matrix is

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{pmatrix}$$

(b) We can write the given system in matrix form as

$$\begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \\ -9 \end{pmatrix} \blacksquare$$

Exercise 13
Write the linear system whose augmented matrix is given by

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

Solution.
The given matrix is the augmented matrix of the linear system

$$\left\{\begin{array}{rcrcrcr} 2x_1 & - & x_2 & & & = & -1 \\ -3x_1 & + & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & + & x_3 & = & 3 \end{array}\right.$$

$\blacksquare$

Exercise 14
(a) Show that if A is an $m \times n$ matrix, x, y are $n \times 1$ matrices and α, β are numbers then $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$.
(b) Using the matrix notation of a linear system, prove that, if a linear system has more than one solution then it must have an infinite number of solutions.

Solution.
Recall that a number is considered an 1×1 matrix. Thus, using matrix multiplication we find

$$\begin{aligned}
A(\alpha x + \beta y) &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \vdots \\ \alpha x_n + \beta y_n \end{pmatrix} \\
&= \begin{pmatrix} \alpha(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) + \beta(a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n) \\ \vdots \\ \alpha(a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n) + \beta(a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n) \end{pmatrix} \\
&= \alpha \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} + \beta \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ \vdots \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n \end{pmatrix} \\
&= \alpha Ax + \beta Ay.
\end{aligned}$$

(b) Let X_1 and X_2 be two solutions of $Ax = b$. For any $t \in \mathbb{R}$ let $X_t = tX_1 + (1-t)X_2$. Then $AX_t = tAX_1 + (1-t)AX_2 = tb + (1-t)b = b$. That is, X_t is a solution to the linear system. Note that for $s \neq t$ we have $X_s \neq X_t$. Since t is arbitrary then there exists infinitely many X_t. In other words, the system has infinitely many solutions ■

1.4 Elementary Row Operations

In this section we introduce the concept of elementary row operations that will be vital for our algebraic method of solving linear systems.
We start with the following definition: Two linear systems are said to be **equivalent** if and only if they have the same set of solutions.

Exercise 15
Show that the system

$$\left\{ \begin{array}{rcrcr} x_1 & - & 3x_2 & = & -7 \\ 2x_1 & + & x_2 & = & 7 \end{array} \right.$$

is equivalent to the system

$$\left\{ \begin{array}{rcrcr} 8x_1 & - & 3x_2 & = & 7 \\ 3x_1 & - & 2x_2 & = & 0 \\ 10x_1 & - & 2x_2 & = & 14. \end{array} \right.$$

Solution.
Solving the first system one finds the solution $x_1 = 2, x_2 = 3$. Similarly, solving the second system one finds the solution $x_1 = 2$ and $x_2 = 3$. Hence, the two systems are equivalent ■

Exercise 16
Show that if $x_1 + kx_2 = c$ and $x_1 + lx_2 = d$ are equivalent then $k = l$ and $c = d$.

Solution.
For arbitrary t the ordered pair $(c - kt, t)$ is a solution to the second equation. That is $c - kt + lt = d$ for all $t \in \mathbb{R}$. In particular, if $t = 0$ we find $c = d$. Thus, $kt = lt$ for all $t \in \mathbb{R}$. Letting $t = 1$ we find $k = l$ ■

Our basic method for solving a linear system is known as the **method of elimination**. The method consists of reducing the original system to an equivalent system that is easier to solve. The reduced system has the shape of an upper (resp. lower) triangle. This new system can be solved by a technique called **backward-substitution** (resp. **forward-substitution**): The unknowns are found starting from the bottom (resp. the top) of the system.
The three basic operations in the above method, known as the **elementary row operations**, are summarized as follows.

(I) Multiply an equation by a non-zero number.
(II) Replace an equation by the sum of this equation and another equation multiplied by a number.
(III) Interchange two equations.

To indicate which operation is being used in the process one can use the following shorthand notation. For example, $r_3 \leftarrow \frac{1}{2}r_3$ represents the row operation of type (I) where each entry of row 3 is being replaced by $\frac{1}{2}$ that entry. Similar interpretations for types (II) and (III) operations.
The following theorem asserts that the system obtained from the original system by means of elementary row operations has the same set of solutions as the original one.

Theorem 1
Suppose that an elementary row operation is performed on a linear system. Then the resulting system is equivalent to the original system.

Proof.
We prove the theorem only for operations of type (II). The cases of operations of types (I) and (III) are left as an exercise for the reader (See Exercise 20 below).
Let

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = d \tag{1.2}$$

and

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1.3}$$

denote two different equations of the original system. Suppose a new system is obtained by replacing (1.3) by (1.4)

$$(a_1 + kc_1)x_1 + (a_2 + kc_2)x_2 + \cdots + (a_n + kc_n)x_n = b + kd \tag{1.4}$$

obtained by adding k times equation (1.2) to equation (1.3). If $s_1, s_2, \cdots, s_n$ is a solution to the original system then

$$c_1s_1 + c_2s_2 + \cdots + c_ns_n = d$$

and

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b.$$

By multiplication and addition, these give

$$(a_1 + kc_1)s_1 + (a_2 + kc_2)s_2 + \cdots + (a_n + kc_n)s_n = b + kd.$$

Hence, $s_1, s_2, \cdots, s_n$ is a solution to the new system. Conversely, suppose that $s_1, s_2, \cdots, s_n$ is a solution to the new system. Then

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b + kd - k(c_1s_1 + c_2s_2 + \cdots + c_ns_n) = b + kd - kd = b.$$

That is, $s_1, s_2, \cdots, s_n$ is a solution to the original system.$\blacksquare$

Exercise 17
Use the elimination method described above to solve the system

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & - & x_3 & = & 3 \\ x_1 & - & 3x_2 & + & 2x_3 & = & 1 \\ 2x_1 & - & 2x_2 & + & x_3 & = & 4. \end{array}\right.$$

Solution.
Step 1: We eliminate x_1 from the second and third equations by performing two operations $r_2 \leftarrow r_2 - r_1$ and $r_3 \leftarrow r_3 - 2r_1$ obtaining

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & - & x_3 & = & 3 \\ & - & 4x_2 & + & 3x_3 & = & -2 \\ & - & 4x_2 & + & 3x_3 & = & -2 \end{array}\right.$$

Step 2: The operation $r_3 \leftarrow r_3 - r_2$ leads to the system

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & - & x_3 & = & 3 \\ & - & 4x_2 & + & 3x_3 & = & -2 \end{array}\right.$$

By assigning x_3 an arbitrary value t we obtain the general solution $x_1 = \frac{t+10}{4}, x_2 = \frac{2+3t}{4}, x_3 = t$. This means that the linear system has infinitely many solutions. Every time we assign a value to t we obtain a different solution $\blacksquare$

Exercise 18
Determine if the following system is consistent or not

$$\left\{\begin{array}{rcrcrcr} 3x_1 & + & 4x_2 & + & x_3 & = & 1\\ 2x_1 & + & 3x_2 & & & = & 0\\ 4x_1 & + & 3x_2 & - & x_3 & = & -2. \end{array}\right.$$

Solution.
Step 1: To eliminate the variable x_1 from the second and third equations we perform the operations $r_2 \leftarrow 3r_2 - 2r_1$ and $r_3 \leftarrow 3r_3 - 4r_1$ obtaining the system

$$\left\{\begin{array}{rcrcrcr} 3x_1 & + & 4x_2 & + & x_3 & = & 1\\ & & x_2 & - & 2x_3 & = & -2\\ & - & 7x_2 & - & 7x_3 & = & -10. \end{array}\right.$$

Step 2: Now, to eliminate the variable x_3 from the third equation we apply the operation $r_3 \leftarrow r_3 + 7r_2$ to obtain

$$\left\{\begin{array}{rcrcrcr} 3x_1 & + & 4x_2 & + & x_3 & = & 1\\ & & x_2 & - & 2x_3 & = & -2\\ & & & - & 21x_3 & = & -24. \end{array}\right.$$

Solving the system by the method of backward substitution we find the unique solution $x_1 = -\frac{3}{7}, x_2 = \frac{2}{7}, x_3 = \frac{8}{7}$. Hence the system is consistent ■

Exercise 19
Determine whether the following system is consistent:

$$\left\{\begin{array}{rcrcr} x_1 & - & 3x_2 & = & 4\\ -3x_1 & + & 9x_2 & = & 8. \end{array}\right.$$

Solution.
Multiplying the first equation by 3 and adding the resulting equation to the second equation we find $0 = 20$ which is impossible. Hence, the given system is inconsistent ■

Exercise 20
(a) Show that the linear system obtained by interchanging two equations is equivalent to the original system.
(b) Show that the linear system obtained by multiplying a row by a scalar is equivalent to the original system.

Solution.
(a) Interchanging two equations in a linear system does yield the same system.
(b) Suppose now a new system is obtained by multiplying the ith row by $\alpha \neq 0$. Then the ith equation of this system looks like

$$(\alpha a_{i1})x_1 + (\alpha a_{i2})x_2 + \cdots + (\alpha a_{in})x_n = \alpha d_i. \tag{1.5}$$

If $s_1, s_2, \cdots, s_n$ is a solution to the original system then

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = d_i.$$

Multiply both sides of this equation by α yields (1.5). That is, $s_1, s_2, \cdots, s_n$ is a solution to the new system.
Now if $s_1, s_2, \cdots, s_n$ is a solution to (1.5) then by dividing through by α we find that $s_1, s_2, \cdots, s_n$ is a solution of the original system ■

1.5 Solving Linear Systems Using Augmented Matrices

In this section we apply the elimination method described in the previous section to the augmented matrix corresponding to a given system rather than to the individual equations. Thus, we obtain a triangular matrix which is row equivalent to the original augmented matrix, a concept that we define next.
We say that a matrix A is **row equivalent** to a matrix B if B can be obtained by applying a finite number of elementary row operations to the matrix A.
This definition combined with Theorem 1 lead to the following

Theorem 2 *Let $Ax = b$ be a linear system. If $[C|d]$ is row equivalent to $[A|b]$ then the system $Cx = d$ is equivalent to $Ax = b$.*

Proof.
The system $Cx = d$ is obtained from the system $Ax = b$ by applying a finite number of elementary row operations. By Theorem 1, this system is equivalent to $Ax = b$ ■

The above theorem provides us with a method for solving a linear system using matrices. It suffices to apply the elementary row operations on the augmented matrix and reduces it to an equivalent triangular matrix. Then the corresponding system is triangular as well. Next, use either the backward-substitution or the forward-substitution technique to find the unknowns.

Exercise 21
Solve the following linear system using elementary row operations on the augmented matrix:

$$\left\{\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9. \end{array}\right.$$

Solution.
The augmented matrix for the system is

$$\left(\begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right)$$

Step 1: The operations $r_2 \leftarrow \frac{1}{2}r_2$ and $r_3 \leftarrow r_3 + 4r_1$ give

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{pmatrix}$$

Step 2: The operation $r_3 \leftarrow r_3 + 3r_2$ gives

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

The corresponding system of equations is

$$\left\{ \begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & & & & x_3 & = & 3 \end{array} \right.$$

Using back-substitution we find the unique solution $x_1 = 29, x_2 = 16, x_3 = 3$ ■

Exercise 22
Solve the following linear system using the method described above.

$$\left\{ \begin{array}{rcrcrcr} & & x_2 & + & 5x_3 & = & -4 \\ x_1 & + & 4x_2 & + & 3x_3 & = & -2 \\ 2x_1 & + & 7x_2 & + & x_3 & = & -1. \end{array} \right.$$

Solution.
The augmented matrix for the system is

$$\begin{pmatrix} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -1 \end{pmatrix}$$

Step 1:The operation $r_2 \leftrightarrow r_1$ gives

$$\begin{pmatrix} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 2 & 7 & 1 & -1 \end{pmatrix}$$

Step 2: The operation $r_3 \leftarrow r_3 - 2r_1$ gives the system

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

Step 3: The operation $r_3 \leftarrow r_3 + r_2$ gives

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

The corresponding system of equations is

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 4x_2 & + & 3x_3 & = & -2\\ & & x_2 & + & 5x_3 & = & -4\\ & & & & 0 & = & -1 \end{array}\right.$$

From the last equation we conclude that the system is inconsistent ■

Exercise 23
Solve the following linear system using the elimination method of this section.

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 2x_2 & & & = & 0\\ -x_1 & + & 3x_2 & + & 3x_3 & = & -2\\ & & x_2 & + & x_3 & = & 0. \end{array}\right.$$

Solution.
The augmented matrix for the system is

$$\left(\begin{array}{rrrr} 1 & 2 & 0 & 0\\ -1 & 3 & 3 & -2\\ 0 & 1 & 1 & 0 \end{array}\right)$$

Step 1: Applying the operation $r_2 \leftarrow r_2 + r_1$ gives

$$\left(\begin{array}{rrrr} 1 & 2 & 0 & 0\\ 0 & 5 & 3 & -2\\ 0 & 1 & 1 & 0 \end{array}\right)$$

Step 2: The operation $r_2 \leftrightarrow r_3$ gives

$$\left(\begin{array}{rrrr} 1 & 2 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 5 & 3 & -2 \end{array}\right)$$

Step 3: Now performing the operation $r_3 \leftarrow r_3 - 5r_2$ yields

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

The system of equations equivalent to the original system is

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 2x_2 & & & = & 0\\ & & x_2 & + & x_3 & = & 0\\ & & & - & 2x_3 & = & -2 \end{array}\right.$$

Using back-substitution we find $x_1 = 2, x_2 = -1, x_3 = 1$ ■

Exercise 24
Determine if the following system is consistent.

$$\left\{\begin{array}{rcrcrcr} & & x_2 & - & 4x_3 & = & 8\\ 2x_1 & - & 3x_2 & + & 2x_3 & = & 1\\ 5x_1 & - & 8x_2 & + & 7x_3 & = & 1. \end{array}\right.$$

Solution.
The augmented matrix of the given system is

$$\left(\begin{array}{cccc} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{array}\right)$$

Step 1: The operation $r_3 \leftarrow r_3 - 2r_2$ gives

$$\left(\begin{array}{cccc} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 1 & -2 & 3 & -1 \end{array}\right)$$

Step 2: The operation $r_3 \leftrightarrow r_1$ leads to

$$\left(\begin{array}{cccc} 1 & -2 & 3 & -1 \\ 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \end{array}\right)$$

Step 3: Applying $r_2 \leftarrow r_2 - 2r_1$ to obtain

$$\left(\begin{array}{cccc} 1 & -2 & 3 & -1 \\ 0 & 1 & -4 & 3 \\ 0 & 1 & -4 & 8 \end{array}\right)$$

Step 4: Finally, the operation $r_3 \leftarrow r_3 - r_2$ gives

$$\left(\begin{array}{cccc} 1 & -2 & 3 & -1 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 0 & 5 \end{array}\right)$$

Hence, the equivalent system is

$$\left\{\begin{array}{ccccccc} x_1 & - & 2x_2 & + & 3x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 3 \\ & & & & 0 & = & 5 \end{array}\right.$$

This last system has no solution (the last equation requires x_1, x_2, and x_3 to satisfy the equation $0x_1 + 0x_2 + 0x_3 = 5$ and no such x_1, x_2, and x_3 exist). Hence the original system is inconsistent ■

Pay close attention to the last row of the row equivalent augmented matrix of the previous exercise. This situation is typical of an inconsistent system.

Exercise 25
Find an equation involving g, h, and k that makes the following augmented matrix corresponds to a consistent system.

$$\left(\begin{array}{cccc} 2 & 5 & -3 & g \\ 4 & 7 & -4 & h \\ -6 & -3 & 1 & k \end{array}\right)$$

Solution.
The augmented matrix for the given system is

$$\left(\begin{array}{rrrc} 2 & 5 & -3 & g \\ 4 & 7 & -4 & h \\ -6 & -3 & 1 & k \end{array}\right)$$

Step 1: Applying the operations $r_2 \leftarrow r_2 - 2r_1$ and $r_3 \leftarrow r_3 + 3r_1$ give

$$\left(\begin{array}{rrrc} 2 & 5 & -3 & g \\ 0 & -3 & 2 & h-2g \\ 0 & 12 & -8 & k+3g \end{array}\right)$$

Step 2: Now, the operation $r_3 \leftarrow r_3 + 4r_2$ gives

$$\left(\begin{array}{rrrc} 2 & 5 & -3 & g \\ 0 & -3 & 2 & h-2g \\ 0 & 0 & 0 & k+4h-5g \end{array}\right)$$

For the system , whose augmented matrix is the last matrix, to be consistent the unknowns x_1, x_2, and x_3 must satisfy the property $0x_1+0x_2+0x_3 = -5g+4h-k$, that is $-5g + 4h + k = 0$ ■

1.6 Echelon Form and Reduced Echelon Form

The elimination method introduced in the previous section reduces the augmented matrix to a "nice" matrix (meaning the corresponding equations are easy to solve). Two of the "nice" matrices discussed in this section are matrices in either row-echelon form or reduced row-echelon form, concepts that we discuss next.
By a **leading entry** of a row in a matrix we mean the leftmost non-zero entry in the row.
A rectangular matrix is said to be in **row-echelon form** if it has the following three characterizations:

(1) All rows consisting entirely of zeros are at the bottom.
(2) The leading entry in each non-zero row is 1 and is located in a column to the right of the leading entry of the row above it.
(3) All entries in a column below a leading entry are zero.

The matrix is said to be in **reduced row-echelon form** if in addition to the above, the matrix has the following additional characterization:

(4) Each leading 1 is the only nonzero entry in its column.

Remark
From the definition above, note that a matrix in row-echelon form has zeros

below each leading 1, whereas a matrix in reduced row-echelon form has zeros both above and below each leading 1.

Exercise 26
Determine which matrices are in row-echelon form (but not in reduced row-echelon form) and which are in reduced row-echelon form
(a)

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

(b)

$$\begin{pmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Solution.
(a)The given matrix is in row-echelon form but not in reduced row-echelon form since the $(1,2)-$entry is not zero.
(b) The given matrix satisfies the characterization of a reduced row-echelon form
■

The importance of the row-echelon matrices is indicated in the following theorem.

Theorem 3
Every nonzero matrix can be brought to (reduced) row-echelon form by a finite number of elementary row operations.

Proof.
The proof consists of the following steps:

Step 1. Find the first column from the left containing a nonzero entry (call it a), and move the row containing that entry to the top position.

Step 2. Multiply the row from Step 1 by $\frac{1}{a}$ to create a leading 1.

Step 3. By subtracting multiples of that row from rows below it, make each entry below the leading 1 zero.

Step 4. This completes the first row. Now repeat steps 1-3 on the matrix consisting of the remaining rows.

The process stops when either no rows remain in step 4 or the remaining rows consist of zeros. The entire matrix is now in row-echelon form.
To find the reduced row-echelon form we need the following additional step.

Step 5. Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1■

The process of reducing a matrix to a row-echelon form discussed in Steps 1 - 4 is known as **Gaussian elimination**. That of reducing a matrix to a reduced row-echelon form, i.e. Steps 1 - 5, is known as **Gauss-Jordan elimination**. We illustrate the above algorithm in the following problems.

Exercise 27
Use Gauss-Jordan elimination to transform the following matrix first into row-echelon form and then into reduced row-echelon form

$$\left(\begin{array}{ccccc} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{array}\right)$$

Solution.
The reduction of the given matrix to row-echelon form is as follows.

Step 1: $r_1 \leftrightarrow r_4$

$$\left(\begin{array}{ccccc} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{array}\right)$$

Step 2: $r_2 \leftarrow r_2 + r_1$ and $r_3 \leftarrow r_3 + 2r_1$

$$\left(\begin{array}{ccccc} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{array}\right)$$

Step 3: $r_2 \leftarrow \frac{1}{2}r_2$ and $r_3 \leftarrow \frac{1}{5}r_3$

$$\left(\begin{array}{ccccc} 1 & 4 & 5 & -9 & -7 \\ 0 & 1 & 2 & -3 & -3 \\ 0 & 1 & 2 & -3 & -3 \\ 0 & -3 & -6 & 4 & 9 \end{array}\right)$$

Step 4: $r_3 \leftarrow r_3 - r_2$ and $r_4 \leftarrow r_4 + 3r_2$

$$\left(\begin{array}{ccccc} 1 & 4 & 5 & -9 & -7 \\ 0 & 1 & 2 & -3 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{array}\right)$$

Step 5: $r_3 \leftrightarrow r_4$

$$\left(\begin{array}{ccccc} 1 & 4 & 5 & -9 & -7 \\ 0 & 1 & 2 & -3 & -3 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Step 6: $r_5 \leftarrow -\frac{1}{5}r_5$

$$\begin{pmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 1 & 2 & -3 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Step 7: $r_1 \leftarrow r_1 - 4r_2$

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

Step 8: $r_1 \leftarrow r_1 - 3r_3$ and $r_2 \leftarrow r_2 + 3r_3$

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

Exercise 28
Use Gauss-Jordan elimination to transform the following matrix first into row-echelon form and then into reduced row-echelon form

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{pmatrix}$$

Solution.
By following the steps in the Gauss-Jordan algorithm we find

Step 1: $r_3 \leftarrow \frac{1}{3}r_3$

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 1 & -3 & 4 & -3 & 2 & 5 \end{pmatrix}$$

Step 2: $r_1 \leftrightarrow r_3$

$$\begin{pmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix}$$

Step 3: $r_2 \leftarrow r_2 - 3r_1$

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

Step 4: $r_2 \leftarrow \frac{1}{2}r_2$

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

Step 5: $r_3 \leftarrow r_3 - 3r_2$

$$\left(\begin{array}{cccccc} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right)$$

Step 6: $r_1 \leftarrow r_1 + 3r_2$

$$\left(\begin{array}{cccccc} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right)$$

Step 7: $r_1 \leftarrow r_1 - 5r_3$ and $r_2 \leftarrow r_2 - r_3$

$$\left(\begin{array}{cccccc} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right) \blacksquare$$

It can be shown that no matter how the elementary row operations are varied, one will always arrive at the same reduced row-echelon form; that is the reduced row echelon form is unique (See Theorem 68). On the contrary row-echelon form is **not** unique. However, the number of leading 1's of two different row-echelon forms is the same (this will be proved in Chapter 4). That is, two row-echelon matrices have the same number of nonzero rows. This number is called the **rank** of A and is denoted by $rank(A)$. In Chapter 6, we will prove that if A is an $m \times n$ matrix then $rank(A) \leq n$ and $rank(A) \leq m$.

Exercise 29
Find the rank of each of the following matrices
(a)

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

(b)

$$B = \left(\begin{array}{ccccc} 3 & 1 & 0 & 1 & -9 \\ 0 & -2 & 12 & -8 & -6 \\ 2 & -3 & 22 & -14 & -17 \end{array}\right)$$

Solution.
(a) We use Gaussian elimination to reduce the given matrix into row-echelon form as follows:

Step 1: $r_2 \leftarrow r_2 - r_1$

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

Step 2: $r_1 \leftrightarrow r_2$

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

Step 3: $r_2 \leftarrow r_2 - 2r_1$

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

Step 4: $r_3 \leftarrow r_3 - r_2$

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

Thus, $rank(A) = 3$.
(b) As in (a), we reduce the matrix into row-echelon form as follows:

Step 1: $r_1 \leftarrow r_1 - r_3$

$$\begin{pmatrix} 1 & 4 & -22 & 15 & 8 \\ 0 & -2 & 12 & -8 & -6 \\ 2 & -3 & 22 & -14 & -17 \end{pmatrix}$$

Step 2: $r_3 \leftarrow r_3 - 2r_1$

$$\begin{pmatrix} 1 & 4 & -22 & 15 & 25 \\ 0 & -2 & 12 & -8 & -6 \\ 0 & -11 & -22 & -44 & -33 \end{pmatrix}$$

Step 3: $r_2 \leftarrow -\frac{1}{2}r_2$

$$\begin{pmatrix} 1 & 4 & -22 & 15 & 8 \\ 0 & 1 & -6 & 4 & 3 \\ 0 & -11 & -22 & -44 & -33 \end{pmatrix}$$

Step 4: $r_3 \leftarrow r_3 + 11r_2$

$$\begin{pmatrix} 1 & 4 & -22 & 15 & 8 \\ 0 & 1 & -6 & 4 & 3 \\ 0 & 0 & -88 & 0 & 0 \end{pmatrix}$$

Step 5: $r_3 \leftarrow \frac{1}{8}r_3$

$$\begin{pmatrix} 1 & 4 & -22 & 15 & 8 \\ 0 & 1 & -6 & 4 & 3 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Hence, $rank(B) = 3$ ■

Exercise 30
Consider the system

$$\left\{\begin{array}{ccccc} ax & + & by & = & k \\ cx & + & dy & = & l. \end{array}\right.$$

Show that if $ad - bc \neq 0$ *then the reduced row-echelon form of the coefficient matrix is the matrix*

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right)$$

Solution.
The coefficient matrix is the matrix

$$\left(\begin{array}{cc} a & b \\ c & d \end{array}\right)$$

Assume first that $a \neq 0$. Using Gaussian elimination we reduce the above matrix into row-echelon form as follows:

Step 1: $r_2 \leftarrow ar_2 - cr_1$

$$\left(\begin{array}{cc} a & b \\ 0 & ad - bc \end{array}\right)$$

Step 2: $r_2 \leftarrow \frac{1}{ad-bc} r_2$

$$\left(\begin{array}{cc} a & b \\ 0 & 1 \end{array}\right)$$

Step 3: $r_1 \leftarrow r_1 - br_2$

$$\left(\begin{array}{cc} a & 0 \\ 0 & 1 \end{array}\right)$$

Step 4: $r_1 \leftarrow \frac{1}{a} r_1$

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right)$$

Next, assume that $a = 0$. Then $c \neq 0$ and $b \neq 0$. Following the steps of Gauss-Jordan elimination algorithm we find

Step 1: $r_1 \leftrightarrow r_2$

$$\left(\begin{array}{cc} c & d \\ 0 & b \end{array}\right)$$

Step 2: $r_1 \leftarrow \frac{1}{c} r_1$ and $r_2 \leftarrow \frac{1}{b} r_2$

$$\left(\begin{array}{cc} 1 & \frac{d}{c} \\ 0 & 1 \end{array}\right)$$

Step 3: $r_1 \leftarrow r_1 - \frac{d}{c} r_2$

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right) \blacksquare$$

1.7 Echelon Forms and Solutions to Linear Systems

In this section we give a systematic procedure for solving systems of linear equations; it is based on the idea of reducing the augmented matrix to either the row-echelon form or the reduced row-echelon form. The new system is equivalent to the original system as provided by the following

Theorem 4
Let $Ax = b$ be a system of linear equations. Let $[C|d]$ be the (reduced) row-echelon form of the augmented matrix $[A|b]$. Then the system $Cx = d$ is equivalent to the system $Ax = b$.

Proof.
This follows from Theorem 2 and Theorem 3 ■

Unknowns corresponding to leading entries in the echelon augmented matrix are called **dependent** or **leading variables**. If an unknown is not dependent then it is called **free** or **independent** variable.

Exercise 31
Find the dependent and independent variables of the following system

$$\begin{cases} x_1 + 3x_2 - 2x_3 + 2x_5 = 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 = -1 \\ 5x_3 + 10x_4 + 15x_6 = 5 \\ 2x_1 + 6x_2 + 8x_4 + 4x_5 + 18x_6 = 6 \end{cases}$$

Solution.
The augmented matrix for the system is

$$\begin{pmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{pmatrix}$$

Using the Gaussian algorithm we bring the augmented matrix to row-echelon form as follows:

Step 1: $r_2 \leftarrow r_2 - 2r_1$ and $r_4 \leftarrow r_4 - 2r_1$

$$\begin{pmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{pmatrix}$$

Step 2: $r_2 \leftarrow -r_2$

$$\begin{pmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{pmatrix}$$

Step 3: $r_3 \leftarrow r_3 - 5r_2$ and $r_4 \leftarrow r_4 - 4r_2$

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

Step 4: $r_3 \leftrightarrow r_4$

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

Step 5: $r_3 \leftarrow \frac{1}{6}r_3$

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

The leading variables are $x_1, x_3,$ and x_6. The free variables are $x_2, x_4,$ and x_5 ■

It follows from Theorem 4 that one way to solve a linear system is to apply the elementary row operations to reduce the augmented matrix to a (reduced) row-echelon form. If the augmented matrix is in reduced row-echelon form then to obtain the general solution one just has to move all independent variables to the right side of the equations and consider them as parameters. The dependent variables are given in terms of these parameters.

Exercise 32
Solve the following linear system.

$$\begin{cases} x_1 + 2x_2 + x_4 = 6 \\ x_3 + 6x_4 = 7 \\ x_5 = 1. \end{cases}$$

Solution.
The augmented matrix is already in row-echelon form. The free variables are x_2 and x_4. So let $x_2 = s$ and $x_4 = t$. Solving the system starting from the bottom we find $x_1 = -2s - t + 6, x_3 = 7 - 6t$, and $x_5 = 1$ ■

If the augmented matrix does not have the reduced row-echelon form but the row-echelon form then the general solution also can be easily found by using the method of backward substitution.

Exercise 33
Solve the following linear system

$$\begin{cases} x_1 - 3x_2 + x_3 - x_4 = 2 \\ x_2 + 2x_3 - x_4 = 3 \\ x_3 + x_4 = 1. \end{cases}$$

Solution.
The augmented matrix is in row-echelon form. The free variable is $x_4 = t$. Solving for the leading variables we find, $x_1 = 11t + 4, x_2 = 3t + 1$, and $x_3 = 1 - t$ ■

The questions of existence and uniqueness of solutions are fundamental questions in linear algebra. The following theorem provides some relevant information.

Theorem 5
A system of n linear equations in m unknowns can have exactly one solution, infinitely many solutions, or no solutions at all.

(1) If the reduced augmented matrix has a row of the form $(0, 0, \cdots, 0, b)$ where b is a nonzero constant, then the system has no solutions.
(2) If the reduced augmented matrix has indepedent variables and no rows of the form $(0, 0, \cdots, 0, b)$ with $b \neq 0$ then the system has infinitely many solutions.
(3) If the reduced augmented matrix has no independent variables and no rows of the form $(0, 0, \cdots, 0, b)$ with $b \neq 0$, then the system has exactly one solution.

Proof
Suppose first that the reduced augmented matrix has a row of the form $(0, \cdots, 0, b)$ with $b \neq 0$. That is, $0x_1 + 0x_2 + ... + 0x_m = b$. Then the left side is 0 whereas the right side is not. This cannot happen. Hence, the system has no solutions. This proves (1).
Now suppose that the reduced augmented matrix has independent variables and no rows of the form $(0, 0, \cdots, 0, b)$ for some $b \neq 0$. Then these variables are treated as parameters and hence the system has infinitely many solutions. This proves (2).
Finally, suppose that the reduced augmented matrix has no row of the form $(0, 0, \cdots, 0, b)$ where $b \neq 0$ and no indepedent variables then the system looks like

$$\begin{array}{ccc} x_1 & = & c_1 \\ x_2 & = & c_2 \\ x_3 & = & c_3 \\ \dots & . & \dots \\ x_m & = & c_m \end{array}$$

i.e. the system has a unique solution. Thus, (3) is established.■

Exercise 34
Find the general solution of the system whose augmented matrix is given by

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

Solution.
We first reduce the system to row-echelon form as follows.

Step 1: $r_2 \leftarrow r_2 + r_1$ and $r_3 \leftarrow r_3 - 2r_1$

$$\begin{pmatrix} 1 & 2 & -7 \\ 0 & 1 & -6 \\ 0 & -3 & 19 \end{pmatrix}$$

Step 2: $r_3 \leftarrow r_3 + 3r_2$

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

The corresponding system is given by

$$\left\{ \begin{array}{ccccc} x_1 & + & 2x_2 & = & -7 \\ & & x_2 & = & -6 \\ & & 0 & = & 1 \end{array} \right.$$

Because of the last equation the system is inconsistent ■

Exercise 35
Find the general solution of the system whose augmented matrix is given by

$$\begin{pmatrix} 1 & -2 & 0 & 0 & 7 & -3 \\ 0 & 1 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 & 5 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Solution.
By adding two times the second row to the first row we find the reduced row-echelon form of the augmented matrix.

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

It follows that the free variables are $x_3 = s$ and $x_5 = t$. Solving for the leading variables we find $x_1 = -1 - t, x_2 = 1 + 3t$, and $x_4 = -4 - 5t$ ■

Exercise 36
Determine the value(s) of h such that the following matrix is the augmented matrix of a consistent linear system

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

Solution.
By adding three times the first row to the second row we find

$$\begin{pmatrix} 1 & 4 & 2 \\ 0 & 12 + h & 5 \end{pmatrix}$$

The system is consistent if and only if $12 + h \neq 0$; that is, $h \neq -12$■

From a computational point of view, most computer algorithms for solving systems use the Gaussian elimination rather than the Gauss-Jordan elimination. Moreover, the algorithm is somehow varied so that to guarantee a reduction in roundoff error, to minimize storage, and to maximize speed. For instance, many algorithms do not normalize the leading entry in each row to be 1.

Exercise 37
Find (if possible) conditions on the numbers a, b, and c such that the following system is consistent

$$\left\{\begin{array}{rcrcrcl} x_1 & + & 3x_2 & + & x_3 & = & a \\ -x_1 & - & 2x_2 & + & x_3 & = & b \\ 3x_1 & + & 7x_2 & - & x_3 & = & c \end{array}\right.$$

Solution.
The augmented matrix of the system is

$$\left(\begin{array}{cccc} 1 & 3 & 1 & a \\ -1 & -2 & 1 & b \\ 3 & 7 & -1 & c \end{array}\right)$$

Now apply Gaussian elimination as follows.

Step 1: $r_2 \leftarrow r_2 + r_1$ and $r_3 \leftarrow r_3 - 3r_1$

$$\left(\begin{array}{cccc} 1 & 3 & 1 & a \\ 0 & 1 & 2 & b+a \\ 0 & -2 & -4 & c-3a \end{array}\right)$$

Step 2: $r_3 \leftarrow r_3 + 2r_2$

$$\left(\begin{array}{cccc} 1 & 3 & 1 & a \\ 0 & 1 & 2 & b+a \\ 0 & 0 & 0 & c-a+2b \end{array}\right)$$

The system has no solution if $c - a + 2b \neq 0$. The system has infinitely many solutions if $c - a + 2b = 0$. In this case, the solution is given by $x_1 = 5t - (2a + 3b), x_2 = (a + b) - 2t, x_3 = t$ ■

1.8 Homogeneous Systems of Linear Equations

So far we have been discussing practical and systematic procedures to solve linear systems of equations. In this section, we will describe a theoretical method for solving systems. The idea consists of finding the general solution of the system with zeros on the right-hand side , call it $(x_1, x_2, \cdots, x_n)$, and then find a particular solution to the given system, say $(y_1, y_2, \cdots, y_n)$. The general solution of the original system is then the sum

$$(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n).$$

A **homogeneous** linear system is any system of the form

$$\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = 0. \end{array}$$

Every homogeneous system is consistent, since $x_1 = 0, x_2 = 0, \cdots, x_n = 0$ is always a solution. This solution is called the **trivial solution**; any other solution is called **nontrivial.** By Theorem 5, a homogeneous system has either a unique solution (the trivial solution) or infinitely many solutions. That is a homogeneous system is always consistent. The following theorem is a case where a homogeneous system is assured to have a nontrivial solution.

Theorem 6
Let $Ax = 0$ be a homogeneous system in m unknowns and n equations.
(1) If $rank(A) < m$ then the system has a nontrivial solution. Hence, by Exercise 14 the system has infinitely many solutions.
(2) If the number of unknowns exceeds the number of equations, i.e. $n < m$, then the system has a nontrivial solution.

Proof.
Applying the Gauss-Jordan elimination to the augmented matrix $[A|0]$ we obtain the matrix $[B|0]$. The number of nonzero rows of B is equals to $rank(A)$. Suppose first that $rank(A) = r < m$. In this case, the system $Bx = 0$ has r equations in m unknowns. Thus, the system has $m - r$ independent variables and consequently the system $Bx = 0$ has a nontrivial solution. By Theorem 4, the system $Ax = 0$ has a nontrivial solution. This proves (1). To prove (2), suppose $n < m$. If the system has only the trivial solution then by (1) we must have $rank(A) \geq m$. This implies that $n < m \leq rank(A) \leq n$, a contradiction. ■

Exercise 38
Solve the following homogeneous system using Gauss-Jordan elimination.

$$\left\{ \begin{array}{rcrcrcrcrcl} 2x_1 & + & 2x_2 & - & x_3 & & & + & x_5 & = & 0 \\ -x_1 & - & x_2 & + & 2x_3 & - & 3x_4 & + & x_5 & = & 0 \\ x_1 & + & x_2 & - & 2x_3 & & & - & x_5 & = & 0 \\ & & & & x_3 & + & x_4 & + & x_5 & = & 0. \end{array} \right.$$

Solution.
The reduction of the augmented matrix to reduced row-echelon form is outlined below.

$$\left(\begin{array}{cccccc} 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \\ 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{array} \right)$$

Step 1: $r_3 \leftarrow r_3 + r_2$

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

Step 2: $r_3 \leftrightarrow r_4$ and $r_1 \leftrightarrow r_2$

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

Step 3: $r_2 \leftarrow r_2 + 2r_1$ and $r_4 \leftarrow -\frac{1}{3}r_4$

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

Step 4: $r_1 \leftarrow -r_1$ and $r_2 \leftarrow \frac{1}{3}r_2$

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

Step 5: $r_3 \leftarrow r_3 - r_2$

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

Step 6: $r_4 \leftarrow r_4 - \frac{1}{3}r_3$ and $r_3 \leftarrow \frac{1}{3}r_3$

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

Step 7: $r_1 \leftarrow r_1 - 3r_3$ and $r_2 \leftarrow r_2 + 2r_3$

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

Step 8: $r_1 \leftarrow r_1 + 2r_2$

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

The corresponding system is

$$\begin{cases} x_1 + x_2 + x_5 = 0 \\ x_3 + x_5 = 0 \\ x_4 = 0 \end{cases}$$

The free variables are $x_2 = s, x_5 = t$ and the general solution is given by the formula: $x_1 = -s - t, x_2 = s, x_3 = -t, x_4 = 0, x_5 = t$ ■

Exercise 39
Solve the following homogeneous system using Gaussian elimination.

$$\begin{cases} x_1 + 3x_2 + 5x_3 + x_4 = 0 \\ 4x_1 - 7x_2 - 3x_3 - x_4 = 0 \\ 3x_1 + 2x_2 + 7x_3 + 8x_4 = 0 \end{cases}$$

Solution.
The augmented matrix for the system is

$$\begin{pmatrix} 1 & 3 & 5 & 1 & 0 \\ 4 & -7 & -3 & -1 & 0 \\ 3 & 2 & 7 & 8 & 0 \end{pmatrix}$$

We reduce this matrix into a row-echelon form as follows.
Step 1: $r_2 \leftarrow r_2 - r_3$

$$\begin{pmatrix} 1 & 3 & 5 & 1 & 0 \\ 1 & -9 & -10 & -9 & 0 \\ 3 & 2 & 7 & 8 & 0 \end{pmatrix}$$

Step 2: $r_2 \leftarrow r_2 - r_1$ and $r_3 \leftarrow r_3 - 3r_1$

$$\begin{pmatrix} 1 & 3 & 5 & 1 & 0 \\ 0 & -12 & -15 & -10 & 0 \\ 0 & -7 & -8 & 5 & 0 \end{pmatrix}$$

Step 3: $r_2 \leftarrow -\frac{1}{12}r_2$

$$\begin{pmatrix} 1 & 3 & 5 & 1 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{5}{6} & 0 \\ 0 & -7 & -8 & 5 & 0 \end{pmatrix}$$

Step 4: $r_3 \leftarrow r_3 + 7r_2$

$$\begin{pmatrix} 1 & 3 & 5 & 1 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{5}{6} & 0 \\ 0 & 0 & \frac{3}{4} & \frac{65}{6} & 0 \end{pmatrix}$$

Step 5: $r_3 \leftarrow \frac{4}{3}r_3$

$$\left(\begin{array}{ccccc} 1 & 3 & 5 & 1 & 0 \\ 0 & 1 & \frac{5}{4} & \frac{5}{6} & 0 \\ 0 & 0 & 1 & \frac{130}{9} & 0 \end{array}\right)$$

We see that $x_4 = t$ is the only free variable. Solving for the leading variables using back substitution we find $x_1 = \frac{176}{9}t, x_2 = \frac{155}{9}t$, and $x_3 = -\frac{130}{9}t$ ■

A **nonhomogeneous system** is a homogeneous system together with a nonzero right-hand side. Theorem 6 (2) applies only to homogeneous linear systems. A nonhomogeneous system with more unknowns than equations need not be consistent.

Exercise 40
Show that the following system is inconsistent.

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 0 \\ 2x_1 & + & 2x_2 & + & 2x_3 & = & 4. \end{array}\right.$$

Solution.
Multiplying the first equation by -2 and adding the resulting equation to the second we obtain $0 = 4$ which is impossible. So the system is inconsistent ■

The fundamental relationship between a nonhomogeneous system and its corresponding homogeneous system is given by the following theorem.

Theorem 7
Let $Ax = b$ be a linear system of equations. If y is a particular solution of the nonhomogeneous system $Ax = b$ and W is the solution set of the homogeneous system $Ax = 0$ then the general solution of $Ax = b$ consists of elements of the form $y + w$ where $w \in W$.

Proof.
Let S be the solution set of $Ax = b$. We will show that $S = \{y + w : w \in W\}$. That is, every element of S can be written in the form $y + w$ for some $w \in W$, and conversely, any element of the form $y + w$, with $w \in W$, is a solution of the nonhomogeneous system, i.e. a member of S. So, let z be an element of S, i.e. $Az = b$. Write $z = y + (z - y)$. Then $A(z - y) = Az - Ay = b + 0 = b$. Thus, $z = y + w$ with $w = z - y \in W$. Conversely, let $z = y + w$, with $w \in W$. Then $Az = A(y + w) = Ay + Aw = b + 0 = b$. That is $z \in S$. ■

We emphasize that the above theorem is of theoretical interest and does not help us to obtain explicit solutions of the system $Ax = b$. Solutions are obtained by means of the methods discussed in this chapter, i.e. Gauss elimination, Gauss-Jordan elimination or by the methods of using determinants to be discussed in Chapter 3.

Exercise 41
Show that if a homogeneous system of linear equations in n unknowns has a nontrivial solution then $rank(A) < n$, where A is the coefficient matrix.

Solution.
Since $rank(A) \leq n$ then either $rank(A) = n$ or $rank(A) < n$. If $rank(A) < n$ then we are done. So suppose that $rank(A) = n$. Then there is a matrix B that is row equivalent to A and that has n nonzero rows. Moreover, B has the following form

$$\begin{pmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} & 0 \\ 0 & 1 & a_{23} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

The corresponding system is triangular and can be solved by back substitution to obtain the solution $x_1 = x_2 = \cdots = x_n = 0$ which is a contradiction. Thus we must have $rank(A) < n$ ■

1.9 Review Problems

Exercise 42
Which of the following equations are not linear and why:
(a) $x_1^2 + 3x_2 - 2x_3 = 5.$
(b) $x_1 + x_1x_2 + 2x_3 = 1.$
(c) $x_1 + \frac{2}{x_2} + x_3 = 5.$

Exercise 43
Show that $(2s + 12t + 13, s, -s - 3t - 3, t)$ *is a solution to the system*

$$\left\{ \begin{array}{rcrcrcrcr} 2x_1 & + & 5x_2 & + & 9x_3 & + & 3x_4 & = & -1 \\ x_1 & + & 2x_2 & + & 4x_3 & & & = & 1 \end{array} \right.$$

Exercise 44
Solve each of the following systems using the method of elimination:
(a)

$$\left\{ \begin{array}{c} 4x_1 - 3x_2 = 0 \\ 2x_1 + 3x_2 = 18 \end{array} \right.$$

(b)

$$\left\{ \begin{array}{c} 4x_1 - 6x_2 = 10 \\ 6x_1 - 9x_2 = 15 \end{array} \right.$$

(c)

$$\left\{ \begin{array}{c} 2x_1 + x_2 = 3 \\ 2x_1 + x_2 = 1 \end{array} \right.$$

Which of the above systems is consistent and which is inconsistent?

Exercise 45
Find the general solution of the linear system

$$\begin{cases} x_1 - 2x_2 + 3x_3 + x_4 = -3 \\ 2x_1 - x_2 + 3x_3 - x_4 = 0 \end{cases}$$

Exercise 46
Find $a, b,$ and c so that the system

$$\begin{cases} x_1 + ax_2 + cx_3 = 0 \\ bx_1 + cx_2 - 3x_3 = 1 \\ ax_1 + 2x_2 + bx_3 = 5 \end{cases}$$

has the solution $x_1 = 3, x_2 = -1, x_3 = 2.$

Exercise 47
Find a relationship between a,b,c so that the following system is consistent

$$\begin{cases} x_1 + x_2 + 2x_3 = a \\ x_1 + x_3 = b \\ 2x_1 + x_2 + 3x_3 = c \end{cases}$$

Exercise 48
For which values of a will the following system have (a) no solutions? (b) exactly one solution? (c) infinitely many solutions?

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 4 \\ 3x_1 - x_2 + 5x_3 = 2 \\ 4x_1 + x_2 + (a^2-14)x_3 = a+2 \end{cases}$$

Exercise 49
Find the values of A,B,C in the following partial fraction

$$\frac{x^2 - x + 3}{(x^2+2)(2x-1)} = \frac{Ax+B}{x^2+2} + \frac{C}{2x-1}.$$

Exercise 50
Find a quadratic equation of the form $y = ax^2 + bx + c$ that goes through the points $(-2, 20), (1, 5),$ and $(3, 25)$.

Exercise 51
For which value(s) of the constant k does the following system have (a) no solutions? (b) exactly one solution? (c) infinitely many solutions?

$$\begin{cases} x_1 - x_2 = 3 \\ 2x_1 - 2x_2 = k \end{cases}$$

Exercise 52
Find a linear equation in the unknowns x_1 and x_2 that has a general solution $x_1 = 5 + 2t, x_2 = t.$

Exercise 53
Consider the linear system

$$\left\{\begin{array}{rcrcrcrcrcr} 2x_1 & + & 3x_2 & - & 4x_3 & + & x_4 & = & 5 \\ -2x_1 & & & + & x_3 & & & = & 7 \\ 3x_1 & + & 2x_2 & & & - & 4x_3 & = & 3 \end{array}\right.$$

(a) Find the coefficient and augmented matrices of the linear system.
(b) Find the matrix notation.

Exercise 54
Solve the following system using elementary row operations on the augmented matrix:

$$\left\{\begin{array}{rcrcrcr} 5x_1 & - & 5x_2 & - & 15x_3 & = & 40 \\ 4x_1 & - & 2x_2 & - & 6x_3 & = & 19 \\ 3x_1 & - & 6x_2 & - & 17x_3 & = & 41 \end{array}\right.$$

Exercise 55
Solve the following system.

$$\left\{\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & x_3 & = & -1 \\ x_1 & + & 2x_2 & + & x_3 & = & 0 \\ 3x_1 & & & - & 2x_3 & = & 5 \end{array}\right.$$

Exercise 56
Which of the following matrices are not in reduced row-ehelon from and why?
(a)

$$\left(\begin{array}{cccc} 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

(b)

$$\left(\begin{array}{cccc} 1 & 0 & 0 & 3 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 3 & 0 \end{array}\right)$$

(c)

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

Exercise 57
Use Gaussian elimination to convert the following matrix into a row-echelon matrix.

$$\left(\begin{array}{cccccc} 1 & -3 & 1 & -1 & 0 & -1 \\ -1 & 3 & 0 & 3 & 1 & 3 \\ 2 & -6 & 3 & 0 & -1 & 2 \\ -1 & 3 & 1 & 5 & 1 & 6 \end{array}\right)$$

Exercise 58
Use Gauss-Jordan elimination to convert the following matrix into reduced row-echelon form.

$$\begin{pmatrix} -2 & 1 & 1 & 15 \\ 6 & -1 & -2 & -36 \\ 1 & -1 & -1 & -11 \\ -5 & -5 & -5 & -14 \end{pmatrix}$$

Exercise 59
Solve the following system using Gauss-Jordan elimination.

$$\begin{cases} 3x_1 + x_2 + 7x_3 + 2x_4 = 13 \\ 2x_1 - 4x_2 + 14x_3 - x_4 = -10 \\ 5x_1 + 11x_2 - 7x_3 + 8x_4 = 59 \\ 2x_1 + 5x_2 - 4x_3 - 3x_4 = 39 \end{cases}$$

Exercise 60
Find the rank of each of the following matrices.
(a)

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

(b)

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

Exercise 61
Choose h and k such that the following system has (a) no solutions, (b) exactly one solution, and (c) infinitely many solutions.

$$\begin{cases} x_1 - 3x_2 = 1 \\ 2x_1 - hx_2 = k \end{cases}$$

Exercise 62
Solve the linear system whose augmented matrix is reduced to the following reduced row-echelon form

$$\begin{pmatrix} 1 & 0 & 0 & -7 & 8 \\ 0 & 1 & 0 & 3 & 2 \\ 0 & 0 & 1 & 1 & -5 \end{pmatrix}$$

Exercise 63
Solve the linear system whose augmented matrix is reduced to the following row-echelon form

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

Exercise 64
Solve the linear system whose augmented matrix is given by

$$\left(\begin{array}{cccc} 1 & 1 & 2 & 8 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & 10 \end{array}\right)$$

Exercise 65
Find the value(s) of a for which the following system has a nontrivial solution. Find the general solution.

$$\left\{\begin{array}{ccccccc} x_1 & + & 2x_2 & + & x_3 & = & 0 \\ x_1 & + & 3x_2 & + & 6x_3 & = & 0 \\ 2x_1 & + & 3x_2 & + & ax_3 & = & 0 \end{array}\right.$$

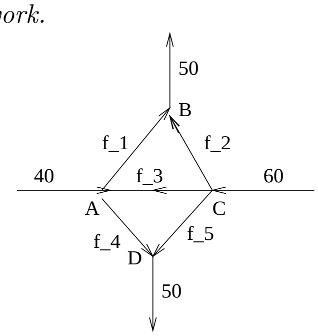
Exercise 66
Solve the following homogeneous system.

$$\left\{\begin{array}{ccccccccc} x_1 & - & x_2 & + & 2x_3 & + & x_4 & = & 0 \\ 2x_1 & + & 2x_2 & & & - & x_4 & = & 0 \\ 3x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 0 \end{array}\right.$$

Exercise 67
Let A be an $m \times n$ matrix.
(a) Prove that if y and z are solutions to the homogeneous system $Ax = 0$ then $y + z$ and cy are also solutions, where c is a number.
(b) Give a counterexample to show that the above is false for nonhomogeneous systems.

Exercise 68
Show that the converse of Theorem 6 is false. That is, show the existence of a nontrivial solution does not imply that the number of unknowns is greater than the number of equations.

Exercise 69 *(Network Flow)*
The junction rule of a network says that at each junction in the network the total flow into the junction must equal the total flows out. To illustrate the use of this rule, consider the network shown in the accompanying diagram. Find the possible flows in the network.



Chapter 2

Matrices

Matrices are essential in the study of linear algebra. The concept of matrices has become a tool in all branches of mathematics, the sciences, and engineering. They arise in many contexts other than as augmented matrices for systems of linear equations. In this chapter we shall consider this concept as objects in their own right and develop their properties for use in our later discussions.

2.1 Matrices and Matrix Operations

In this section, we discuss several types of matrices. We also examine four operations on matrices- addition, scalar multiplication, trace, and the transpose operation- and give their basic properties. Also, we introduce symmetric, skew-symmetric matrices.

A **matrix A of size** $m \times n$ is a rectangular array of the form

$$A = \left(\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right)$$

where the a_{ij}'s are the **entries** of the matrix, m is the number of rows, n is the number of columns. The **zero matrix 0** is the matrix whose entries are all 0. The $n \times n$ **identity matrix** I_n is a square matrix whose main diagonal consists of $1's$ and the off diagonal entries are all 0. A matrix A can be represented with the following compact notation $A = (a_{ij})$. The **ith row** of the matrix A is

$$[a_{i1}, a_{i2}, ..., a_{in}]$$

and the **jth column** is

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

In what follows we discuss the basic arithmetic of matrices.

Two matrices are said to be **equal** if they have the same size and their corresponding entries are all equal. If the matrix A is not equal to the matrix B we write $A \neq B$.

Exercise 70
Find x_1, x_2 and x_3 such that

$$\begin{pmatrix} x_1+x_2+2x_3 & 0 & 1 \\ 2 & 3 & 2x_1+4x_2-3x_3 \\ 4 & 3x_1+6x_2-5x_3 & 5 \end{pmatrix} = \begin{pmatrix} 9 & 0 & 1 \\ 2 & 3 & 1 \\ 4 & 0 & 5 \end{pmatrix}$$

Solution.
Because corresponding entries must be equal, this gives the following linear system

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & + & 2x_3 & = & 9 \\ 2x_1 & + & 4x_2 & - & 3x_3 & = & 1 \\ 3x_1 & + & 6x_2 & - & 5x_3 & = & 0 \end{array}\right.$$

The augmented matrix of the system is

$$\begin{pmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{pmatrix}$$

The reduction of this matrix to row-echelon form is

Step 1: $r_2 \leftarrow r_2 - 2r_1$ and $r_3 \leftarrow r_3 - 3r_1$

$$\begin{pmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{pmatrix}$$

Step 2: $r_2 \leftrightarrow r_3$

$$\begin{pmatrix} 1 & 1 & 2 & 9 \\ 0 & 3 & -11 & -27 \\ 0 & 2 & -7 & -17 \end{pmatrix}$$

Step 3: $r_2 \leftarrow r_2 - r_3$

$$\begin{pmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -4 & -10 \\ 0 & 2 & -7 & -17 \end{pmatrix}$$

Step 4: $r_3 \leftarrow r_3 - 2r_2$

$$\left(\begin{array}{cccc} 1 & 1 & 2 & 9 \\ 0 & 1 & -4 & -10 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

The corresponding system is

$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & + & 2x_3 & = & 9 \\ & & x_2 & - & 4x_3 & = & -10 \\ & & & & x_3 & = & 3 \end{array}\right.$$

Using backward substitution we find: $x_1 = 1, x_2 = 2, x_3 = 3$ ■

Exercise 71

Solve the following matrix equation for $a, b, c,$ and d

$$\left(\begin{array}{cc} a-b & b+c \\ 3d+c & 2a-4d \end{array}\right) = \left(\begin{array}{cc} 8 & 1 \\ 7 & 6 \end{array}\right)$$

Solution.

Equating corresponding entries we get the system

$$\left\{\begin{array}{ccccccccc} a & - & b & & & & & = & 8 \\ & & b & + & c & & & = & 1 \\ & & & & c & + & 3d & = & 7 \\ 2a & & & & & - & 4d & = & 6 \end{array}\right.$$

The augmented matrix is

$$\left(\begin{array}{ccccc} 1 & -1 & 0 & 0 & 8 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & 7 \\ 1 & 0 & 0 & -4 & 6 \end{array}\right)$$

We next apply Gaussian elimination as follows.

Step 1: $r_4 \leftarrow r_4 - r_1$

$$\left(\begin{array}{ccccc} 1 & -1 & 0 & 0 & 8 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & 7 \\ 0 & 1 & 0 & -4 & -2 \end{array}\right)$$

Step 2: $r_4 \leftarrow r_4 - r_2$

$$\left(\begin{array}{ccccc} 1 & -1 & 0 & 0 & 8 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & 7 \\ 0 & 0 & -1 & -4 & -3 \end{array}\right)$$

Step 3: $r_4 \leftarrow r_4 + r_3$

$$\left(\begin{array}{ccccc} 1 & -1 & 0 & 0 & 8 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & 7 \\ 0 & 0 & 0 & -1 & 4 \end{array}\right)$$

Using backward substitution to find: $a = -10, b = -18, c = 19, d = -4$ ■

Next, we introduce the operation of addition of two matrices. If A and B are two matrices of the same size, then the **sum** $A + B$ is the matrix obtained by adding together the corresponding entries in the two matrices. Matrices of different sizes cannot be added.

Exercise 72
Consider the matrices

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

Compute, if possible, $A + B$, $A + C$ *and* $B + C$.

Solution.
We have

$$A + B = \begin{pmatrix} 4 & 2 \\ 6 & 10 \end{pmatrix}$$

$A + B$ and $B + C$ are undefined since A and B are of different sizes as well as A and C ■

From now on, a constant number will be called a **scalar.** If A is a matrix and c is a scalar, then the product cA is the matrix obtained by multiplying each entry of A by c. Hence, $-A = (-1)A$. We define, $A - B = A + (-B)$. The matrix cI_n is called a **scalar** matrix.

Exercise 73
Consider the matrices

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

Compute $A - 3B$.

Solution.
Using the above definitions we have

$$A - 3B = \begin{pmatrix} 2 & -3 & -17 \\ -2 & 11 & -14 \end{pmatrix} ■$$

Let M_{mn} be the collection of all $m \times n$ matrices. This set under the operations of addition and scalar multiplication satisfies algebraic properties which will remind us of the system of real numbers. The proofs of these properties depend on the properties of real numbers. Here we shall assume that the reader is familiar with the basic algebraic properties of $\mathbb{R}$. The following theorem list the properties of matrix addition and multiplication of a matrix by a scalar.

Theorem 8
Let $A, B,$ and C be $m \times n$ and let c, d be scalars. Then

(i) $A + B = B + A$,
(ii) $(A + B) + C = A + (B + C) = A + B + C$,
(iii) $A + \mathbf{0} = \mathbf{0} + A = A$,
(iv) $A + (-A) = \mathbf{0}$,
(v) $c(A + B) = cA + cB$,
(vi) $(c + d)A = cA + dA$,
(vii) $(cd)A = c(dA)$,
(viii) $I_m A = AI_n = A$.

Proof.
(i) $A+B = (a_{ij})+(b_{ij}) = (a_{ij}+b_{ij}) = (b_{ij}+a_{ij}) = (b_{ij})+(a_{ij}) = B+A$, since addition of scalars is commutative.
(ii) Use the fact that addition of scalars is associative.
(iii) $A + \mathbf{0} = (a_{ij}) + (0) = (a_{ij} + 0) = (a_{ij}) = A$.
We leave the proofs of the remaining properties to the reader ■

Exercise 74
Solve the following matrix equation.

$$\left(\begin{array}{cc} 3 & 2 \\ -1 & 1 \end{array}\right) + \left(\begin{array}{cc} a & b \\ c & d \end{array}\right) = \left(\begin{array}{cc} 1 & 0 \\ -1 & 2 \end{array}\right)$$

Solution.
Adding and then equating corresponding entries we obtain $a = -2, b = -2, c = 0$, and $d = 1$ ■

If A is a square matrix then the sum of the entries on the main diagonal is called the **trace** of A and is denoted by $tr(A)$.

Exercise 75
Find the trace of the coefficient matrix of the system

$$\left\{\begin{array}{rcrcrcl} & - & x_2 & + & 3x_3 & = & 1 \\ x_1 & & & + & 2x_3 & = & 2 \\ -3x_1 & - & 2x_2 & & & = & 4 \end{array}\right.$$

Solution.
If A is the coefficient matrix of the system then

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

The trace of A is the number $tr(A) = 0 + 0 + 0 = 0$ ■

Two useful properties of the trace of a matrix are given in the following theorem.

Theorem 9
Let $A = (a_{ij})$ *and* $B = (b_{ij})$ *be two* $n \times n$ *matrices and* c *be a scalar. Then*
(i) $tr(A + B) = tr(A) + tr(B)$,
(ii) $tr(cA) = c\ tr(A)$.

Proof.
(i) $tr(A + B) = \sum_{i=1}^n (a_{ii} + b_{ii}) = \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = tr(A) + tr(B)$.
(ii) $tr(cA) = \sum_{i=1}^n ca_{ii} = c\sum_{i=1}^n a_{ii} = c\ tr(A)$.■

If A is an $m \times n$ matrix then the **transpose** of A, denoted by A^T, is defined to be the $n \times m$ matrix obtained by interchanging the rows and columns of A, that is the first column of A^T is the first row of A, the second column of A^T is the second row of A, etc. Note that, if $A = (a_{ij})$ then $A^T = (a_{ji})$. Also, if A is a square matrix then the diagonal entries on both A and A^T are the same.

Exercise 76
Find the transpose of the matrix

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

Solution.
The transpose of A is the matrix

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

The following result lists some of the properties of the transpose of a matrix.

Theorem 10
Let $A = (a_{ij})$, *and* $B = (b_{ij})$ *be two* $m \times n$ *matrices,* $C = (c_{ij})$ *be an* $n \times n$ *matrix, and* c *a scalar. Then*
(i) $(A^T)^T = A$,
(ii) $(A + B)^T = A^T + B^T$,
(iii) $(cA)^T = cA^T$,
(iv) $tr(C^T) = tr(C)$.

Proof.
(i) $(A^T)^T = (a_{ji})^T = (a_{ij}) = A$.
(ii) $(A + B)^T = (a_{ij} + b_{ij})^T = (a_{ji} + b_{ji}) = (a_{ji}) + (b_{ji}) = A^T + B^T$.
(iii) $(cA)^T = (ca_{ij})^T = (ca_{ji}) = c(a_{ji}) = cA^T$.
(iv) $tr(C^T) = \sum_{i=1}^n c_{ii} = tr(C)$. ■

Exercise 77
A square matrix A *is called* **symmetric** *if* $A^T = A$. *A square matrix* A *is called*

skew-symmetric *if* $A^T = -A$.
(a) Show that the matrix

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

is a symmetric matrix.
(b) Show that the matrix

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

is a skew-symmetric matrix.
(c) Show that for any square matrix A the matrix $S = \frac{1}{2}(A + A^T)$ *is symmetric and the matrix* $K = \frac{1}{2}(A - A^T)$ *is skew-symmetric.*
(d) Show that if A is a square matrix, then $A = S + K$, *where S is symmetric and K is skew-symmetric.*
(e) Show that the representation in (d) is unique.

Solution.
(a) A is symmetric since

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

(b) A is skew-symmetric since

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

(c) Because $S^T = \frac{1}{2}(A + A^T)^T = \frac{1}{2}(A + A^T)$ then S is symmetric. Similarly, $K^T = \frac{1}{2}(A - A^T)^T = \frac{1}{2}(A^T - A) = -\frac{1}{2}(A - A^T) = -K$ so that K is skew-symmetric.
(d) $S + K = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) = A$.
(e) Let S' be a symmetric matrix and K' be skew-symmetric such that $A = S' + K'$. Then $S + K = S' + K'$ and this implies that $S - S' = K - K'$. But the matrix $S - S'$ is symmetric and the matrix $K' - K$ is skew-symmetric. This equality is true only when $S - S'$ is the zero matrix. That is $S = S'$. Hence, $K = K'$ ■

Exercise 78
Let A be an $n \times n$ *matrix.*
(a) Show that if A is symmetric then A and A^T *have the same main diagonal.*
(b) Show that if A is skew-symmetric then the entries on the main diagonal are 0.
(c) If A and B are symmetric then so is $A + B$.

Solution.
(a) Let $A = (a_{ij})$ be symmetric. Let $A^T = (b_{ij})$. Then $b_{ij} = a_{ji}$ for all $1 \leq i, j \leq n$. In particular, when $i = j$ we have $b_{ii} = a_{ii}$. That is, A and A^T have the same main diagonal.
(b) Since A is skew-symmetric then $a_{ij} = -a_{ji}$. In particular, $a_{ii} = -a_{ii}$ and this implies that $a_{ii} = 0$.
(c) Suppose A and B are symmetric. Then $(A + B)^T = A^T + B^T = A + B$. That is, $A + B$ is symmetric ■

Exercise 79
Let A be an $m \times n$ matrix and α a real number. Show that if $\alpha A = \mathbf{0}$ then either $\alpha = 0$ or $A = \mathbf{0}$.

Solution.
Let $A = (a_{ij})$. Then $\alpha A = (\alpha a_{ij})$. Suppose $\alpha A = \mathbf{0}$. Then $\alpha a_{ij} = 0$ for all $0 \leq i \leq m$ and $0 \leq j \leq n$. If $\alpha \neq 0$ then $a_{ij} = 0$ for all indices i and j. In this case, $A = \mathbf{0}$ ■

2.2 Properties of Matrix Multiplication

In the previous section we discussed some basic properties associated with matrix addition and scalar multiplication. Here we introduce another important operation involving matrices-the product.
We have already introduced the concept of matrix multiplication in Section 1.3. For the sake of completeness we review this concept.
Let $A = (a_{ij})$ be a matrix of size $m \times n$ and $B = (b_{ij})$ be a matrix of size $n \times p$. Then the **product** matrix is a matrix of size $m \times p$ and entries

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{in}b_{nj},$$

that is, c_{ij} is obtained by multiplying componentwise the entries of the ith row of A by the entries of the jth column of B. It is very important to keep in mind that the number of columns of the first matrix must be equal to the number of rows of the second matrix; otherwise the product is undefined.
An interesting question associated with matrix multiplication is the following: If A and B are square matrices then is it always true that $AB = BA$?
The answer to this question is negative. In general, matrix multiplication is not commutative, as the following exercise shows.

Exercise 80

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

Show that $AB \neq BA$. Hence, matrix multiplication is not commutative.

Solution.
Using the definition of matrix multiplication we find

$$AB = \begin{pmatrix} -4 & 7 \\ 0 & 5 \end{pmatrix}, BA = \begin{pmatrix} -1 & 2 \\ 9 & 2 \end{pmatrix}$$

Hence, $AB \neq BA$ ■

Exercise 81
Consider the matrices

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix}, C = \begin{pmatrix} -1 & -2 \\ 11 & 4 \end{pmatrix}$$

(a) Compare $A(BC)$ and $(AB)C$.
(b) Compare $A(B+C)$ and $AB + AC$.
(c) Compute I_2A and AI_2, where I_2 is the 2×2 identity matrix.

Solution.
(a)

$$A(BC) = (AB)C = \begin{pmatrix} 70 & 14 \\ 235 & 56 \end{pmatrix}$$

(b)

$$A(B+C) = AB + AC = \begin{pmatrix} 16 & 7 \\ 59 & 33 \end{pmatrix}$$

(c) $AI_2 = I_2A = A$ ■

Exercise 82
Let A be an $m \times n$ matrix and B an $n \times p$ matrix.

(a) Show that the entries in the jth column of AB are the entries in the product AB_j where B_j is the jth column of B. Thus, $AB = [AB_1, AB_2, \cdots, AB_p]$.
(b) Show that the entries in the ith row of AB are the entries of the product A_iB where A_i is the ith row of A. That is,

$$AB = \begin{pmatrix} A_1B \\ A_2B \\ \vdots \\ A_mB \end{pmatrix}$$

Solution.
Let C_j be the jth column of AB then

$$C_j = \begin{pmatrix} a_{11}b_{1j} + a_{12}b_{2j} + \cdots + a_{1n}b_{nj} \\ a_{21}b_{1j} + a_{22}b_{2j} + \cdots + a_{2n}b_{nj} \\ \vdots \\ a_{n1}b_{1j} + a_{n2}b_{2j} + \cdots + a_{mn}b_{nj} \end{pmatrix}$$

On the other hand,

$$\begin{aligned} AB_j &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{1j} + a_{12}b_{2j} + \cdots + a_{1n}b_{nj} \\ a_{21}b_{1j} + a_{22}b_{2j} + \cdots + a_{2n}b_{nj} \\ \vdots \\ a_{m1}b_{1j} + a_{m2}b_{2j} + \cdots + a_{mn}b_{nj} \end{pmatrix} = C_j \end{aligned}$$

(b) Let $r_1, r_2, \cdots, r_m$ be the rows of AB. Then

$$r_i^T = \left\{ \begin{matrix} a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1} \\ a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2} \\ \vdots \\ a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np} \end{matrix} \right)$$

If A_i is the ith row of A then

$$A_iB = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix} = r_i \blacksquare$$

Exercise 83
(a) If B has a column of zeros, show that the same is true of AB for any A.
(b) If A has a row of zeros, show that the same is true of AB for any B.

Solution.
(a) Follows from (a) of the previous exercise.
(b) Follows from (b) of the previous exercise $\blacksquare$

Exercise 84
Let A be an $m \times n$ matrix. Show that if $Ax = \mathbf{0}$ for all $n \times 1$ matrix x then $A = \mathbf{0}$.

Solution.
Suppose the contrary. Then there exist indices i and j such that $a_{ij} \neq 0$. Let x be the $n \times 1$ matrix whose jth row is 1 and 0 elsewhere. Then the jth entry of Ax is $0 = a_{ij} \neq 0$, a contradiction $\blacksquare$

As the reader has noticed so far, most of the basic rules of arithmetic of real numbers also hold for matrices but a few do not. In Exercise 80 we have seen that matrix multiplication is not commutative. The following exercise shows that the cancellation law of numbers does not hold for matrix product.

Exercise 85
(a) Consider the matrices

$$A = \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right), B = \left(\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right), C = \left(\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right)$$

Compare AB and AC. Is it true that $B = C$?
(b) Find two square matrices A and B such that $AB = \mathbf{0}$ but $A \neq \mathbf{0}$ and $B \neq \mathbf{0}$.

Solution.
(a) Note that $B \neq C$ even though $AB = AC = \mathbf{0}$.
(b) The given matrices satisfy $AB = \mathbf{0}$ with $A \neq \mathbf{0}$ and $B \neq \mathbf{0}$ ■

Matrix multiplication shares many properties of the product of real numbers which are listed in the following theorem

Theorem 11
Let A be a matrix of size $m \times n$. Then
(a) $A(BC) = (AB)C$, where B is of size $n \times p$, C of size $p \times q$.
(b) $A(B + C) = AB + AC$, where B and C are of size $n \times p$.
(c) $(B + C)A = BA + CA$, where B and C are of size $l \times m$.
(d) $c(AB) = (cA)B = A(cB)$, where c denotes a scalar.

Proof.
Let $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$.
(a) Let $AB = (d_{ij}), BC = (e_{ij}), A(BC) = (f_{ij})$, and $(AB)C = (g_{ij})$. Then from the definition of matrix multiplication we have

$$\begin{array}{rcl} f_{ij} & = & (ith\ row\ of\ A)(jth\ column\ of\ BC) \\ & = & a_{i1}e_{1j} + a_{i2}e_{2j} + \cdots + a_{in}e_{nj} \\ & = & a_{i1}(b_{11}c_{1j} + b_{12}c_{2j} + \cdots + b_{1p}c_{pj}) \\ & + & a_{i2}(b_{21}c_{1j} + b_{22}c_{2j} + \cdots + b_{2p}c_{pj}) \\ & + & \cdots + a_{in}(b_{n1}c_{1j} + b_{n2}c_{2j} + \cdots + b_{np}c_{pj}) \\ & = & (a_{i1}b_{11} + a_{i2}b_{21} + \cdots + a_{in}b_{n1})c_{1j} \\ & + & (a_{i1}b_{12} + a_{i2}b_{22} + \cdots + a_{in}b_{n2})c_{2j} \\ & + & \cdots + (a_{i1}b_{1p} + a_{i2}b_{2p} + \cdots + a_{in}b_{np})c_{pj} \\ & = & d_{i1}c_{1j} + d_{i2}c_{2j} + \cdots + d_{ip}c_{pj} \\ & = & (ith\ row\ of\ AB)(jth\ column\ of\ C) \\ & = & g_{ij}. \end{array}$$

(b) Let $A(B + C) = (k_{ij}), AC = (h_{ij})$ Then

$$\begin{array}{rcl} k_{ij} & = & (ith\ row\ of\ A)(jth\ column\ of\ B + C) \\ & = & a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \cdots + a_{in}(b_{nj} + c_{nj}) \\ & = & (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}) \\ & + & (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{in}c_{nj}) \\ & = & d_{ij} + h_{ij}. \end{array}$$

(c) Similar to (b).
(d) The proof is left for the reader. ■

In the next theorem we establish a property about the transpose of a matrix.

Theorem 12
Let $A = (a_{ij}), B = (b_{ij})$ be matrices of sizes $m \times n$ and $n \times m$ respectively. Then $(AB)^T = B^T A^T$.

proof
Let $AB = (c_{ij})$. Then $(AB)^T = (c_{ji})$. Let $B^T A^T = (c'_{ij})$.Then

$$\begin{array}{rcl} c'_{ij} & = & (ith\ row\ of\ B^T)(jth\ column\ of\ A^T) \\ & = & a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni} \\ & = & c_{ji} \end{array}$$

This ends a proof of the theorem ■

Exercise 86
Let A be any matrix. Show that AA^T and $A^T A$ are symmetric matrices.

Solution.
First note that for any matrix A the matrices AA^T and $A^T A$ are well-defined. Since $(AA^T)^T = (A^T)^T A^T = AA^T$ then AA^T is symmetric. Similarly, $(A^T A)^T = A^T (A^T)^T = A^T A$ ■

Finally, we discuss the powers of a square matrix. Let A be a square matrix of size $n \times n$. Then the non-negative powers of A are defined as follows: $A^0 = I_n, A^1 = A$, and for $k \geq 2$, $A^k = (A^{k-1})A$.

Exercise 87
suppose that

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

Compute A^3.

Solution.
Multiplying the matrix A by itself three times we obtain

$$A^3 = \begin{pmatrix} 37 & 54 \\ 81 & 118 \end{pmatrix} \blacksquare$$

For the sake of completeness, we mention few words regarding the mathematical proof by induction which will be used in the next theorem as well as in the rest of these notes. Let S be a statement that depends on a non-negative integer $n \geq n_0$. To prove that S is valid for all $n \geq n_0$ one has to show the following:

1. Show that S is valid for n_0.
2. Supposing that S is valid for $n > n_0$ show that S is valid for $n + 1$.

Theorem 13
For any non-negative integers s, t we have
(a) $A^{s+t} = A^s A^t$
(b) $(A^s)^t = A^{st}$.

Proof.
(a) Fix s. Let $t = 1$. Then by definition above we have $A^{s+1} = A^s A$. Now, we prove by induction on t that $A^{s+t} = A^s A^t$. The equality holds for $t = 1$. As the induction hypothesis, suppose that $A^{s+t} = A^s A^t$. Then $A^{s+(t+1)} = A^{(s+t)+1} = A^{s+t}A = (A^s A^t)A = A^s(A^t A) = A^s A^{t+1}$.
(b) The proof is similar to (a) and is left to the reader. ■

Exercise 88
Let A and B be two $n \times n$ matrices.
(a) Show that $tr(AB) = tr(BA)$.
(b) Show that $AB - BA = I_n$ is impossible.

Solution.
(a) Let $A = (a_{ij})$ and $B = (b_{ij})$. Then
$tr(AB) = \sum_{i=1}^n (\sum_{k=1}^n a_{ik}b_{ki}) = \sum_{i=1}^n (\sum_{k=1}^n b_{ik}a_{ki}) = tr(BA)$.
(b) If $AB - BA = I_n$ then $0 = tr(AB) - tr(BA) = tr(AB - BA) = tr(I_n) = n \geq 1$, a contradiction ■

2.3 The Inverse of a Square Matrix

Most problems in practice reduces to a system with matrix notation $Ax = b$. Thus, in order to get x we must somehow be able to eliminate the coefficient matrix A. One is tempted to try to divide by A. Unfortunately such an operation has not been defined for matrices. In this section we introduce a special type of square matrices and formulate the matrix analogue of numerical division. Recall that the $n \times n$ identity square matrix is the matrix I_n whose main diagonal entries are 1 and off diagonal entries are 0.
A square matrix A of size n is called **invertible** or **non-singular** if there exists a square matrix B of the same size such that $AB = BA = I_n$. In this case B is called the **inverse** of A. A square matrix that is not invertible is called **singular.**

Exercise 89
Show that the matrix

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

is the inverse of the matrix

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

Solution.
Using matrix multiplication one checks that $AB = BA = I_2$ ■

Exercise 90
Show that the matrix

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

is singular.

Solution.
Let $B = (b_{ij})$ be a 2×2 matrix. If $BA = I_2$ then the $(2,2)$-th entry of BA is zero while the $(2,2)-$entry of I_2 is 1, which is impossible. Thus, A is singular ■

In Section 2.5 (See Theorem 20), we will prove that for a square matrix to be invertible it suffices to prove the existence of a matrix B such that either $AB = I_n$ or $BA = I_n$. In other words, there is no need to check the double equality $AB = BA = I_n$.
However, it is important to keep in mind that the concept of invertibility is defined only for square matrices. In other words, it is possible to have a matrix A of size $m \times n$ and a matrix B of size $n \times m$ such that $AB = I_m$. It would be wrong to conclude that A is invertible and B is its inverse.

Exercise 91
Let

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

Show that $AB = I_2$.

Solution.
Simple matrix multiplication shows that $AB = I_2$. However, this does not imply that B is the inverse of A since BA is undefnied so that the condition $BA = I_2$ fails ■

Exercise 92
Show that the identity matrix is invertible but the zero matrix is not.

Solution.
Since $I_n I_n = I_n$ then I_n is nonsingular and its inverse is I_n. Now, for any $n \times n$ matrix B we have $B\mathbf{0} = \mathbf{0} \neq I_n$ so that the zero matrix is not invertible ■

Now if A is a nonsingular matrix then how many different inverses does it possess? The answer to this question is provided by the following theorem.

Theorem 14
The inverse of a matrix is unique.

Proof.
Suppose A has two inverses B and C. We will show that $B = C$. Indeed, $B = BI_n = B(AC) = (BA)C = I_nC = C$. ■

Since an invertible matrix A has a unique inverse then we will denote it from now on by A^{-1}.
For an invertible matrix A one can now define the negative power of a square matrix as follows: For any positive integer $n \geq 1$, we define $A^{-n} = (A^{-1})^n$.
The next theorem lists some of the useful facts about inverse matrices.

Theorem 15
Let A and B be two square matrices of the same size $n \times n$.
(a) If A and B are invertible matrices then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.
(b) If A is invertible then A^{-1} is invertible and $(A^{-1})^{-1} = A$.
(c) If A is invertible then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$

Proof.
(a) If A and B are invertible then $AA^{-1} = A^{-1}A = I_n$ and $BB^{-1} = B^{-1}B = I_n$. In This case, $(AB)(B^{-1}A^{-1}) = A[B(B^{-1}A^{-1})] = A[(BB^{-1})A^{-1}] = A(I_nA^{-1}) = AA^{-1} = I_n$. Similarly, $(B^{-1}A^{-1})(AB) = I_n$. It follows that $B^{-1}A^{-1}$ is the inverse of AB.
(b) Since $A^{-1}A = AA^{-1} = I_n$ then A is the inverse of A^{-1}, i.e. $(A^{-1})^{-1} = A$.
(c) Since $AA^{-1} = A^{-1}A = I_n$ then by taking the transpose of both sides we get $(A^{-1})^TA^T = A^T(A^{-1})^T = I_n$. This shows that A^T is invertible with inverse $(A^{-1})^T$. ■

Exercise 93
(a) Under what conditions a diagonal matrix is invertible?
(b) Is the sum of two invertible matrices necessarily invertible?

Solution.
(a) Let $D = (d_{ii})$ be a diagonal $n \times n$ matrix. Let $B = (b_{ij})$ be an $n \times n$ matrix such that $DB = I_n$ and let $DB = (c_{ij})$. Then using matrix multiplication we find $c_{ij} = \sum_{k=1}^n d_{ik}b_{kj}$. If $i \neq j$ then $c_{ij} = d_{ii}b_{ij} = 0$ and $c_{ii} = d_{ii}b_{ii} = 1$. If $d_{ii} \neq 0$ for all $1 \leq i \leq n$ then $b_{ij} = 0$ for $i \neq j$ and $b_{ii} = \frac{1}{d_{ii}}$. Thus, if $d_{11}d_{22}\cdots d_{nn} \neq 0$ then D is invertible and its inverse is the diagonal matrix $D^{-1} = (\frac{1}{d_{ii}})$.
(b) The following two matrices are invertible but their sum , which is the zero matrix, is not.

$$A = \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right), \left(\begin{array}{cc} -1 & 0 \\ 0 & 1 \end{array}\right) \blacksquare$$

Exercise 94
Consider the 2×2 matrix

$$A = \left(\begin{array}{cc} a & b \\ c & d \end{array}\right).$$

Show that if $ad - bc \neq 0$ then A^{-1} exists and is given by

$$A^{-1} = \frac{1}{ad - bc}\left(\begin{array}{cc} d & -b \\ -c & a \end{array}\right).$$

Solution.
Let

$$B = \left(\begin{array}{cc} x & y \\ z & w \end{array}\right)$$

be a matrix such that $BA = I_2$. Then using matrix multiplication we find

$$\left(\begin{array}{cc} ax + cy & bx + dy \\ az + cw & bz + dw \end{array}\right) = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right)$$

Equating corresponding entries we obtain the following systems of linear equations in the unknowns x, y, z and w.

$$\left\{\begin{array}{ccccc} ax & + & cy & = & 1 \\ bx & + & dy & = & 0 \end{array}\right.$$

and

$$\left\{\begin{array}{ccccc} az & + & cw & = & 0 \\ bz & + & dw & = & 0 \end{array}\right.$$

In the first system, using elimination we find $(ad-bc)y = -b$ and $(ad-bc)x = d$. Similarly, using the second system we find $(ad - bc)z = -c$ and $(ad - bc)w = a$. If $ad - bc \neq 0$ then one can solve for $x, y, z,$ and w and in this case $B = A^{-1}$ as given in the statement of the problem ■

Finally, we mention here that matrix inverses can be used to solve systems of linear equations as suggested by the following theorem.

Theorem 16
If A is an $n \times n$ invertible matrix and b is a column matrix then the equation $Ax = b$ has a unique solution $x = A^{-1}b$.

Proof.
Since $A(A^{-1}b) = (AA^{-1})b = I_n b = b$ then $A^{-1}b$ is a solution to the equation $Ax = b$. Now, if y is another solution then $y = I_n y = (A^{-1}A)y = A^{-1}(Ay) = A^{-1}b$. ■

Exercise 95
Let A and B be $n \times n$ matrices.
(a) Verify that $A(I_n + BA) = (I_n + AB)A$ and that $(I_n + BA)B = B(I_n + AB)$.
(b) If $I_n + AB$ is invertible show that $I_n + BA$ is also invertible and $(I_n + BA)^{-1} = I_n - B(I_n + AB)^{-1}A$.

Solution.
(a) $A(I_n + BA) = AI_n + ABA = I_nA + ABA = (I_n + AB)$. Similar argument for the second equality.
(b) Suppose $I_n + AB$ is invertible. Then postmultiplying the second equality in (a) by $-(I_n + AB)^{-1}A$ to obtain $-(I_n + BA)B(I_n + AB)^{-1}A = -BA$. Now add $I_n + BA$ to both sides to obtain

$$(I_n + BA)[I_n - B(I_n + AB)^{-1}A] = I_n$$

This says that $I_n + BA$ is invertible and $(I + n + BA)^{-1} = I_n - B(I_n + AB)^{-1}A$ ■

Exercise 96
If A is invertible and $k \neq 0$ show that $(kA)^{-1} = \frac{1}{k}A^{-1}$.

Solution.
Suppose that A is invertible and $k \neq 0$. Then $(kA)A^{-1} = k(AA^{-1}) = kI_n$. This implies $(kA)(\frac{1}{k}A^{-1}) = I_n$. Thus, kA is invertible with inverse equals to $\frac{1}{k}A^{-1}$ ■

Exercise 97
Prove that if R is an $n \times n$ matrix in reduced row-echelon form and with no zero rows then $R = I_n$.

Solution.
Since R has no zero rows then R^T has no zero rows as well. Since the first row of R is nonzero then the leading 1 must be in the $(1, 1)$ position; otherwise, the first column will be zero and this shows that R^T has a zero row which is a contradiction. By the definition of reduced row-echelon form all entries below the leading entry in the first column must be 0 and all entries to the right of the leading entry in the first row are 0. Now, we repeat this argument to the $(n - 1) \times (n - 1)$ matrix obtained by deleting the first row and the first column of R and so on ■

2.4 Elementary Matrices

In this section we introduce a special type of invertible matrices, the so-called elementary matrices, and we discuss some of their properties. As we shall see, elementary matrices will be used in the next section to develop an algorithm for finding the inverse of a square matrix.
An $n \times n$ **elementary matrix** is a matrix obtained from the identity matrix by performing *one* single elementary row operation.

Exercise 98
Show that the following matrices are elementary matrices
(a)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

(b)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

(c)

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

Solution.
We list the operations that produce the given elementary matrices.
(a) $r_1 \leftarrow 1r_1$.
(b) $r_2 \leftrightarrow r_3$.
(c) $r_1 \leftarrow r_1 + 3r_3$ ■

Exercise 99
Consider the matrix

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

(a) Find the row equivalent matrix to A obtained by adding 3 times the first row of A to the third row. Call the equivalent matrix B.
(b) Find the elementary matrix E corresponding to the above elementary row operation.
(c) Compare EA and B.

Solution.
(a)

$$B = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{pmatrix}$$

(b)

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

(c) $EA = B$ ■

The conclusion of the above exercise holds for any matrix of size $m \times n$.

Theorem 17
If the elementary matrix E results from performing a certain row operations on I_m and if A is an $m \times n$ matrix, then the product of EA is the matrix that results when this same row operation is performed on A.

Proof.
We prove the theorem only for type II operations. Let E be the elementary matrix of size $m \times m$ obtained by adding k times row p of I_m to row q. Let B be the matrix obtained from A by adding k times row p to row q. We must show that $B = EA$. Let $A = (a_{ij})$. Then $B = (b_{ij})$, where $b_{ij} = a_{ij}$ for $i \neq q$, and $b_{qj} = ka_{pj} + a_{qj}$. Let $EA = (d_{ij})$. We will show that $b_{ij} = d_{ij}$. If $i \neq q$ then $d_{ij} = (ith\ row\ of\ E)(jth\ column\ of\ A) = a_{ij} = b_{ij}$. If $i = q$, then $d_{qj} = (qth\ row\ of\ E)(jth\ column\ of\ A) = ka_{pj} + a_{qj} = b_{qj}$. It follows that $EA = B$. Types I and III are left for the reader ■

It follows from the above theorem that a matrix A is row equivalent to a matrix B if and only if $B = E_kE_{k-1}\cdots E_1A$, where $E_1, E_2, \cdots, E_k$ are elementary matrices.
The above theorem is primarily of theoretical interest and will be used for developping some results about matrices and systems of linear equations. From a computational point of view, it is preferred to perform row operations directly rather than multiply on the left by an elementary matrix. Also, this theorem says that an elementary row operation on A can be achieved by premultiplying A by the corresponding elementary matrix E.
Given any elementary row operation, there is another row operation (called its **inverse**) that reverse the effect of the first operation. The inverses are described in the following chart.

Type	*Operation*	*Inverse operation*
I	$r_i \leftarrow cr_i$	$r_i \leftarrow \frac{1}{c}r_i$
II	$r_j \leftarrow cr_i + r_j$	$r_j \leftarrow -cr_i + r_j$
III	$r_i \leftrightarrow r_j$	$r_i \leftrightarrow r_j$

The following theorem gives an important property of elementary matrices.

Theorem 18
Every elementary matrix is invertible, and the inverse is an elementary matrix.

Proof.
Let A be any $n \times n$ matrix. Let E be an elementary matrix obtained by applying a row elementary operation ρ on I_n. By Theorem 17, applying ρ on A produces a matrix EA. Applying the inverse operation ρ^{-1} to EA gives $F(EA)$ where F is the elementary matrix obtained from I_n by applying the operation ρ^{-1}. Since inverse row operations cancel the effect of each other, it follows that $FEA = A$. Since A was arbitrary, we can choose $A = I_n$. Hence, $FE = I_n$. A similar argument shows that $EF = I_n$. Hence E is invertible and $E^{-1} = F$ ■

Exercise 100
Show the following
(a) A is row equivalent to A.
(b) If A is row equivalent to B then B is row equivalent to A.
(c) If A is row equivalent to B and B is row equivalent to C then A is row equivalent to C.

Solution.
(a) Suppose that A is an $m \times n$ matrix. Then $A = I_m A$. Since I_m is an elementary matrix then $A \sim A$.
(b) Suppose that A and B are two matrices of sizes $m \times n$. If $A \sim B$ then $B = E_k E_{k-1} \cdots E_1 A$ where $E_1, E_2, \cdots, E_k$ are $m \times m$ elementary matrices. Since these matrices are invertible elementary matrices then $A = E_1^{-1} E_2^{-1} \cdots E_k^{-1} B$. That is, $B \sim A$.
(c) Suppose that $A \sim B$ and $B \sim C$. Then $B = E_k \cdots E_1 A$ and $C = F_l F_{l-1} \cdots F_1 B$. But then we have $C = (F)l F_{l-1} \cdots F_1 E_k E_{k-1} \cdots E_1 A$. That is, $A \sim C$ ■

Exercise 101
Write down the inverses of the following elementary matrices:

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Solution.
(a) $E_1^{-1} = E_1$.
(b)

$$E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{9} \end{pmatrix}$$

(c)

$$E_3^{-1} = \begin{pmatrix} 1 & 0 & -5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ■$$

Exercise 102
If E is an elementary matrix show that E^T is also an elementary matrix of the same type.

Solution.
Suppose that E is the elementary matrix obtained by interchanging rows i and j of I_n with $i < j$. This is equivalent to interchanging columns i and j of I_n. But then E^T is obtained by interchanging rows i and j of I_n and so is an elementary matrix. If E is obtained by multiplying the ith row of I_n by a nonzero constant k then this is the same thing as multiplying the ith column of I_n by k. Thus, E^T is obtained by multiplying the ith row of I_n by k and so is an elementary matrix. Finally, if E is obtained by adding k times the ith row of I_n to the jth row then E^T is obtained by adding k times the jth row of I_n to the ith row. Note that if E is of Type I or Type III then $E^T = E$ ■

2.5 An Algorithm for Finding A^{-1}

Before we establish the main results of this section, we recall the reader of the following method of mathematical proofs. To say that statements $p_1, p_2, \cdots, p_n$

are all equivalent means that either they are all true or all false. To prove that they are equivalent, one assumes p_1 to be true and proves that p_2 is true, then assumes p_2 to be true and proves that p_3 is true, continuing in this fashion, assume that p_{n-1} is true and prove that p_n is true and finally, assume that p_n is true and prove that p_1 is true. This is known as the proof by circular argument. Now, back to our discussion of inverses. The following result establishes relationships between square matrices and systems of linear equations. These relationships are very important and will be used many times in later sections.

Theorem 19
If A is an $n \times n$ matrix then the following statements are equivalent.

(a) A is invertible.
(b) $Ax = \mathbf{0}$ has only the trivial solution.
(c) A is row equivalent to I_n.
(d) $rank(A) = n$.

Proof.
$(a) \Rightarrow (b)$: Suppose that A is invertible and x_0 is a solution to $Ax = \mathbf{0}$. Then $Ax_0 = \mathbf{0}$. Multiply both sides of this equation by A^{-1} to obtain $A^{-1}Ax_0 = A^{-1}\mathbf{0}$, that is $x_0 = \mathbf{0}$. Hence, the trivial solution is the only solution.

$(b) \Rightarrow (c)$: Suppose that $Ax = \mathbf{0}$ has only the trivial solution. Then the reduced row-echelon form of the augmented matrix has no rows of zeros or free variables. Hence it must look like

$$\left(\begin{array}{cccccc} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & 1 & 0 \end{array}\right)$$

If we disregard the last column of the previous matrix we can conclude that A can be reduced to I_n by a sequence of elementary row operations, i.e. A is row equivalent to I_n.

$(c) \Rightarrow (d)$: Suppose that A is row equivalent to I_n. Then $rank(A) = rank(I_n) = n$.

$(d) \Rightarrow (a)$: Suppose that $rank(A) = n$. Then A is row equivalent to I_n. That is I_n is obtained by a finite sequence of elementary row operations performed on A. Then by Theorem 17, each of these operations can be accomplished by premultiplying on the left by an appropriate elementary matrix. Hence, obtaining

$$E_kE_{k-1}\dots E_2E_1A = I_n,$$

where k is the necessary number of elementary row operations needed to reduce A to I_n. Now, by Theorem 18 each E_i is invertible. Hence, $E_kE_{k-1}\dots E_2E_1$ is invertible and $A^{-1} = E_k E_{k-1} \dots E_2 E_1$. ■

Using the definition, to show that an $n \times n$ matrix A is invertible we find a matrix B of the same size such that $AB = I_n$ and $BA = I_n$. The next theorem shows that one of these equality is enough to assure invertibilty.

Theorem 20
If A and B are two square matrices of size $n \times n$ such that $AB = I$ then $BA = I_n$ and $B^{-1} = A$.

Proof
Suppose that $Bx = \mathbf{0}$. Multiply both sides by A to obtain $ABx = \mathbf{0}$. That is, $x = \mathbf{0}$. This shows that the homogenenous system $Bx = \mathbf{0}$ has only the trivial solution so by Theorem 9 we see that B is invertible, say with inverse C. Hence, $C = I_n C = (AB)C = A(BC) = AI_n = A$ so that $B^{-1} = A$. Thus, $BA = BB^{-1} = I_n$.■

As an application of Theorem 19, we describe an algorithm for finding A^{-1}. We perform elementary row operations on A until we get I_n; say that the product of the elementary matrices is $E_k E_{k-1} \dots E_2 E_1$. Then we have

$$\begin{aligned}(E_k E_{k-1} \dots E_2 E_1)[A|I_n] &= [(E_k E_{k-1} \dots E_2 E_1)A|(E_k E_{k-1} \dots E_2 E_1)I_n] \\ &= [I_n|A^{-1}]\end{aligned}$$

We ask the reader to carry the above algorithm in solving the following problems.

Exercise 103
Find the inverse of

$$A = \left(\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{array}\right)$$

Solution.
We first construct the matrix

$$\left(\begin{array}{ccccccc} 1 & 2 & 3 & | & 1 & 0 & 0 \\ 2 & 5 & 3 & | & 0 & 1 & 0 \\ 1 & 0 & 8 & | & 0 & 0 & 1 \end{array}\right)$$

Applying the above algorithm to obtain

Step 1: $r_2 \leftarrow r_2 - 2r_1$ and $r_3 \leftarrow r_3 - r_1$

$$\left(\begin{array}{ccccccc} 1 & 2 & 3 & | & 1 & 0 & 0 \\ 0 & 1 & -3 & | & -2 & 1 & 0 \\ 0 & -2 & 5 & | & -1 & 0 & 1 \end{array}\right)$$

Step 2: $r_3 \leftarrow r_3 + 2r_2$

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

Step 3: $r_1 \leftarrow r_1 - 2r_2$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 9 & 5 & -2 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right)$$

Step 4: $r_2 \leftarrow r_2 - 3r_3$ and $r_1 \leftarrow r_1 + 9r_3$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right)$$

Step 5: $r_3 \leftarrow -r_3$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right)$$

It follows that

$$A^{-1} = \left(\begin{array}{ccc} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{array}\right) \blacksquare$$

Exercise 104

Show that the following homogeneous system has only the trivial solution.

$$\begin{array}{ccccccc} x_1 & + & 2x_2 & + & 3x_3 & = & 0 \\ 2x_1 & + & 5x_2 & + & 3x_3 & = & 0 \\ x_1 & & & + & 8x_3 & = & 0. \end{array}$$

Solution.

The coefficient matrix of the given system is invertible by the previous exercise. Thus, by Theorem 19 the system has only the trivial solution $\blacksquare$

The following result exhibit a criterion for checking the singularity of a square matrix.

Theorem 21

If A is a square matrix with a row consisting entirely of zeros then A is singular.

Proof.

The reduced row-echelon form will have a row of zeros. So the rank of the coefficient matrix of the homogeneous system $Ax = \mathbf{0}$ is less than n. By Theorem 5, $Ax = \mathbf{0}$ has a nontrivial solution and as a result of Theorem 19 the matrix A

must be singular. ■

How can we tell when a square matrix A is singular? i.e., when does the algorithm of finding A^{-1} fail? The answer is provided by the following theorem

Theorem 22
An $n \times n$ matrix A is singular if and only if A is row equivalent to a matrix B that has a row of zeros.

Proof.
Suppose first that A is singular. Then by Theorem 19, A is not row equivalent to I_n. Thus, A is row equivalent to a matrix $B \neq I_n$ which is in reduced echelon form. By Theorem 19, B must have a row of zeros.
Conversely, suppose that A is row equivalent to matrix B with a row consisting entirely of zeros. Then B is singular by Theorem 251. Now, $B = E_kE_{k-1}\ldots E_2E_1A$. If A is nonsingular then B is nonsingular, a contradiction. Thus, A must be singular ■

The following theorem establishes a result of the solvability of linear systems using the concept of invertibility of matrices.

Theorem 23
An $n \times n$ square matrix A is invertible if and only if the linear system $Ax = b$ is consistent for every $n \times 1$ matrix b.

Proof.
Suppose first that A is invertible. Then for any $n \times 1$ matrix b the linear system $Ax = b$ has a unique solution, namely $x = A^{-1}b$.
Conversely, suppose that the system $Ax = b$ is solvable for any $n \times 1$ matrix b. In particular, $Ax_i = e_i, 1 \leq i \leq n$, has a solution, where e_i is the ith column of I_n. Construct the matrix

$$C = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

Then

$$AC = \begin{pmatrix} Ax_1 & Ax_2 & \cdots & Ax_n \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & \cdots & e_n \end{pmatrix} = I_n.$$

Hence, by Theorem 20, A is non-singular ■

Exercise 105
Solve the following system by using the previous theorem

$$\left\{\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & = & 5 \\ 2x_1 & + & 5x_2 & + & 3x_3 & = & 3 \\ x_1 & & & + & 8x_3 & = & 17 \end{array}\right.$$

Solution.
Using Exercise 103 and Theorem 23 we have

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 17 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \blacksquare$$

Exercise 106
If P is an $n \times n$ matrix suxh that $P^TP = I_n$ then the matrix $H = I_n - 2PP^T$ is called the **Householder matrix**. *Show that H is symmetric and $H^TH = I_n$.*

Solution.
Taking the transpose of H we have $H^T = I_n^T - 2(P^T)^TP^T = H$. That is, H is symmetric. On the other hand, $H^TH = H^2 = (I_n - 2PP^T)^2 = I_n - 4PP^T + 4(PP^T)^2 = I_n - 4PP^T + 4P(P^TP)P^T = I_n - 4PP^T + 4PP^T = I_n$ $\blacksquare$

Exercise 107
Let A and B be two square matrices. Show that AB is nonsingular if and only if both A and B are nonsingular.

Solution.
Suppose that AB is nonsingular. Then $(AB)^{-1}AB = AB(AB)^{-1} = I_n$. Suppose that A is singular. Then A has a row consisting of 0. Then AB has a row consisting of 0 (Exercise 90) but then AB is singular according to Theorem 251, a contradiction. Since $(AB)^{-1}AB = I_n$ then by Theorem 20, B is nonsingular. The converse is just Theorem 15 (a) $\blacksquare$

2.6 Review Problems

Exercise 108
Compute the matrix

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

Exercise 109
Find $w, x, y,$ and z.

$$\begin{pmatrix} 1 & 2 & w \\ 2 & x & 4 \\ y & -4 & z \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -3 & 4 \\ 0 & -4 & 5 \end{pmatrix}$$

Exercise 110
Determine two numbers s and t such that the following matrix is symmetric.

$$A = \begin{pmatrix} 2 & s & t \\ 2s & 0 & s+t \\ 3 & 3 & t \end{pmatrix}$$

Exercise 111
Let A be a 2×2 matrix. Show that

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

Exercise 112
Let $A = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}, C = \begin{bmatrix} 3 & 0 & 1 \end{bmatrix}$. If $rA+sB+tC = \mathbf{0}$ show that $s = r = t = 0$.

Exercise 113
Show that the product of two diagonal matrices is again a diagonal matrix.

Exercise 114
Let A be an arbitrary matrix. Under what conditions is the product AA^T defined?

Exercise 115
(a) Show that $AB = BA$ if and only if $(A-B)(A+B) = A^2 - B^2$.
(b) Show that $AB = BA$ if and only if $(A+B)^2 = A^2 + 2AB + B^2$.

Exercise 116
Let A be a matrix of size $m \times n$. Denote the columns of A by $C_1, C_2, \cdots, C_n$. Let x be the $n \times 1$ matrix with entries $x_1, x_2, \cdots, x_n$. Show that $Ax = x_1C_1 + x_2C_2 + \cdots + x_nC_n$.

Exercise 117
Let A be an $m \times n$ matrix. Show that if $yA = \mathbf{0}$ for all $y \in \mathbb{R}^m$ then $A = \mathbf{0}$.

Exercise 118
An $n \times n$ matrix A is said to be **idempotent** *if $A^2 = A$.*
(a) Show that the matrix

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

is idempotent
(b) Show that if A is idempotent then the matrix $(I_n - A)$ is also idempotent.

Exercise 119
The purpose of this exercise is to show that the rule $(ab)^n = a^n b^n$ *does not hold with matrix multiplication. Consider the matrices*

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

Show that $(AB)^2 \neq A^2B^2$.

Exercise 120
Show that $AB = BA$ *if and only if* $A^TB^T = B^TA^T$.

Exercise 121
Suppose $AB = BA$ *and* n *is a non-negative integer.*
(a) Use induction to show that $AB^n = B^nA$.
(b) Use induction and (a) to show that $(AB)^n = A^nB^n$.

Exercise 122
Let A *and* B *be symmetric matrices. Show that* AB *is symmetric if and only if* $AB = BA$.

Exercise 123
Show that $tr(AA^T)$ *is the sum of the squares of all the entries of* A.

Exercise 124
(a) Find two 2×2 *singular matrices whose sum in nonsingular.*
(b) Find two 2×2 *nonsingular matrices whose sum is singular.*

Exercise 125
Show that the matrix

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

is singular.

Exercise 126
Let

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

Find A^{-3}.

Exercise 127
Let

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

Find A.

Exercise 128
Let A and B be square matrices such that $AB = \mathbf{0}$. Show that if A is invertible then B is the zero matrix.

Exercise 129
Find the inverse of the matrix

$$A = \left(\begin{array}{cc} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{array}\right)$$

Exercise 130
Which of the following are elementary matrices?
(a)

$$\left(\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right)$$

(b)

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

(c)

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 9 \\ 0 & 0 & 1 \end{array}\right)$$

(d)

$$\left(\begin{array}{cccc} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Exercise 131
Let A be a 4×3 matrix. Find the elementary matrix E, which as a premultiplier of A, that is, as EA, performs the following elementary row operations on A:
(a) Multiplies the second row of A by -2.
(b) Adds 3 times the third row of A to the fourth row of A.
(c) Interchanges the first and third rows of A.

Exercise 132
For each of the following elementary matrices, describe the corresponding elementary row operation and write the inverse.
(a)

$$E = \left(\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right)$$

(b)

$$E = \left(\begin{array}{ccc} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

(c)

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Exercise 133
Consider the matrices

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

Find elementary matrices $E_1, E_2, E_3,$ *and* E_4 *such that*
$(a) E_1 A = B, (b) E_2 B = A, (c) E_3 A = C, (d) E_4 C = A.$

Exercise 134
Determine if the following matrix is invertible.

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

Exercise 135
For what values of a *does the following homogeneous system have a nontrivial solution?*

$$\left\{ \begin{array}{lllll} (a-1)x_1 & + & 2x_2 & = & 0 \\ 2x_1 & + & (a-1)x_2 & = & 0 \end{array} \right.$$

Exercise 136
Find the inverse of the matrix

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

Exercise 137
What conditions must b_1, b_2, b_3 *satisfy in order for the following system to be consistent?*

$$\left\{ \begin{array}{lllllll} x_1 & + & x_2 & + & 2x_3 & = & b_1 \\ x_1 & & & + & x_3 & = & b_2 \\ 2x_1 & + & x_2 & + & 3x_3 & = & b_3 \end{array} \right.$$

Exercise 138
Prove that if A *is symmetric and nonsingular than* A^{-1} *is symmetric.*

Exercise 139
If

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

find D^{-1}.

Exercise 140
Prove that a square matrix A is nonsingular if and only if A is a product of elementary matrices.

Exercise 141
Prove that two $m \times n$ matrices A and B are row equivalent if and only if there exists a nonsingular matrix P such that $B = PA$.

Exercise 142
Let A and B be two $n \times n$ matrices. Suppose A is row equivalent to B. Prove that A is nonsingular if and only if B is nonsingular.

Exercise 143
Show that the product of two lower (resp. upper) triangular matrices is again lower (resp. upper) triangular.

Exercise 144
Show that a 2×2 lower triangular matrix is invertible if and only if $a_{11}a_{22} \neq 0$ and in this case the inverse is also lower triangular.

Exercise 145
Let A be an $n \times n$ matrix and suppose that the system $Ax = \mathbf{0}$ has only the trivial solution. Show that $A^k x = \mathbf{0}$ has only the trivial solution for any positive integer k.

Exercise 146
Show that if A and B are two $n \times n$ matrices then $A \sim B$.

Exercise 147
Show that any $n \times n$ invertible matrix A satisfies the property (P): If $AB = AC$ then $B = C$.

Exercise 148
Show that an $n \times n$ matrix A is invertible if and only if the equation $yA = \mathbf{0}$ has only the trivial solution.

Exercise 149
Let A be an $n \times n$ matrix such that $A^n = \mathbf{0}$. Show that $I_n - A$ is invertible and find its inverse.

Exercise 150
Let $A = (a_{ij}(t))$ be an $m \times n$ matrix whose entries are differentiable functions of the variable t. We define $\frac{dA}{dt} = (\frac{da_{ij}}{dt})$. Show that if the entries in A and B are differentiable functions of t and the sizes of the matrices are such that the stated operations can be performed, then
(a) $\frac{d}{dt}(kA) = k\frac{dA}{dt}$.
(b) $\frac{d}{dt}(A + B) = \frac{dA}{dt} + \frac{dB}{dt}$.
(c) $\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}$.

Exercise 151
Let A be an $n \times n$ invertible matrix with entries being differentiable functions of t. Find a formula for $\frac{dA^{-1}}{dt}$.

Exercise 152 *(Sherman-Morrison formula)*
Let A be an $n \times n$ invertible matrix. Let u, v be $n \times 1$ matrices such that $v^T A^{-1} u + 1 \neq 0$. Show that
(a) $(A + uv^T)(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}) = I_n$.
(b) Deduce from (a) that A is invertible.

Exercise 153
Let $x = (x_1, x_2, \cdots x_m)^T$ be an $m \times 1$ matrix and $y = (y_1, y_2, \cdots, y_n)^T$ be an $n \times n$ matrix. Construct the $m \times n$ matrix xy^T.

Exercise 154
Show that a triangular matrix A with the property $A^T = A^{-1}$ is always diagonal.

Chapter 3

Determinants

With each square matrix we can associate a real number called the determinant of the matrix. Determinants have important applications to the theory of systems of linear equations. More specifically, determinants give us a method (called Cramer's method) for solving linear systems. Also, determinant tells us whether or not a matrix is invertible.

3.1 Definition of the Determinant

Determinants can be used to find the inverse of a matrix and to determine if a linear system has a unique solution. Throughout this chapter we use only square matrices.

Let us first recall from probability theory the following important formula known as the **multiplication rule of counting**: If a choice consists of k steps, of which the first can be made in n_1 ways, for each of these the second step can be made in n_2 ways, $\cdots$, and for each of these the kth can be made in n_k ways, then the choice can be made in $n_1 n_2 \cdots n_k$ ways.

Exercise 155
If a new-car buyer has the choice of four body styles, three engines, and ten colors, in how many different ways can s/he order one of these cars?

Solution.
By the multiplication rule of counting, there are $(4)(3)(10) = 120$ choices ■

A **permutation** of the set $S = \{1, 2, \ldots, n\}$ is an arrangement of the elements of S in some order without omissions or repetitions. We write $\sigma = (\sigma(1)\sigma(2)\cdots\sigma(n))$.
In terms of functions, a permutation is a function on S with range equals to S. Thus, any σ on S possesses an inverse σ^{-1}, that is $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = id$

where $id = (123\cdots n)$. So if $\sigma = (613452)$ then $\sigma^{-1} = (263451)$.(Show that $\sigma\circ\sigma^{-1} = \sigma^{-1}\circ\sigma = id$).

Exercise 156
Let $\sigma = (24513)$ *and* $\tau = (41352)$ *be permutations in* S_5. *Find*
(a) $\sigma\circ\tau$ *and* $\tau\circ\sigma$,
(b) σ^{-1}.

Solution.
(a) Using the definition of composition of two functions we find, $\sigma(\tau(1)) = \sigma(4) = 1$. A similar argument for $\sigma(\tau(2))$, etc. Thus, $\sigma\circ\tau = (12534)$ and $\tau\circ\sigma = (15243)$.
(b) Since $\sigma(4) = 1$ then $\sigma^{-1}(1) = 4$. Similarly, $\sigma^{-1}(2) = 1$, etc. Hence, $\sigma^{-1} = (41523)$ ■

Let S_n denote the set of all permutations on S. How many permutations are there in S_n? We have n positions to be filled by n numbers. For the first position, there are n possibilities. For the second there are $n-1$ possibilities, etc. Thus, according to the multiplication rule of counting there are

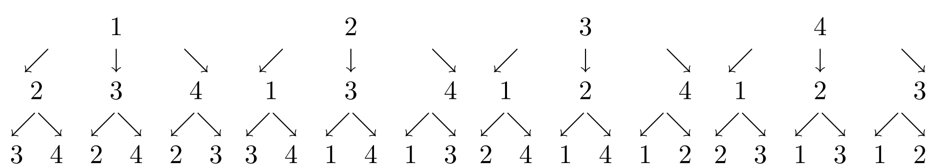
$$n(n-1)(n-2)\dots 2.1 = n!$$

permutations.

Is there a way to list all the permutations of S_n? The answer is yes and one can find the permutations by using a **permutation tree** which we describe in the following exercise

Exercise 157
List all the permutations of $S = \{1, 2, 3, 4\}$.

Solution.



An **inversion** is said to occur whenever a larger integer precedes a smaller one. If the number of inversions is even (resp. odd) then the permutation is said to be **even** (**resp. odd**). We define the **sign** of a permutation to be a function sgn with domain S_n and range $\{-1, 1\}$ such that $sgn(\sigma) = -1$ if σ is odd and $sgn(\sigma) = +1$ if σ is even.

Exercise 158
Classify each of the permutations in S_6 *as even or odd.*

(i) (613452).
(ii) (2413).
(iii) (1234).
(iv) $id = (123456)$.

Solution.
(i) $sgn((613452)) = +1$ since there are 8 inversions, namely, $(61), (63), (64), (65), (62), (32), (42)$, and (52).
(ii) $sgn((2413)) = -1$. The inversions are: $(21), (41), (43)$.
(iii) $sgn((1234)) = +1$. There are 0 inversions.
(iv) $sgn(id) = +1$ ■

The following two theorems are of theoretical interest.

Theorem 24
Let τ and σ be two elements of S_n. Then $sgn(\tau \circ \sigma) = sgn(\tau)sgn(\sigma)$. Hence, the composition of two odd or even permutations is always even.

Proof.
Define the polynomial in n indeterminates $g(x_1, x_2, \cdots, x_n) = \Pi_{i<j}(x_i - x_j)$. For any permutation $\sigma \in S_n$ we define $\sigma(g) = \Pi_{i<j}(x_{\sigma(i)} - x_{\sigma(j)})$. We will show that for any $\sigma \in S_n$ one has $\sigma(g) = sgn(\sigma)g$. Clearly, $\sigma(g) = (\Pi_{i<j}(x_i - x_j))(\Pi_{i>j}(x_i - x_j))$. But for $i > j$ one can write $x_i - x_j = -(x_j - x_i)$, with $(x_j - x_i)$ being a factor of g. Let k be the number of factors of the form $(x_i - x_j)$ with $i > j$. Then $\sigma(g) = (-1)^k g$. It follows that if k is even then $\sigma(g) = g$ and if k is odd $\sigma(g) = -g$.
Now, let $1 \leq i < j \leq n$ be such that $\sigma(i) > \sigma(j)$. Since $Range(\sigma) = \{1, 2, \cdots, n\}$ then there exit $i^*, j^* \in \{1, 2, \cdots, n\}$ such that $\sigma(i^*) = i$ and $\sigma(j^*) = j$. Since $i < j$ then $\sigma(j^*) > \sigma(i^*)$. Thus for every pair (i, j) with $i < j$ and $\sigma(i) > \sigma(j)$ there is a factor of $\sigma(g)$ of the form $(x_{\sigma(j^*)} - x_{\sigma(i^*)})$ with $\sigma(j^*) > \sigma(i^*)$. If σ is even (resp. odd) then the number of (i, j) with the property $i < j$ and $\sigma(i) > \sigma(j)$ is even (resp. odd). This implies that k is even (resp. odd). Thus, $\sigma(g) = g$ if σ is even and $\sigma(g) = -g$ if σ is odd.
Finally, for any $\sigma, \tau \in S_n$ we have

$$sgn(\sigma \circ \tau)g = (\sigma \circ \tau)g = \sigma(\tau(g)) = \sigma(sgn(\tau)g) = (sgn(\sigma))(sgn(\tau))g.$$

Since g was arbitrary then $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$. This concludes a proof of the theorem ■

Theorem 25
For any $\sigma \in S_n$, we have $sgn(\sigma) = sgn(\sigma^{-1})$.

proof.
We have $\sigma \circ \sigma^{-1} = id$. By Theorem 24 we have $sgn(\sigma)sgn(\sigma^{-1}) = 1$ since id is even. Hence, σ and σ^{-1} are both even or both odd ■

Let A be an $n \times n$ matrix. An **elementary product** from A is a product of n entries from A, no two of which come from the same row or same column.

Exercise 159
List all elementary products from the matrices
(a)

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

(b)

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Solution.
(a) The only elementary products are $a_{11}a_{22}, a_{12}a_{21}$.
(b) An elementary product has the form $a_{1*}a_{2*}a_{3*}$. Since no two factors come from the same column, the column numbers have no repetitions; consequently they must form a permutation of the set $\{1, 2, 3\}$. The $3! = 6$ permutations yield the following elementary products:
$a_{11}a_{22}a_{33}, a_{11}a_{23}a_{32}, a_{12}a_{23}a_{31}, a_{12}a_{21}a_{33}, a_{13}a_{21}a_{32}, a_{13}a_{22}a_{31}$ ■

Let A be an $n \times n$ matrix. Consider an elementary product of entries of A. For the first factor, there are n possibilities for an entry from the first row. Once selected, there are $n - 1$ possibilities for an entry from the second row for the second factor. Continuing, we find that there are $n!$ elementary products. They are the products of the form $a_{1\sigma(1)}a_{2\sigma(2)} \dots a_{n\sigma(n)}$, where σ is a permutation of $\{1, 2, \dots, n\}$, i.e. a member of S_n.

Let A be an $n \times n$ matrix. Then we define the **determinant** of A to be the number

$$|A| = \sum sgn(\sigma)a_{1\sigma(1)}a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

where the sum is over all permutations σ of $\{1, 2, \dots, n\}$.

Exercise 160
Find $|A|$ if
(a)

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

(b)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Solution.
By using the definition of a determinant and Exercise 159 we obtain
(a) $|A| = a_{11}a_{22} - a_{21}a_{12}$.
(b) $|A| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$
■

Exercise 161
Find all values of λ such that $|A| = 0$.
(a)

$$A = \left(\begin{array}{cc} \lambda - 1 & -2 \\ 1 & \lambda - 4 \end{array}\right),$$

(b)

$$A = \left(\begin{array}{ccc} \lambda - 6 & 0 & 0 \\ 0 & \lambda & -1 \\ 0 & 4 & \lambda - 4 \end{array}\right)$$

Solution.
(a) Using Exercise 160 (a) we find $|A| = (\lambda - 3)(\lambda - 2) = 0$. Therefore, $\lambda = 3$ or $\lambda = 2$.
(b) By Exercise 160 (b) we have $|A| = (\lambda - 2)^2(\lambda - 6) = 0$ and this yields $\lambda = 2$ or $\lambda = 6$ ■

Exercise 162
Prove that if a square matrix A has a row of zeros then $|A| = 0$.

Solution.
Suppose that the ith row of A consists of 0. By the definition of the determinant each elementary product will have one factor belonging to the ith row. Thus, each elementary product is zero and consequently $|A| = 0$ ■

3.2 Evaluating Determinants by Row Reduction

In this section we provide a simple procedure for finding the determinant of a matrix. The idea is to reduce the matrix into row-echelon form which in this case is a triangular matrix. Recall that a matrix is said to be **triangular** if it is upper triangular, lower triangular or diagonal. The following theorem provides a formula for finding the determinant of a triangular matrix.

Theorem 26
If A is an $n \times n$ triangular matrix then $|A| = a_{11}a_{22}\dots a_{nn}$.

Proof.
Assume that A is lower triangular. We will show that the only elementary

product that appears in the formula of $|A|$ is $a_{11}a_{22}\cdots a_{nn}$. To see this, consider a typical elementary product

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}. \tag{3.1}$$

Since $a_{12} = a_{13} = \cdots = a_{1n} = 0$ then for the above product to appear we must have $\sigma(1) = 1$. But then $\sigma(2) \neq 1$ since no two factors come from the same column. Hence, $\sigma(2) \geq 2$. Further, since $a_{23} = a_{24} = \cdots = a_{2n} = 0$ we must have $\sigma(2) = 2$ in order to have (3.1). Continuing in this fashion we obtain $\sigma(3) = 3, \cdots, \sigma(n) = n$. That is, $|A| = a_{11}a_{22}\cdots a_{nn}$. The proof for an upper triangular matrix is similar.■

Exercise 163
Compute $|A|$.
(a)

$$A = \left(\begin{array}{ccc} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{array}\right),$$

(b)

$$A = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{array}\right),$$

Solution.
(a) Since A is triangular then $|A| = (1)(4)(6) = 24$.
(b) $|A| = (1)(3)(6) = 18$ ■

Exercise 164
Compute the determinant of the identity matrix I_n.

Solution.
Since the identity matrix is triangular with entries equal to 1 on the main diagonal then $|I_n| = 1$ ■

The following theorem is of practical use. It provides a technique for evaluating determinants by greatly reducing the labor involved. We shall show that the determinant can be evaluated by reducing the matrix to row-echelon form. Before proving the next theorem, we make the following remark. Let τ be a permutation that interchanges only two numbers i and j with $i < j$. That is, $\tau(k) = k$ if $k \neq i, j$ and $\tau(i) = j, \tau(j) = i$. We call such a permutation a **transposition**. In this case, $\tau = (12\cdots j(i+1)(i+2)\cdots(j-1)i(j+1)\cdots n)$. Hence, there are $2(j-i-1)+1$ inversions, namely, $(j, i+1), \cdots, (j, j-1), (j, i), (i+1, i), (i+2, i), \cdots, (j-1, i)$. This means that τ is odd. Now, for any permutation $\sigma \in S_n$ we have $sgn(\tau \circ \sigma) = sgn(\tau)sgn(\sigma) = -sgn(\sigma)$.(See Theorem 24)

Theorem 27
Let A be an $n \times n$ matrix.
(a) Let B be the matrix obtained from A by multiplying a row by a scalar c. Then $|B| = c|A|$.
(b) Let B be the matrix obtained from A by interchanging two rows of A. Then $|B| = -|A|$.
(c) If a square matrix has two identical rows then its determinant is zero.
(d) Let B be the matrix obtained from A by adding c times a row to another row. Then $|B| = |A|$.

Proof.
(a) Multiply row r of $A = (a_{ij})$ by a scalar c. Let $B = (b_{ij})$ be the resulting matrix. Then $b_{ij} = a_{ij}$ for $i \neq r$ and $b_{rj} = ca_{rj}$. Using the definition of a determinant we have

$$\begin{array}{rcl} |B| & = & \sum_{\sigma \in S_n} sgn(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{r\sigma(r)} \cdots b_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots c(a_{r\sigma(r)}) \cdots a_{n\sigma(n)} \\ & = & c\sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} \cdots a_{n\sigma(n)} = c|A|. \end{array}$$

(b) Interchange rows r and s of A with $r < s$. Let $B = (b_{ij})$ be the resulting matrix. Let τ be the transposition that interchanges the numbers r and s. Then $b_{ij} = a_{i\tau(j)}$. Using the definition of determinant we have

$$\begin{array}{rcl} |B| & = & \sum_{\sigma \in S_n} sgn(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a_{1\tau\circ\sigma(1)} a_{2\tau\circ\sigma(2)} \cdots a_{n\tau\circ\sigma(n)} \\ & = & -\sum_{\sigma \in S_n} sgn(\tau \circ \sigma) a_{1\tau\circ\sigma(1)} a_{2\tau\circ\sigma(2)} \cdots a_{n\tau\circ\sigma(n)} \\ & = & -\sum_{\delta \in S_n} sgn(\delta) a_{1\delta(1)} a_{2\delta(2)} \cdots a_{n\delta(n)} = -|A|. \end{array}$$

Note that as σ runs through all the elements of S_n, $\tau \circ \sigma$ runs through all the elements of S_n as well. This ends a proof of (b).
(c) Suppose that rows r and s of A are equal. Let B be the matrix obtained from A by interchanging rows r and s. Then $B = A$ so that $|B| = |A|$. But by (b), $|B| = -|A|$. Hence, $|A| = -|A|, 2|A| = 0$ and hence $|A| = 0$.
(d) Let $B = (b_{ij})$ be obtained from A by adding to each element of the sth row of A c times the corresponding element of the rth row of A. That is, $b_{ij} = a_{ij}$ if $i \neq s$ and $b_{sj} = ca_{rj} + a_{sj}$. Then

$$\begin{array}{rcl} |B| & = & \sum_{\sigma \in S_n} sgn(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{s\sigma(s)} \cdots b_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots (ca_{r\sigma(s)} + a_{s\sigma(s)}) \cdots a_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} \cdots a_{s\sigma(s)} \cdots a_{n\sigma(n)} \\ & + & c(\sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} \cdots a_{r\sigma(s)} \cdots a_{n\sigma(n)}). \end{array}$$

The first term in the last expression is $|A|$. The second is c times the determinant of a matrix with two identical rows, namely, rows r and s with entries $(a_{r1}, a_{r2}, \cdots, a_{rn})$. By (c), this quantity is zero. Hence, $|B| = |A|$.■

Exercise 165

Use Theorem 27 to evaluate the determinant of the following matrix

$$A = \begin{pmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{pmatrix}$$

Solution.
We use Gaussian elimination as follows.

Step 1: $r_1 \leftrightarrow r_2$

$$\begin{vmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix} = -|A|$$

Step 2: $r_1 \leftarrow r_1 - r_3$

$$\begin{vmatrix} 1 & -12 & 8 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix} = -|A|$$

Step 3: $r_3 \leftarrow r_3 - 2r_1$

$$\begin{vmatrix} 1 & -12 & 8 \\ 0 & 1 & 5 \\ 0 & 30 & -15 \end{vmatrix} = -|A|$$

Step 4: $r_3 \leftarrow r_3 - 30r_2$

$$\begin{vmatrix} 1 & -12 & 8 \\ 0 & 1 & 5 \\ 0 & 0 & -165 \end{vmatrix} = -|A|$$

Thus,

$$|A| = -\begin{vmatrix} 1 & -12 & 8 \\ 0 & 1 & 5 \\ 0 & 0 & -165 \end{vmatrix} = 165 \blacksquare$$

Exercise 166
Show that if a square matrix has two proportional rows then its determinant is zero.

Solution.
Suppose that A is a square matrix such that row j is k times row i with $k \neq 0$. By adding $-\frac{1}{k}r_j$ to r_i then the ith row will consist of 0. By Thereom 27 (c), $|A| = 0$ $\blacksquare$

Exercise 167
Show that if A is an $n \times n$ matrix and c is a scalar then $|cA| = c^n|A|$.

Solution.
The matrix cA is obtained from the matrix A by multiplying the rows of A by $c \neq 0$. By mutliplying the first row of cA by $\frac{1}{c}$ we obtain $|B| = \frac{1}{c}|cA|$ where B

is obtained from the matrix A by multiplying all the rows of A, except the first one, by c. Now, divide the second row of B by $\frac{1}{c}$ to obtain $|B'| = \frac{1}{c}|B|$, where B' is the matrix obtained from A by multiplying all the rows of A, except the first and the second, by c. Thus, $|B'| = \frac{1}{c^2}|cA|$. Repeating this process, we find $|A| = \frac{1}{c^n}|cA|$ or $|cA| = c^n|A|$ ■

Exercise 168
(a) Let E_1 be the elementary matrix corresponding to type I elementary row operation. Find $|E_1|$.
(b) Let E_2 be the elementary matrix corresponding to type II elementary row operation. Find $|E_2|$.
(c) Let E_3 be the elementary matrix corresponding to type III elementary row operation. Find $|E_3|$.

Solution.
(a) The matrix E_1 is obtained from the identity matrix by multiplying a row of I_n by a nonzero scalar c. In this case,$|E_1| = c|I_n| = c$.
(b) E_2 is obtained from I_n by adding a multiple of a row to another row. Thus, $|E_2| = |I_n| = 1$.
(c) The matrix E_3 is obtained from the matrix I_n by interchanging two rows. In this case, $|E_3| = -|I_n| = -1$ ■

Exercise 169
Find, by inspection, the determinant of the following matrix.

$$A = \begin{pmatrix} 3 & -1 & 4 & -2 \\ 6 & -2 & 5 & 2 \\ 5 & 8 & 1 & 4 \\ -9 & 3 & -12 & 6 \end{pmatrix}$$

Solution.
Since the first and the fourth rows are proportional then the determinant is zero by Exercise 166 ■

3.3 Properties of the Determinant

In this section we shall derive some of the fundamental properties of the determinant. One of the immediate consequences of these properties will be an important determinant test for the invertibility of a square matrix.

Theorem 28
Let $A = (a_{ij})$ be an $n \times n$ square matrix. Then $|A^T| = |A|$.

Proof.
Since $A = (a_{ij})$ then $A^T = (a_{ji})$. Using the definition of determinant we have

$$|A^T| = \sum_{\sigma \in S_n} sgn(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

Now, since σ is a permutation then $a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n} = a_{1k_1}a_{2k_2}\cdots a_{nk_n}$, where $\sigma(k_i) = i, 1 \leq i \leq n$. We claim that $\tau = (k_1k_2\cdots k_n) = \sigma^{-1}$. Indeed, let $\tau(i) = k_i$ for $1 \leq i \leq n$. Then $(\sigma \circ \tau)(i) = \sigma(\tau(i)) = \sigma(k_i) = i$, i.e. $\sigma \circ \tau = id$. Hence, $\tau = \sigma^{-1}$.
Next, by Theorem 25, $sgn(\tau) = sgn(\sigma)$ and therefore

$$\begin{array}{rcl} |A^T| &=& \sum_{\sigma\in S_n} sgn(\tau)a_{1,\tau(1)}a_{2,\tau(2)}\cdots a_{n,\tau(n)} \\ &=& \sum_{\tau\in S_n} sgn(\tau)a_{1,\tau(1)}a_{2,\tau(2)}\cdots a_{n,\tau(n)} = |A|. \end{array}$$

This ends a proof of the theorem ■

It is worth noting here that Theorem 28 says that every property about determinants that contains the word "row" in its statement is also true when the word "column" is substituted for "row".

Exercise 170
Prove that if two columns of a square matrix are proportional then the matrix has determinant equals to zero.

Solution.
If A is a square matrix with two proportional columns then A^T has two proportional rows. By Exercise 166 and Theorem 28 we have $|A| = |A^T| = 0$ ■

Exercise 171
Show that if A is matrix with a column consisting of 0 then $|A| = 0$.

Solution.
If A has a column consisting entirely of 0 then A^T has a row consisitng entirely of 0. By Theorem 28 and Exercise 162 we have $|A| = |A^T| = 0$ ■

Next, we prove that the determinant of the product of two matrices is the product of determinants and that A is invertible if and only if A has nonzero determinant. The next result relates the determinant of a product of a matrix with an elementary matrix.

Theorem 29
If E is an elementary matrix then $|EA| = |E||A|$.

Proof.
If E is an elementary matrix of type I then $|E| = c$ and by Theorem 27 $|EA| = c|A| = |E||A|$. If E is an elementary matrix of type II, then $|E| = 1$ and $|EA| = |A| = |E||A|$. Finally, if E is an elementary matrix of type III then $|E| = -1$ and $|EA| = -|A| = |E||A|$■

By induction one can extend the above result.

Theorem 30
If $B = E_1E_2\cdots E_nA$, where $E_1, E_2, \cdots, E_n$ are elementary matrices, then $|B| = |E_1||E_2|\cdots|E_n||A|$.

Proof.
This follows from Theorem 29. Indeed, $|B| = |E_1(E_2\cdots E_nA)| = |E_1||E_2(E_3\cdots E_nA)| = \cdots = |E_1||E_2|\cdots|E_n||A|$ ■

Next we establish a condition for a matrix to be invertible.

Theorem 31
If A is an $n\times n$ matrix then A is nonsingular if and only if $|A| \neq 0$.

Proof
If A is nonsingular then A is row equivalent to I_n (Theorem 19). That is, $A = E_kE_{k-1}\cdots E_1I_n$. By Theorem 30 we have $|A| = |E_k|\cdots|E_1||I_n| \neq 0$ since if E is of type I then $|E| = c \neq 0$; if E is of type II then $|E| = 1$, and if E is of type III then $|E| = -1$.
Now, suppose that $|A| \neq 0$. If A is singular then by Theorem 22, A is row equivalent to a matrix B that has a row of zeros. Also, $B = E_kE_{k-1}\cdots E_1A$. By Theorem 30, $0 = |B| = |E_k||E_{k-1}|\cdots|A| \neq 0$, a contradiction. Hence, A must be nonsingular. ■

The following result follows from Theorems 19 and 31.

Theorem 32
The following statements are all equivalent:
(i) A is nonsingular.
(ii) $|A| \neq 0$.
(iii) $rank(A) = n$.
(iv) A is row equivalent to I_n.
(v) The homogeneous systen $Ax = \mathbf{0}$ has only the trivial solution.

Exercise 172
Prove that $|A| = 0$ if and only if $Ax = \mathbf{0}$ has a nontrivial solution.

Solution.
If $|A| = 0$ then according to Theorem 32 the homogeneous system $Ax = \mathbf{0}$ must have a nontrivial solution. Conversely, if the homogeneous system $Ax = \mathbf{0}$ has a nontrivial solution then A must be singular by Theorem 32. By Theorem 32 (a), $|A| = 0$ ■

Theorem 33
If A and B are $n\times n$ matrices then $|AB| = |A||B|$.

Proof.
If A is singular then AB is singular and therefore $|AB| = 0 = |A||B|$ since $|A| =$

0. So assume that A is nonsingular. Then by Theorem 32, $A = E_k E_{k-1} \cdots E_1 I_n$. By Theorem 30 we have $|A| = |E_k||E_{k-1}| \cdots |E_1|$. Thus, $|AB| = |E_k E_{k-1} \cdots E_1 B| = |E_k||E_{k-1}| \cdots |E_1||B| = |A||B|$. ■

Exercise 173
Is it true that $|A + B| = |A| + |B|$?

Solution.
No. Consider the following matrices.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Then $|A + B| = |\mathbf{0}| = 0$ and $|A| + |B| = -2$ ■

Exercise 174
Show that if A is invertible then $|A^{-1}| = \frac{1}{|A|}$.

Solution.
If A is invertible then $A^{-1}A = I_n$. Taking the determinant of both sides we find $|A^{-1}||A| = 1$. That is, $|A^{-1}| = \frac{1}{|A|}$. Note that since A is invertible then $|A| \neq 0$ ■

Exercise 175
Let A and B be two **similar** *matrices, i.e. there exists a nonsingular matrix P such that $A = P^{-1}BP$. Show that $|A| = |B|$.*

Solution.
Using Theorem 33 and Exercise 174 we have, $|A| = |P^{-1}BP| = |P^{-1}||B||P| = \frac{1}{|P|}|B||P| = |B|$. Note that since P is nonsingular then $|P| \neq 0$ ■

3.4 Finding A^{-1} Using Cofactor Expansions

In Section 3.2 we discussed the row reduction method for computing the determinant of a matrix. This method is well suited for computer evaluation of determinants because it is systematic and easily programmed. In this section we introduce a method for evaluating determinants that is useful for hand computations and is important theoretically. Namely, we will obtain a formula for the inverse of an invertible matrix as well as a formula for the solution of square systems of linear equations.
Let $A = (a_{ij})$ be an $n \times n$ matrix. Let M_{ij} be the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the ith row and the jth column of A. The determinant of M_{ij} is called the **minor of the entry** a_{ij}. The number $C_{ij} = (-1)^{i+j}|M_{ij}|$ is called the **cofactor of the entry** a_{ij}.

Exercise 176
Let

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

Find the minor and the cofactor of the entry $a_{32} = 4$.

Solution.
The minor of the entry a_{32} is

$$|M_{32}| = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

and the cofactor is $C_{32} = (-1)^{3+2}|M_{32}| = -26$ ■

Note that the cofactor and the minor of an entry differ only in sign. A quick way for determining whether to use the + or − is to use the following checkboard array

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

In order to prove the main result of this section, we first prove the following

Theorem 34
Let $A = (a_{ij}), A' = (a'_{ij})$, *and* $A'' = (a''_{ij})$ *be* $n \times n$ *matrices such that* $a_{ij} = a'_{ij} = a''_{ij}$ *if* $i \neq r$ *and* $a''_{rj} = a_{rj} + a'_{rj}$. *Then*

$$|A''| = |A| + |A'|.$$

The same result holds for columns.

Proof.
Using the definition of determinant we have

$$\begin{array}{rcl} |A''| & = & \sum_{\sigma \in S_n} sgn(\sigma) a''_{1\sigma(1)} a''_{2\sigma(2)} \cdots a''_{r\sigma(r)} \cdots a''_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a''_{1\sigma(1)} a''_{2\sigma(2)} \cdots (a_{r\sigma(r)} + a'_{r\sigma(r)}) \cdots a''_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a''_{1\sigma(1)} a''_{2\sigma(2)} \cdots a_{r\sigma(r)} \cdots a''_{n\sigma(n)} \\ & + & \sum_{\sigma \in S_n} sgn(\sigma) a''_{1\sigma(1)} a''_{2\sigma(2)} \cdots a'_{r\sigma(r)} \cdots a''_{n\sigma(n)} \\ & = & \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} \cdots a_{n\sigma(n)} \\ & + & \sum_{\sigma \in S_n} sgn(\sigma) a'_{1\sigma(1)} a'_{2\sigma(2)} \cdots a'_{r\sigma(r)} \cdots a'_{n\sigma(n)} \\ & = & |A| + |A'|. \end{array}$$

Since $|A^T| = |A|$ the same result holds for columns. ■

Exercise 177
Prove the following identity without evaluating the determinants .

$$\begin{vmatrix} a_1 & b_1 & a_1+b_1+c_1 \\ a_2 & b_2 & a_2+b_2+c_2 \\ a_3 & b_3 & a_3+b_3+c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Solution.
By the previous theorem and Exercise 170 we have

$$\begin{aligned} \begin{vmatrix} a_1 & b_1 & a_1+b_1+c_1 \\ a_2 & b_2 & a_2+b_2+c_2 \\ a_3 & b_3 & a_3+b_3+c_3 \end{vmatrix} &= \begin{vmatrix} a_1 & b_1 & a_1 \\ a_2 & b_2 & a_2 \\ a_3 & b_3 & a_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & b_1 \\ a_2 & b_2 & b_2 \\ a_3 & b_3 & b_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \\ &= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \blacksquare \end{aligned}$$

The following result is known as the **Laplace expansion** of $|A|$ along a row (respectively a column).

Theorem 35
If a_{ij} denotes the ijth entry of A and C_{ij} is the cofactor of the entry a_{ij} then the expansion along row i is

$$|A| = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

The expansion along column j is given by

$$|A| = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

Proof.
We prove the first formula. The second formula follows from the first because of the fact that $|A^T| = |A|$. Let $r_1, r_2, \cdots, r_n$ denote the rows of A. By Theorem 34 we have

$$|A| = \begin{vmatrix} r_1 \\ r_2 \\ \vdots \\ \sum_{j=1}^n a_{ij}e_j \\ \vdots \\ r_n \end{vmatrix} = \sum_{j=1}^n a_{ij} \begin{vmatrix} r_1 \\ r_2 \\ \vdots \\ e_j \\ \vdots \\ r_n \end{vmatrix}$$

where e_j is the $1 \times n$ matrix with 1 at the jth position and zero elsewhere. We will show that

$$\begin{vmatrix} r_1 \\ r_2 \\ \vdots \\ e_j \\ \vdots \\ r_n \end{vmatrix}$$

is just the cofactor of the entry a_{ij}.
By Theorem 27 (d) we have

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}$$

Now, move the ith row to the first row using $(i-1)$ successive row interchanges. Then, move the jth column to the first using $(j-1)$ successive column interchanges. Thus, we have

$$\begin{vmatrix} r_1 \\ r_2 \\ \vdots \\ e_j \\ \vdots \\ r_n \end{vmatrix} = (-1)^{i+j} \begin{vmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & M_{ij} \end{vmatrix}$$

where M_{ij} is the $(n-1)\times(n-1)$ matrix obtained from A by deleting the ith row and the jth column of A.
Let $E_1, E_2, \cdots, E_k$ be $(n-1)\times(n-1)$ elementary matrices such that $E_kE_{k-1}\cdots E_1M_{ij} = C_{ij}$ is an upper triangular matrix. Then the matrices $E'_1, E'_2, \cdots, E'_k$ are $n \times n$ elementary matrices such that

$$\left| E'_kE'_{k-1}\cdots E'_1 \begin{pmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & M_{ij} \end{pmatrix} \right| = \begin{vmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & C_{ij} \end{vmatrix}$$

where

$$E'_i = \begin{pmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & E_i \end{pmatrix}$$

Therefore,

$$\begin{vmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & M_{ij} \end{vmatrix} = \frac{|C_{ij}|}{|E_k||E_{k-1}\cdots|E_1|} = |M_{ij}|.$$

Hence,

$$\begin{vmatrix} r_1 \\ r_2 \\ \vdots \\ e_j \\ \vdots \\ r_n \end{vmatrix} = (-1)^{i+j}|M_{ij}| = C_{ij}$$

This ends a proof of the theorem ■

Remark
In general, the best strategy for evaluating a determinant by cofactor expansion is to expand along a row or a column having the largest number of zeros.

Exercise 178
Find the determinant of each of the following matrices.
(a)

$$A = \left(\begin{array}{ccc} 0 & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right)$$

(b)

$$A = \left(\begin{array}{cccc} 0 & 0 & 0 & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right)$$

Solution.
(a) Expanding along the first column we find

$$|A| = -a_{31}a_{22}a_{13}.$$

(b) Again, by expanding along the first column we obtain

$$|A| = a_{41}a_{32}a_{23}a_{34} \blacksquare$$

Exercise 179
Use cofactor expansion along the first column to find $|A|$ *where*

$$A = \left(\begin{array}{cccc} 3 & 5 & -2 & 6 \\ 1 & 2 & -1 & 1 \\ 2 & 4 & 1 & 5 \\ 3 & 7 & 5 & 3 \end{array}\right)$$

Solution.
Expanding along the first column we find

$$\begin{array}{rcr} |A| & = & 3C_{11} + C_{21} + 2C_{31} + 3C_{41} \\ & = & 3|M_{11}| - |M_{21}| + 2|M_{31}| - 3|M_{41}| \\ & = & 3(-54) + 78 + 2(60) - 3(18) = -18 \blacksquare \end{array}$$

If A is an $n \times n$ square matrix and C_{ij} is the cofactor of the entry a_{ij} then the transpose of the matrix

$$\left(\begin{array}{cccc} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{array}\right)$$

is called the **adjoint** of A and is denoted by $adj(A)$.

Exercise 180
Let

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

Find $adj(A)$.

Solution.
We first find the matrix of cofactors of A.

$$\begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix} = \begin{pmatrix} 12 & 6 & -16 \\ 4 & 2 & 16 \\ 12 & -10 & 16 \end{pmatrix}$$

The adjoint of A is the transpose of this cofactor matrix.

$$adj(A) = \begin{pmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{pmatrix} \blacksquare$$

Our next goal is to find another method for finding the inverse of a nonsingular square matrix. To this end, we need the following result.

Theorem 36
For $i \neq j$ *we have*

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots + a_{in}C_{jn} = 0.$$

Proof.
Let B be the matrix obtained by replacing the jth row of A by the ith row of A. Then B has two identical rows and therefore $|B| = 0$(See Theorem 27 (c)). Expand $|B|$ along the jth row. The elements of the jth row of B are $a_{i1}, a_{i2}, \ldots, a_{in}$. The cofactors are $C_{j1}, C_{j2}, \ldots, C_{jn}$. Thus

$$0 = |B| = a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots + a_{in}C_{jn}$$

This concludes a proof of the theorem $\blacksquare$

The following theorem states that the product $A \cdot adj(A)$ is a scalar matrix.

Theorem 37 *If* A *is an* $n \times n$ *matrix then* $A \cdot adj(A) = |A|I_n$.

Proof.
The (i, j) entry of the matrix

$$A.adj(A) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} C_{11} & C_{21} & \ldots & C_{n1} \\ C_{12} & C_{22} & \ldots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \ldots & C_{nn} \end{pmatrix}$$

is given by the sum

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots + a_{in}C_{jn} = |A|$$

if $i = j$ and 0 if $i \neq j$. Hence,

$$A.adj(A) = \begin{pmatrix} |A| & 0 & \ldots & 0 \\ 0 & |A| & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & |A| \end{pmatrix} = |A|I_n.$$

This ends a proof of the theorem ■

The following theorem provides a way for finding the inverse of a matrix using the notion of the adjoint.

Theorem 38
If $|A| \neq 0$ *then* A *is invertible and* $A^{-1} = \frac{adj(A)}{|A|}$. *Hence,* $adj(A) = A^{-1}|A|$.

Proof.
By Theorem 37 we have that $A(adj(A)) = |A|I_n$. If $|A| \neq 0$ then $A(\frac{adj(A)}{|A|}) = I_n$. By Theorem 20, A is invertible with inverse $A^{-1} = \frac{adj(A)}{|A|}$. ■

Exercise 181
Let

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

Use Theorem 38 to find A^{-1}.

Solution.
First we find the determinant of A given by $|A| = 64$. By Theorem 38

$$A^{-1} = \frac{1}{|A|}adj(A) = \begin{pmatrix} \frac{3}{16} & \frac{1}{16} & \frac{3}{16} \\ \frac{3}{32} & \frac{1}{32} & -\frac{5}{32} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} ■$$

In the next theorem we discuss three properties of the adjoint matrix.

Theorem 39
Let A *and* B *denote invertible* $n \times n$ *matrices. Then,*

(a) $adj(A^{-1}) = (adj(A))^{-1}$.
(b) $adj(A^T) = (adj(A))^T$.
(c) $adj(AB) = adj(B)adj(A)$.

Proof.
(a) Since $A(adj(A)) = |A|I_n$ then $adj(A)$ is invertible and $(adj(A))^{-1} = \frac{A}{|A|} = (A^{-1})^{-1}|A^{-1}| = adj(A^{-1})$.

(b) $adj(A^T) = (A^T)^{-1}|A^T| = (A^{-1})^T|A| = (adj(A))^T$.

(c) We have $adj(AB) = (AB)^{-1}|AB| = B^{-1}A^{-1}|A||B| = (B^{-1}|B|)(A^{-1}|A|) = adj(B)adj(A)$. ■

Exercise 182
Show that if A is singular then $A.adj(A) = \mathbf{0}$.

Solution.
If A is singular then $|A| = 0$. But then $A \cdot adj(A) = |A|I_n = \mathbf{0}$ ■

3.5 Cramer's Rule

Cramer's rule is another method for solving a linear system of n equations in n unknowns. This method is reasonable for inverting, for example, a 3×3 matrix by hand; however, the inversion method discussed before is more efficient for larger matrices.

Theorem 40
Let $Ax = b$ be a matrix equation with $A = (a_{ij}), x = (x_i), b = (b_i)$. Then we have the following matrix equation

$$\begin{pmatrix} |A|x_1 \\ |A|x_2 \\ \vdots \\ |A|x_n \end{pmatrix} = \begin{pmatrix} |A_1| \\ |A_2| \\ \vdots \\ |A_n| \end{pmatrix}$$

where A_i is the matrix obtained from A by replacing its ith column by b. It follows that

(1) If $|A| \neq 0$ then the above system has a unique solution given by

$$x_i = \frac{|A_i|}{|A|},$$

where $1 \leq i \leq n$.
(2) If $|A| = 0$ and $|A_i| \neq 0$ for some i then the system has no solution.
(3) If $|A| = |A_1| = \ldots = |A_n| = 0$ then the system has an infinite number of solutions.

Proof.
We have the following chain of equalities

$$|A|x = |A|(I_n x)$$

$$\begin{aligned} &= (|A|I_n)x \\ &= adj(A)Ax \\ &= adj(A)b \end{aligned}$$

The ith entry of the vector $|A|x$ is given by

$$|A|x_i = b_1C_{1i} + b_2C_{2i} + \ldots + b_nC_{ni}.$$

On the other hand by expanding $|A_i|$ along the ith column we find that

$$|A_i| = C_{1i}b_1 + C_{2i}b_2 + \ldots + C_{ni}b_n.$$

Hence

$$|A|x_i = |A_i|.$$

Now, (1), (2), and (3) follow easily. This ends a proof of the theorem ■

Exercise 183
Use Cramer's rule to solve

$$\left\{\begin{array}{rcrcrcr} -2x_1 & + & 3x_2 & - & x_3 & = & 1 \\ x_1 & + & 2x_2 & - & x_3 & = & 4 \\ -2x_1 & - & x_2 & + & x_3 & = & -3. \end{array}\right.$$

Solution.
By Cramer's rule we have

$$A = \left(\begin{array}{rrr} -2 & 3 & -1 \\ 1 & 2 & -1 \\ -2 & -1 & 1 \end{array}\right), |A| = -2.$$

$$A_1 = \left(\begin{array}{rrr} 1 & 3 & -1 \\ 4 & 2 & -1 \\ -3 & -1 & 1 \end{array}\right), |A_1| = -4.$$

$$A_2 = \left(\begin{array}{rrr} -2 & 1 & -1 \\ 1 & 4 & -1 \\ -2 & -3 & 1 \end{array}\right), |A_2| = -6.$$

$$A_3 = \left(\begin{array}{rrr} -2 & 3 & 1 \\ 1 & 2 & 4 \\ -2 & -1 & -3 \end{array}\right), |A_3| = -8.$$

Thus, $x_1 = \frac{|A_1|}{|A|} = 2, x_2 = \frac{|A_2|}{|A|} = 3, x_3 = \frac{|A_3|}{|A|} = 4$ ■

Exercise 184
Use Cramer's rule to solve

$$\left\{\begin{array}{rcrcrcr} 5x_1 & - & 3x_2 & - & 10x_3 & = & -9 \\ 2x_1 & + & 2x_2 & - & 3x_3 & = & 4 \\ -3x_1 & - & x_2 & + & 5x_3 & = & 1. \end{array}\right.$$

Solution.
By Cramer's rule we have

$$A = \left(\begin{array}{ccc} 5 & -3 & -10 \\ 2 & 2 & -3 \\ -3 & -1 & 5 \end{array}\right), |A| = -2.$$

$$A_1 = \left(\begin{array}{ccc} -9 & -3 & -10 \\ 4 & 2 & -3 \\ 1 & -1 & 5 \end{array}\right), |A_1| = 66.$$

$$A_2 = \left(\begin{array}{ccc} 5 & -9 & -10 \\ 2 & 4 & -3 \\ -3 & 1 & 5 \end{array}\right), |A_2| = -16.$$

$$A_3 = \left(\begin{array}{ccc} 5 & -3 & -9 \\ 2 & 2 & 4 \\ -3 & -1 & 1 \end{array}\right), |A_3| = 36.$$

Thus, $x_1 = \frac{|A_1|}{|A|} = -33, x_2 = \frac{|A_2|}{|A|} = 8, x_3 = \frac{|A_3|}{|A|} = -18$ ■

Exercise 185
Let ABC be a triangle such that $dist(A,B) = c, dist(A,C) = b, dist(B,C) = a$, the angle between AB and AC is α, between BA and BC is β, and that between CA and CB is γ.
(a) Use trigonometry to show that

$$\left\{\begin{array}{ccccc} b\cos\gamma & + & c\cos\beta & = & a \\ c\cos\alpha & + & a\cos\gamma & = & b \\ a\cos\beta & + & b\cos\alpha & = & c \end{array}\right.$$

(b) Use Cramer's rule to express $\cos\alpha, \cos\beta$, and $\cos\gamma$ in terms of a, b, and c.

Solution.
(a) Let A_1 be the leg of the perpendicular to BC through A. Then the triangles A_1AB and A_1AC are right triangles. In this case, we have $\cos\beta = \frac{dist(B,A_1)}{c}$ or $dist(B,A_1) = c\cos\beta$. Simlarly, $dist(C,A_1) = b\cos\gamma$. But $a = dist(B,A_1) + dist(C,A_1) = b\cos\gamma + c\cos\beta$. Similar argument for the remaining two equalitites.
(b)

$$A = \left(\begin{array}{ccc} 0 & c & b \\ c & 0 & a \\ b & a & 0 \end{array}\right), |A| = 2abc.$$

$$A_1 = \left(\begin{array}{ccc} a & c & b \\ b & 0 & a \\ c & a & 0 \end{array}\right), |A_1| = a(c^2 + b^2) - a^3.$$

$$A_2 = \begin{pmatrix} 0 & a & b \\ c & b & a \\ b & c & 0 \end{pmatrix}, |A_2| = b(a^2 + c^2) - b^3.$$

$$A_3 = \begin{pmatrix} 0 & c & a \\ c & 0 & b \\ b & a & c \end{pmatrix}, |A_3| = c(a^2 + b^2) - c^3.$$

Thus, $\cos\alpha = \frac{|A_1|}{|A|} = \frac{c^2+b^2-a^2}{2bc}, \cos\beta = \frac{|A_2|}{|A|} = \frac{a^2+c^2-b^2}{2ac}, \cos\gamma = \frac{|A_3|}{|A|} = \frac{a^2+b^2-c^2}{2ab}$
■

3.6 Review Problems

Exercise 186
(a) Find the number of inversions in the permutation (41352).
(b) Is this permutation even or odd?

Exercise 187
Evaluate the determinant of each of the following matrices
(a)

$$A = \begin{pmatrix} 3 & 5 \\ -2 & 4 \end{pmatrix}$$

(b)

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

Exercise 188
Find all values of t for which the determinant of the following matrix is zero.

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

Exercise 189
Solve for x

$$\begin{vmatrix} x & -1 \\ 3 & 1-x \end{vmatrix} = \begin{vmatrix} 1 & 0 & -3 \\ 2 & x & -6 \\ 1 & 3 & x-5 \end{vmatrix}$$

Exercise 190
Evaluate the determinant of the following matrix

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

Exercise 191
Evaluate the determinant of the following matrix.

$$\left|\begin{array}{ccccc} 2 & 7 & -3 & 8 & 3 \\ 0 & -3 & 7 & 5 & 1 \\ 0 & 0 & 6 & 7 & 6 \\ 0 & 0 & 0 & 9 & 8 \\ 0 & 0 & 0 & 0 & 4 \end{array}\right|$$

Exercise 192
Use the row reduction technique to find the determinant of the following matrix.

$$A = \left(\begin{array}{cccc} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{array}\right)$$

Exercise 193
Given that

$$\left|\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array}\right| = -6,$$

find
(a)

$$\left|\begin{array}{ccc} d & e & f \\ g & h & i \\ a & b & c \end{array}\right|,$$

(b)

$$\left|\begin{array}{ccc} 3a & 3b & 3c \\ -d & -e & -f \\ 4g & 4h & 4i \end{array}\right|$$

(c)

$$\left|\begin{array}{ccc} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{array}\right|$$

(d)

$$\left|\begin{array}{ccc} -3a & -3b & -3c \\ d & e & f \\ g-4d & h-4e & i-4f \end{array}\right|$$

Exercise 194
Determine by inspection the determinant of the following matrix.

$$\left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 2 & 4 & 6 & 8 & 10 \end{array}\right)$$

Exercise 195
Find the determinant of the 1×1 *matrix* $A = (3)$.

Exercise 196
Let A *be a* 3×3 *matrix such that* $|2A| = 6$. *Find* $|A|$.

Exercise 197
Show that if n *is any positive integer then* $|A^n| = |A|^n$.

Exercise 198
Show that if A *is an* $n \times n$ *skew-symmetric and* n *is odd then* $|A| = 0$.

Exercise 199
Show that if A *is* **orthogonal**, *i.e.* $A^TA = AA^T = I_n$ *then* $|A| = \pm 1$. *Note that* $A^{-1} = A^T$.

Exercise 200
If A *is a nonsingular matrix such that* $A^2 = A$, *what is* $|A|$?

Exercise 201
True or false: If

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

then $rank(A) = 3$. *Justify your answer.*

Exercise 202
Find out, without solving the system, whether the following system has a non-trivial solution

$$\left\{ \begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 0 \\ 3x_1 & + & x_2 & + & 2x_3 & = & 0 \end{array} \right.$$

Exercise 203
For which values of c *does the matrix*

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

have an inverse.

Exercise 204
If $|A| = 2$ *and* $|B| = 5$, *calculate* $|A^3B^{-1}A^TB^2|$.

Exercise 205
Show that $|AB| = |BA|$.

Exercise 206
Show that $|A + B^T| = |A^T + B|$ *for any* $n \times n$ *matrices* A *and* B.

Exercise 207
Let $A = (a_{ij})$ *be a triangular matrix. Show that* $|A| \neq 0$ *if and only if* $a_{ii} \neq 0$, *for* $1 \leq i \leq n$.

Exercise 208
Express

$$\left| \begin{array}{cc} a_1 + b_1 & c_1 + d_1 \\ a_2 + b_2 & c_2 + d_2 \end{array} \right|$$

as a sum of four determinants whose entries contain no sums.

Exercise 209
Let

$$A = \left(\begin{array}{ccc} 3 & -2 & 1 \\ 5 & 6 & 2 \\ 1 & 0 & -3 \end{array} \right)$$

(a) Find adj(A).
(b) Compute $|A|$.

Exercise 210
Find the determinant of the matrix

$$A = \left(\begin{array}{cccc} 3 & 0 & 0 & 0 \\ 5 & 1 & 2 & 0 \\ 2 & 6 & 0 & -1 \\ -6 & 3 & 1 & 0 \end{array} \right)$$

Exercise 211
Find the determinant of the following **Vandermonde** *matrix.*

$$A = \left(\begin{array}{ccc} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{array} \right)$$

Exercise 212
Let A *be an* $n \times n$ *matrix. Show that* $|adj(A)| = |A|^{n-1}$.

Exercise 213
If

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

find $adj(A)$.

Exercise 214
If $|A| = 2$, *find* $|A^{-1} + adj(A)|$.

Exercise 215
Show that $adj(\alpha A) = \alpha^{n-1} adj(A)$, *where* n *is a positive integer.*

Exercise 216
Consider the matrix

$$A = \left(\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 5 & 7 \end{array}\right)$$

(a) Find $|A|$.
(b) Find $adj(A)$.
(c) Find A^{-1}.

Exercise 217
Prove that if A *is symmetric then* $adj(A)$ *is also symmetric.*

Exercise 218
Prove that if A *is a nonsingular triangular matrix then* A^{-1} *is also triangular.*

Exercise 219
Let A *be an* $n \times n$ *matrix.*
(a) Show that if A *has integer entries and* $|A| = 1$ *then* A^{-1} *has integer entries as well.*
(b) Let $Ax = b$. *Show that if the entries of* A *and* b *are integers and* $|A| = 1$ *then the entries of* x *are also integers.*

Exercise 220
Show that if $A^k = \mathbf{0}$ *for some positive integer* k *then* A *is singular.*

Exercise 221
Use Cramer's Rule to solve

$$\left\{\begin{array}{rcrcrcr} x_1 & & & + & 2x_3 & = & 6 \\ -3x_1 & + & 4x_2 & + & 6x_3 & = & 30 \\ -x_1 & - & 2x_2 & + & 3x_3 & = & 8 \end{array}\right.$$

Exercise 222
Use Cramer's Rule to solve

$$\left\{\begin{array}{rcrcrcr} 5x_1 & + & x_2 & - & x_3 & = & 4 \\ 9x_1 & + & x_2 & - & x_3 & = & 1 \\ x_1 & - & x_2 & + & 5x_3 & = & 2 \end{array}\right.$$

Chapter 4

The Theory of Vector Spaces

In Chapter 2, we saw that the operations of addition and scalar multiplication on the set M_{mn} of $m \times n$ matrices possess many of the same algebraic properties as addition and scalar multiplication on the set $\mathbb{R}$ of real numbers. In fact, there are many other sets with operations that share these same properties. Instead of studying these sets individually, we study them as a class.
In this chapter, we define vector spaces to be sets with algebraic operations having the properties similar to those of addition and scalar multiplication on $\mathbb{R}$ and M_{mn}. We then establish many important results that apply to all vector spaces, not just $\mathbb{R}$ and M_{mn}.

4.1 Vectors in Two and Three Dimensional Spaces

Many Physical quantities are represented by vectors. For examples, the velocity of a moving object, displacement, force of gravitation, etc. In this section we discuss the geometry of vectors in two and three dimensional spaces, discuss the arithmetic operations of vectors and study some of the properties of these operations. As we shall see later in this chapter, the present section will provide us with an example of a vector space, a concept that will be defined in the next section, that can be illustrated geometrically.
A **vector** in space is a directed segment with a tail, called the **initial point**, and a tip, called the **terminal point**. A vector will be denoted by $\vec{v}$ and scalars by greek letters $\alpha, \beta, \cdots$ Finally, we point out that most of the results of this section hold for vectors in $\mathbb{R}^2$.

Two vectors $\vec{v}$ and $\vec{w}$ are said to be **equivalent** if they have the same length and the same direction. We write $\vec{v} = \vec{w}$.

Since we can always draw a vector with initial point O that is equivalent to a given vector then from now on we assume that the vectors have all the same initial point O unless indicated otherwise. The coordinates (x, y, z) of the terminal point of a vector $\vec{v}$ are called the **components** of $\vec{v}$ and we write

$$\vec{v} = (x, y, z).$$

In terms of components, two vectors $\vec{v} = (x, y, z)$ and $\vec{w} = (x', y', z')$ are equivalent if and only if $x = x', y = y'$, and $z = z'$.
Given two vectors $\vec{v} = (x_1, y_1, z_1)$ and $\vec{w} = (x_2, y_2, z_2)$, we define the **sum** $\vec{v} + \vec{w}$ to be the vector

$$\vec{v} + \vec{w} = (x_1 + x_2, y_1 + y_2, z_1 + z_2).$$

Geometrically, the $v \vec{+} w$ is the diagonal of the parallelogram with sides the vectors $\vec{v}$ and $\vec{w}$.
We define the **negative** vector $-\vec{v}$ to be the vector

$$-\vec{v} = (-x_1, -y_1, -z_1)$$

and we define

$$\vec{v} - \vec{w} = \vec{v} + (-\vec{w}) = (x_1 - x_2, y_1 - y_2, z_1 - z_2).$$

Assuming that $\vec{v}$ and $\vec{w}$ have the same initial point, the vector $\vec{v} - \vec{w}$ is the vector with initial point the terminal point of $\vec{w}$ and its terminal point is the terminal point of $\vec{v}$.
If α is a scalar we define $\alpha\vec{v}$ to be the vector

$$\alpha\vec{v} = (\alpha x_1, \alpha y_1, \alpha z_1).$$

With the above definitions, every vector $\vec{u} = (x, y, z)$ can be expressed as a combination of the three vectors $\vec{i} = (1, 0, 0), \vec{j} = (0, 1, 0)$, and $\vec{k} = (0, 0, 1)$. That is

$$\vec{u} = x\vec{i} + y\vec{j} + z\vec{k}.$$

Exercise 223
Given two points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$, find the components of the vector $\overrightarrow{P_1P_2}$.

Solution.
The vector $\overrightarrow{P_1P_2}$ is the difference of the vectors $\overrightarrow{OP_1}$ and $\overrightarrow{OP_2}$ so
$\overrightarrow{P_1P_2} = \overrightarrow{OP_2} - \overrightarrow{OP_1} = (x_2, y_2, z_2) - (x_1, y_1, z_1) = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ ■

Exercise 224
Let $\vec{u} = (1, 2, 3), \vec{v} = (2, -3, 1)$, and $\vec{w} = (3, 2, -1)$.

(a) Find the components of the vector $\vec{u} - 3\vec{v} + 8\vec{w}$.
(b) Find scalars c_1, c_2, c_3 such that

$$c_1\vec{u} + c_2\vec{v} + c_3\vec{w} = (6, 14, -2).$$

Solution.
(a) $\vec{u} - 3\vec{v} + 8\vec{w} = (1,2,3) - 3(2,-3,1) + 8(3,2,-1) = (1,2,3) - (6,-9,3) + (24,16,-8) = (19,27,-8)$.
(b) We have $(c_1, 2c_1, 3c_1) + (2c_2, -3c_2, c_2) + (3c_3, 2c_3, -c_3) = (c_1 + 2c_2 + 3c_3, 2c_1 - 3c_2 + 2c_3, 3c_1 + c_2 - c_3) = (6, 14, -2)$. This yields the linear system

$$\left\{\begin{array}{rcrcrcr} c_1 & + & 2c_2 & + & 3c_3 & = & 6 \\ 2c_1 & - & 3c_2 & + & 2c_3 & = & 14 \\ 3c_1 & + & c_2 & - & c_3 & = & -2 \end{array}\right.$$

The augmented matrix of the system is

$$\left(\begin{array}{rrrr} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

The reduction of this matrix to row-echelon form is

Step 1: $r_2 \leftarrow r_2 - 2r_1$ and $r_3 \leftarrow r_3 - 3r_1$

$$\left(\begin{array}{rrrr} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{array}\right)$$

Step 2: $r_2 \leftarrow 2r_2 - 3r_3$

$$\left(\begin{array}{rrrr} 1 & 2 & 3 & 6 \\ 0 & 1 & 22 & 64 \\ 0 & -5 & -10 & -20 \end{array}\right)$$

Step 3: $r_3 \leftarrow r_3 + 5r_2$

$$\left(\begin{array}{rrrr} 1 & 2 & 3 & 6 \\ 0 & 1 & 22 & 64 \\ 0 & 0 & 100 & 300 \end{array}\right)$$

Step 4: $r_3 \leftarrow \frac{1}{100} r_3$

$$\left(\begin{array}{rrrr} 1 & 2 & 3 & 6 \\ 0 & 1 & 22 & 64 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

The corresponding system is

$$\left\{\begin{array}{rcrcrcr} c_1 & + & 2c_2 & + & 3c_3 & = & 6 \\ & & c_2 & + & 22c_3 & = & 64 \\ & & & & c_3 & = & 3 \end{array}\right.$$

Solving the above triangular system to obtain: $c_1 = 1, c_2 = -2, c_3 = 3$ ■

The basic properties of vector addition and scalar multiplication are collected in the following theorem.

Theorem 41
If $\vec{u}, \vec{v}$, and $\vec{w}$ are vectors and α, β are scalars then the following properties hold.

(a) $\vec{u}+\vec{v}=\vec{v}+\vec{u}$.
(b) $(\vec{u}+\vec{v})+\vec{w}=\vec{u}+(\vec{v}+\vec{w})$.
(c) $\vec{u}+\vec{0}=\vec{u}$, *where* $\vec{0}=(0,0,0)$.
(d) $\vec{u}+(-\vec{u})=\vec{0}$.
(e) $\alpha(\beta\vec{u})=(\alpha\beta)\vec{u}$.
(f) $\alpha(\vec{u}+\vec{v})=\alpha\vec{u}+\alpha\vec{v}$.
(g) $(\alpha+\beta)\vec{u}=\alpha\vec{u}+\beta\vec{u}$.
(h) $1\vec{u}=\vec{u}$.

Proof.
The prove of this theorem uses the properties of addition and scalar multiplication of real numbers. We prove (a). The remaining properties can be proved easily. Suppose that $\vec{u}=(x_1,y_1,z_1)$ and $\vec{v}=(x_2,y_2,z_2)$. Then using the fact that the addition of real numbers is commutative we have

$$\begin{aligned}\vec{u}+\vec{v} &= (x_1,y_1,z_1)+(x_2,y_2,z_2)\\ &= (x_1+x_2,y_1+y_2,z_1+z_2)\\ &= (x_2+x_1,y_2+y_1,z_2+z_1)\\ &= (x_2,y_2,z_2)+(x_1,y_1,z_1)\\ &= \vec{v}+\vec{u}\end{aligned}$$

This ends the proof of (a) ■

Exercise 225
(a) Let $\vec{u}=(x_1,y_1,z_1)$. Show that the length of $\vec{u}$, known as the **norm** *of $\vec{u}$, is given by the expression*

$$||\vec{u}||=\sqrt{x_1^2+y_1^2+z_1^2}.$$

(b) Given two points $P_1(x_1,y_1,z_1)$ and $P_2(x_2,y_2,z_2)$ show that the distance between these points is given by the formula

$$||\overrightarrow{P_1P_2}||=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}.$$

Solution.
(a) Let $\vec{u}=\overrightarrow{OP}$. Let Q be the orthogonal projection of P onto the xy-plane. Then by the Pythagorean Theorem we have

$$||\overrightarrow{OQ}||^2+||\overrightarrow{PQ}||^2=||\overrightarrow{OP}||^2$$

But $||\overrightarrow{OQ}||^2=x_1^2+y_1^2$ and $||\overrightarrow{PQ}||^2=z_1^2$. Thus,

$$||\vec{u}||^2=x_1^2+y_1^2+z_1^2$$

Now take the square root of both sides.
(b) By Exercise 223, $\overrightarrow{P_1P_2}=(x_2-x_1,y_2-y_1,z_2-z_1)$. Hence, by part (a) we have

$$||\overrightarrow{P_1P_2}||=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}\blacksquare$$

Exercise 226
Show that if $\vec{u}$ is a non zero vector then the length of the vector $\frac{\vec{u}}{||\vec{u}||}$ is 1.

Solution.
If $\vec{u} = (x, y, z)$ then $\frac{\vec{u}}{||\vec{u}||} = (\frac{x}{||\vec{u}||}, \frac{y}{||\vec{u}||}, \frac{z}{||\vec{u}||})$. Hence,

$$||\frac{\vec{u}}{||\vec{u}||}|| = \sqrt{\frac{x^2}{||\vec{u}||^2} + \frac{y^2}{||\vec{u}||^2} + \frac{z^2}{||\vec{u}||^2}} = 1\blacksquare$$

Exercise 227
Suppose that an $xyz-$coordinate system is translated to obtain a new system $x'y'z'$ with origin the point $O'(k, l, m)$. Let P be a point in the space. Suppose that (x, y, z) are the coordinates of P in $xyz-$system and (x', y', z') are the coordinates of P in the $x'y'z'-$ system. Show that $x' = x - k, y' = y - l, z' = z - m$.

Solution.
By Exercise 223, $\overrightarrow{O'P} = (x - k, y - l, z - m) = (x', y', z')$. Thus, $x' = x - k, y' = y - l, z' = z - m$ $\blacksquare$

Next, we shall discuss a way for multiplying vectors. If $\vec{u}$ and $\vec{v}$ are two vectors in the space and θ is the angle between them we define the **inner product** (sometimes called **scalar product** or **dot product**) of $\vec{u}$ and $\vec{v}$ to be the number

$$< \vec{u}, \vec{v} >= ||\vec{u}||||\vec{v}|| \cos\theta \quad 0 \le \theta \le \pi.$$

The following theorem lists the most important properties of the dot product.

Theorem 42
Let $\vec{u} = (x_1, y_1, z_1)$, $\vec{v} = (x_2, y_2, z_2)$ and $\vec{w} = (x_3, y_3, z_3)$ be vectors and α be a scalar. Then
(a) $< \vec{u}, \vec{v} >= x_1x_2 + y_1y_2 + z_1z_2$.
(b) $< \vec{u}, \vec{u} >= ||\vec{u}||^2$.
(c) $< \vec{u}, \vec{v} >=< \vec{v}, \vec{u} >$.
(d) $< \vec{u}, \vec{v} + \vec{w} >=< \vec{u}, \vec{v} > + < \vec{u}, \vec{w} >$.
(e) $\alpha < \vec{u}, \vec{v} >=< \alpha\vec{u}, \vec{v} >=< \vec{u}, \alpha\vec{v} >$.
(f) $< \vec{u}, \vec{u} >= 0$ if and only if $\vec{u} = \vec{0}$.

Proof.
(a) Let $\overrightarrow{OP} = \vec{u} = (x_1, y_1, z_1)$ and $\overrightarrow{OQ} = \vec{v} = (x_2, y_2, z_2)$. Let θ be the angle between them. Then $\overrightarrow{PQ} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$. By the law of cosines we have

$$\begin{aligned} ||\overrightarrow{PQ}||^2 &= ||\overrightarrow{OP}||^2 + ||\overrightarrow{OQ}||^2 - 2||\overrightarrow{OP}||||\overrightarrow{OQ}|| \cos\theta \\ &= ||\vec{u}||^2 + ||\vec{v}||^2 - 2||\vec{u}||||\vec{v}|| \cos\theta. \end{aligned}$$

But $||\overrightarrow{PQ}|| = ||\vec{u} - \vec{v}||$. Hence,

$$\begin{aligned} <\vec{u}, \vec{v}> &= ||\vec{u}||||\vec{v}|| \cos\theta \\ &= \frac{1}{2}(||\vec{u}||^2 + ||\vec{v}||^2 - ||\vec{u} - \vec{v}||^2) \\ &= \frac{1}{2}(x_1^2 + y_1^2 + z_1^2 + x_2^2 + y_2^2 + z_2^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2) \\ &= x_1x_2 + y_1y_2 + z_1z_2. \end{aligned}$$

(b) $<\vec{u}, \vec{u}> = ||\vec{u}||^2 = x_1^2 + y_1^2 + z_1^2$.
(c) $<\vec{u}, \vec{v}> = x_1x_2 + y_1y_2 + z_1z_2 = x_2x_1 + y_2y_1 + z_2z_1 = <\vec{v}, \vec{u}>$.
(d) $<\vec{u}, \vec{v} + \vec{w}> = x_1(x_2 + x_3) + y_1(y_2 + y_3) + z_1(z_2 + z_3) = (x_1x_2 + y_1y_2 + z_1z_2) + (x_1x_3 + y_1y_3 + z_1z_3) = <\vec{u}, \vec{v}> + <\vec{u}, \vec{w}>$.
(e) $\alpha <\vec{u}, \vec{v}> = \alpha(x_1x_2 + y_1y_2 + z_1z_2) = (\alpha x_1)x_2 + (\alpha y_1)y_2 + (\alpha z_1)z_2 = <\alpha\vec{u}, \vec{v}>$.
(f) If $<\vec{u}, \vec{u}> = 0$ then $x_1^2 + y_1^2 + z_1^2 = 0$ and this implies that $x_1 = y_1 = z_1 = 0$. That is, $\vec{u} = \vec{0}$. The converse is straightforward. ■

Exercise 228
Consider the vectors $\vec{u} = (2, -1, 1)$ *and* $\vec{v} = (1, 1, 2)$. *Find* $<\vec{u}, \vec{v}>$ *and the angle* θ *between these two vectors.*

Solution.
Using the above theorem we find

$$<\vec{u}, \vec{v}> = (2)(1) + (-1)(1) + (1)(2) = 3.$$

Also, $||\vec{u}|| = ||\vec{v}|| = \sqrt{6}$. Hence,

$$\cos\theta = \frac{3}{6} = \frac{1}{2}$$

This implies that $\theta = \frac{\pi}{3}$ radians. ■

Exercise 229
Two vectors are said to be **orthogonal** *if* $<\vec{u}, \vec{v}> = 0$. *Show that the two vectors* $\vec{u} = (2, -1, 1)$ *and* $\vec{v} = (1, 1, -1)$ *are orthogonal.*

Solution.
Indeed, $<\vec{u}, \vec{v}> = (2)(1) + (-1)(1) + (1)(-1) = 0$ ■

Next, we discuss the decomposition of a vector into a sum of two vectors. To be more precise, let $\vec{u}$ and $\vec{w}$ be two vectors with the same initial point Q and with $\vec{w}$ being horizontal. From the tip of $\vec{u}$ drop a perpendicular to the line through $\vec{w}$, and construct the vector $\vec{u_1}$ from Q to the foot of this perpendicular. Next form the vector $\vec{u_2} = \vec{u} - \vec{u_1}$. Clearly, $\vec{u} = \vec{u_1} + \vec{u_2}$, where $\vec{u_1}$ is parallel to $\vec{w}$ and $\vec{u_2}$ is perpendicular to $\vec{w}$. We call $\vec{u_1}$ the **orthogonal projection of $\vec{u}$ on $\vec{w}$** and we denote it by $proj_{\vec{w}}\vec{u}$.
The following theorem gives formulas for calculating $\vec{u_1}$ and $\vec{u_2}$.

Theorem 43
For any two vectors $\vec{u}$ *and* $\vec{w} \neq 0$ *we have* $\vec{u_1} = \frac{<\vec{u},\vec{w}>}{||\vec{w}||^2}\vec{w}$ *and* $\vec{u_2} = \vec{u} - \frac{<\vec{u},\vec{w}>}{||\vec{w}||^2}\vec{w}$.

Proof.
Since $\vec{u_1}$ is parallel to $\vec{w}$ then there is a scalar t such that $\vec{u_1} = t\vec{w}$. Thus, $<\vec{u},\vec{w}>=< t\vec{w}+\vec{u_2},\vec{w}>= t||\vec{w}||^2+<\vec{u_2},\vec{w}>$. But $\vec{u_2}$ and $\vec{w}$ are orthogonal so that the second term on the above equality is zero. Hence, $t = \frac{<\vec{u},\vec{w}>}{||\vec{w}||^2}$. It follows that $\vec{u_1} = t\vec{w} = \frac{<\vec{u},\vec{w}>}{||\vec{w}||^2}\vec{w}$ ■

Exercise 230
Let $\vec{u} = (2,-1,3)$ *and* $\vec{w} = (4,-1,2)$. *Find* $\vec{u_1}$ *and* $\vec{u_2}$.

Solution.
We have: $<\vec{u},\vec{w}>= 15$ and $||\vec{w}||^2 = 21$. Thus, $\vec{u_1} = (\frac{20}{7}, -\frac{5}{7}, \frac{10}{7})$ and $\vec{u_2} = \vec{u} - \vec{u_1} = (-\frac{6}{7}, -\frac{2}{7}, \frac{11}{7})$ ■

Exercise 231
Given a vector $\vec{v} = (a,b,c)$ *in* $\mathbb{R}^3$. *The angles* α,β,γ *between* $\vec{v}$ *and the unit vectors* $\vec{i},\vec{j}$, *and* $\vec{k}$, *respectively, are called the* **direction angles** *of* $\vec{v}$ *and the numbers* $\cos\alpha, \cos\beta$, *and* $\cos\gamma$ *are called the* **direction cosines** *of* $\vec{v}$.
(a) Show that $\cos\alpha = \frac{a}{||\vec{v}||}$.
(b) Find $\cos\beta$ *and* $\cos\gamma$.
(c) Show that $\frac{\vec{v}}{||\vec{v}||} = (\cos\alpha, \cos\beta, \cos\gamma)$.
(d) Show that $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

Solution.
(a) $\cos\alpha = \frac{<\vec{v},\vec{i}>}{||\vec{v}||||\vec{i}||} = \frac{a}{||\vec{v}||}$.
(b) By repeating the arithmetic of part (a) one finds $\cos\beta = \frac{b}{||\vec{v}||}$ and $\cos\gamma = \frac{c}{||\vec{v}||}$.
(c) $\frac{\vec{v}}{||\vec{v}||} = (\frac{a}{||\vec{v}||}, \frac{b}{||\vec{v}||}, \frac{c}{||\vec{v}||}) = (\cos\alpha, \cos\beta, \cos\gamma)$.
(d) $\cos\alpha^2 + \cos\beta^2 + \cos\gamma^2 = ||\frac{\vec{v}}{||\vec{v}||}||^2 = 1$ ■

Exercise 232
An interesting application of determinants and vectors is the construction of a vector orthogonal to two given vectors.

(a) Given two vectors $\vec{u} = (x_1,y_1,z_1)$ *and* $\vec{v} = (x_2,y_2,z_2)$ *then we define the* **cross product** *of* $\vec{u}$ *and* $\vec{v}$ *to be the vector*

$$\vec{u} \times \vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}$$

Find the components of the vector $\vec{u} \times \vec{v}$.
(b) Find $\vec{u} \times \vec{v}$, *where* $\vec{u} = (1,2,-2), \vec{v} = (3,0,1)$.
(c) Show that $<\vec{u},\vec{u}\times\vec{v}>= 0$ *and* $<\vec{v},\vec{u}\times\vec{v}>= 0$. *Hence,* $\vec{u}\times\vec{v}$ *is orthogonal to both* $\vec{u}$ *and* $\vec{v}$

Solution.
(a) Evaluating the determinant defining $\vec{u}\times\vec{v}$ we find $\vec{u}\times\vec{v} = (y_1z_2 - y_2z_1, x_2z_1 - x_1z_2, x_1y_2 - x_2y_1)$.
(b) Substituting in (a) we find $\vec{u}\times\vec{v} = (2, -7, -6)$.
(c) Indeed, $<\vec{u}, \vec{u}\times\vec{v}> = x_1(y_1z_2 - y_2z_1) + y_1(x_2z_1 - x_1z_2) + z_1(x_1y_2 - x_2y_1) = 0$. Similarly, $<\vec{v}, \vec{u}\times\vec{v}> = 0$ ■

4.2 Vector Spaces, Subspaces, and Inner Product Spaces

The properties listed in Theorem 41 of the previous section generalizes to $\mathbb{R}^n$, the Euclidean space to be discussed below. Also, these properties hold for the collection M_{mn} of all $m \times n$ matrices with the operation of addition and scalar multiplication. In fact there are other sets with operations satisfying the conditions of Theorem 41. Thus, it make sense to study the group of sets with these properties. In this section, we define vector spaces to be sets with algebraic operations having the properties similar to those of addition and scalar multiplication on $\mathbb{R}^n$ and M_{mn}.

Let n be a positive integer. Let $\mathbb{R}^n$ be the collection of elements of the form $(x_1, x_2, \ldots, x_n)$, where the x_is are real numbers. Define the following operations on $\mathbb{R}^n$:

(a) Addition: $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$
(b) Multiplication of a vector by a scalar:

$$\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n).$$

The basic properties of addition and scalar multiplication of vectors in $\mathbb{R}^n$ are listed in the following theorem.

Theorem 44
The following properties hold, for u, v, w in $\mathbb{R}^n$ and α, β scalars:

(a) $u + v = v + u$
(b) $u + (v + w) = (u + v) + w$
(c) $u + 0 = 0 + u = u$ *where* $0 = (0, 0, \ldots, 0)$
(d) $u + (-u) = 0$
(e) $\alpha(u + v) = \alpha u + \alpha v$
(f) $(\alpha + \beta)u = \alpha u + \beta u$
(g) $\alpha(\beta u) = (\alpha\beta)u$
(h) $1u = u$.

Proof.
This is just a generalization of Theorem 41 ■

The set $\mathbb{R}^n$ with the above operations and properties is called the **Euclidean space.**

A **vector space** is a set V together with the following operations:
(i) Addition: If $u, v \in V$ then $u + v \in V$. We say that V is **closed under addition**.
(ii) Multiplication of an element by a scalar: If $\alpha \in \mathbb{R}$ and $u \in V$ then $\alpha u \in V$. That is, V is **closed under scalar multiplication.**
(iii) These operations satisfy the properties (a) - (h) of Theorem 44.

Exercise 233
Let $F(\mathbb{R})$ be the set of functions $f : \mathbb{R} \to \mathbb{R}$. Define the operations

$$(f + g)(x) = f(x) + g(x)$$

and

$$(\alpha f)(x) = \alpha f(x).$$

Show that $F(\mathbb{R})$ is a vector space under these operations.

Solution.
The proof is based on the properties of the vector space $\mathbb{R}$.
(a) $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$ where we have used the fact that the addition of real numbers is commutative.
(b) $[(f + g) + h](x) = (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = f(x) + (g + h)(x) = [f + (g + h)](x)$.
(c) Let $\mathbf{0}$ be the zero function. Then for any $f \in F(\mathbb{R})$ we have $(f + \ 0)(x) = f(x) + \mathbf{0}(x) = f(x) = (\mathbf{0} + f)(x)$.
(d) $[f + (-f)](x) = f(x) + (-f(x)) = f(x) - f(x) = 0 = \mathbf{0}(x)$.
(e) $[\alpha(f + g)](x) = \alpha(f + g)(x) = \alpha f(x) + \alpha g(x) = (\alpha f + \alpha g)(x)$.
(f) $[(\alpha + \beta)f](x) = (\alpha + \beta)f(x) = \alpha f(x) + \beta f(x) = (\alpha f + \beta f)(x)$.
(g) $[\alpha(\beta f)](x) = \alpha(\beta f)(x) = (\alpha\beta)f(x) = [(\alpha\beta)f](x)$
$(h)(1f)(x) = 1f(x) = f(x)$.
Thus, $F(\mathbb{R})$ is a vector space ■

Exercise 234
Let M_{mn} be the collection of all $m \times n$ matrices. Show that M_{mn} is a vector space using matrix addition and scalar multiplication.

Solution.
This follows from Theorem 8 ■

Exercise 235
Let $V = \{(x, y) : x \geq 0, y \geq 0\}$. Show that the set V fails to be a vector space under the standard operations on $\mathbb{R}^2$.

Solution.
For any $(x, y) \in V$ with $x, y > 0$ we have $-(x, y) \notin V$. Thus, V is not a vector space ■

The following theorem exhibits some properties which follow directly from the axioms of the definition of a vector space and therefore hold for every vector space.

Theorem 45
Let V be a vector space, u a vector in V and α is a scalar. Then the following properties hold:

(a) $0u = 0$.
(b) $\alpha 0 = 0$
(c) $(-1)u = -u$
(d) If $\alpha u = 0$ *then* $\alpha = 0$ *or* $u = 0$.

Proof.
(a) For any scalar $\alpha \in \mathbb{R}$ we have $0u = (\alpha + (-\alpha))u = \alpha u + (-\alpha)u = \alpha u + (-(\alpha u)) = 0$.
(b) Let $u \in V$. Then $\alpha 0 = \alpha(u + (-u)) = \alpha u + \alpha(-u) = \alpha u + (-(\alpha u)) = 0$.
(c) $u + (-u) = u + (-1)u = 0$. So that $-u = (-1)u$.
(d) Suppose $\alpha u = 0$. If $\alpha \neq 0$ then α^{-1} exists and $u = 1u = (\alpha^{-1}\alpha)u = \alpha^{-1}(\alpha u) = \alpha^{-1}0 = 0$. ■

Now, it is possible that a vector space in contained in a larger vector space. A subset W of a vector space V is called a **subspace** of V if the following two properties are satisfied:

(i) If u, v are in W then $u + v$ is also in W.
(ii) If α is a scalar and u is in W then αu is also in W.

Every vector space V has at least two subspaces:V itself and the subspace consisting of the zero vector of V. These are called the **trivial** subspaces of V.

Exercise 236
Show that a subspace of a vector space is itself a vector space.

Solution.
All the axioms of a vector space hold for the elements of a subspace ■

The following exercise provides a criterion for deciding whether a subset S of a vector space V is a subspace of V.

Exercise 237
Show that W is a subspace of V if and only if $\alpha u + v \in W$ for all $u, v \in W$ and $\alpha \in \mathbb{R}$.

Solution.
Suppose that W is a subspace of V. If $u, v \in W$ and $\alpha \in \mathbb{R}$ then $\alpha u \in W$ and therefore $\alpha u + v \in W$. Conversely, suppose that for all $u, v \in W$ and $\alpha \in \mathbb{R}$ we have $\alpha u + v \in \mathbb{R}$. In particular, if $\alpha = 1$ then $u + v \in W$. If $v = 0$ then $\alpha u + v = \alpha u \in W$. Hence, W is a subspace ■

Exercise 238
Let M_{22} be the collection of 2×2 matrices. Show that the set W of all 2×2 matrices having zeros on the main diagonal is a subspace of M_{22}.

Solution.
The set W is the set

$$W = \left\{ \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\}$$

Clearly, the 2×2 zero matrix belongs to W. Also,

$$\alpha \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} + \begin{pmatrix} 0 & a' \\ b' & 0 \end{pmatrix} = \begin{pmatrix} 0 & \alpha a + a' \\ \alpha b + b' & 0 \end{pmatrix} \in W$$

Thus, W is a subspace of M_{22} ■

Exercise 239
Let $D([a, b])$ be the collection of all differentiable functions on $[a, b]$. Show that $D([a, b])$ is a subspace of the vector space of all functions defined on $[a, b]$.

Solution.
We know from calculus that if f, g are differentiable functions on $[a, b]$ and $\alpha \in \mathbb{R}$ then $\alpha f + g$ is also differentiable on $[a, b]$. Hence, $D([a, b])$ is a subspace of $F([a, b])$ ■

Exercise 240
Let A be an $m \times n$ matrix. Show that the set $S = \{x \in \mathbb{R}^n : Ax = \mathbf{0}\}$ is a subspace of $\mathbb{R}^n$.

Solution.
See Exercise 67 (a) ■

Vector spaces are important in defining spaces with some kind of geometry. For example, the Euclidean space $\mathbb{R}^n$. Such spaces are called inner product spaces, a notion that we discuss next.
Recall that if $\vec{u} = (x_1, y_1, z_1)$ and $\vec{v} = (x_2, y_2, z_2)$ are two vectors in $\mathbb{R}^3$ then the inner product of $\vec{u}$ and $\vec{v}$ is given by the formula

$$< \vec{u}, \vec{v} >= x_1 x_2 + y_1 y_2 + z_1 z_2.$$

We have seen that the inner product satisfies the following axioms:

(I1) $<\vec{u},\vec{v}>=<\vec{v},\vec{u}>$ (symmetry axiom)
(I2) $<\vec{u},\vec{v}+\vec{w}>=<\vec{u},\vec{v}>+<\vec{u},\vec{w}>$ (additivity axiom)
(I3) $<\alpha\vec{u},\vec{v}>=\alpha<\vec{u},\vec{v}>$ (homogeneity axiom)
(I4) $<\vec{u},\vec{u}>\geq 0$ and $<\vec{u},\vec{u}>=0$ if and only if $\vec{u}=\vec{0}$.

We say that $\mathbb{R}^3$ with the inner product operation is an inner product space. In general, a vector space with a binary operation that satisfies axioms (I1) - (I4) is called an **inner product space.**
Some important properties of inner products are listed in the following theorem.

Theorem 46
In an inner product space the following properties hold:

(a) $<0,u>=<u,0>=0$.
(b) $<u+v,w>=<u,w>+<v,w>$.
(c) $<u,\alpha v>=\alpha<u,v>$.

Proof.
(a) $<0,u>=<u+(-u),u>=<u,u+(-u)>=<u,u>+<u,-u>=<u,u>+<-u,u>=<u,u>-<u,u>=0$. Similarly, $<u,0>=0$.
(b) $<u+v,w>=<w,u+v>=<w,u>+<w,v>=<u,w>+<v,w>$.
(c) $<u,\alpha v>=<\alpha v,u>=\alpha<v,u>=\alpha<u,v>$. ■

We shall now prove a result that will enable us to give a definition for the cosine of an angle between two nonzero vectors in an inner product space. This result has many applications in mathematics.

Theorem 47
If u and v are vectors in an inner product space then

$$<u,v>^2\leq<u,u><v,v> .$$

This is known as the **Cauchy-Schwarz inequality.**

Proof.
If $<u,u>=0$ or $<v,v>=0$ then either $u=0$ or $v=0$. In this case, $<u,v>=0$ and the inequality holds. So suppose that $<u,u>\neq 0$ and $<v,v>\neq 0$. Then $<\frac{u}{||u||},\frac{v}{||v||}>\leq 1$. Since, $||\frac{u}{||u||}||=||\frac{v}{||v||}||=1$ then it suffices to show that $<u,v>\leq 1$ for all u,v of norm 1.
We have,

$$\begin{array}{rcl} 0 & \leq & <u-v,u-v> \\ & = & <u,u>+<v,v>-2<u,v> \end{array}$$

This implies that $2<u,v>\leq ||u||^2+||v||^2=1+1=2$. That is, $<u,v>\leq 1$. ■

Exercise 241 *(Normed Vector Spaces)*
Let V be an inner product space. For u in V define $||u||=<u,u>^{\frac{1}{2}}$. Show that the norm function $||.||$ satisfies the following axioms:

(a) $||u|| \geq 0$.
(b) $||u|| = 0$ *if and only if* $u = 0$.
(c) $||\alpha u|| = |\alpha|||u||$.
(d) $||u + v|| \leq ||u|| + ||v||$. *(Triangle Inequality)*

Solution.
(a) Follows from I4.
(b) Follows from I4.
(c) $||\alpha u||^2 =< \alpha u, \alpha u >= \alpha^2 < u, u >$ and therefore $||\alpha u|| = |\alpha| < u, u >^{\frac{1}{2}} = |\alpha|||u||$.
(d) $||u + v||^2 =< u + v, u + v >=< u, u > +2 < u, v > + < v, v >\leq< u, u > +2 < u, u >^{\frac{1}{2}} < v, v >^{\frac{1}{2}} + < v, v >= ||u||^2 + 2||u||||v|| + ||v||^2 = (||u|| + ||v||)^2$. Now take the square root of both sides ■

Exercise 242
The purpose of this exercise is to provide a matrix formula for the dot product in $\mathbb{R}^n$. *Let*

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

be two vectors in $\mathbb{R}^n$. *Show that* $< \vec{u}, \vec{v} >= \vec{v}^T \vec{u}$.

Solution.
On one hand, we have $< \vec{u}, \vec{v} >= u_1v_1 + u_2v_2 + \cdots u_nv_n$. On the other hand,

$$\vec{v}^T \vec{u} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n ■$$

Exercise 243
Show that $< A, B >= tr(AB^T)$, *where* tr *is the trace function, is an inner product in* M_{mn}.

Solution.
(a) $< A, A >=$ sum of the squares of the entries of A and therefore is non-negative.
(b) $< A, A >= 0$ if and only if $\sum_{i=1}^n \sum_{k=1}^n a_{ik}^2 = 0$ and this is equivalent to $A = \mathbf{0}$.
(c) $< A, B >= tr(AB^T) = tr((AB^T)^T) = tr(BA^T) =< B, A >$.
(d) $< A, B+C >= tr(A(B+C)^T) = tr(AB^T+AC^T) = tr(AB^T)+tr(AC^T) =< A, B > + < A, C >$.
(e) $< \alpha A, B >= tr(\alpha AB^T) = \alpha tr(AB^T) = \alpha < A, B >$. Thus, M_{mn} is an inner product space ■

Exercise 244
Show that $< p, q >= a_0b_0 + a_1b_1 + \cdots + a_nb_n$ *defines an inner product in* P_n, *the vector space of all polynomials of degree* n.

Solution.
(a) $< p, p >= a_0^2 + a_1^2 + \cdots + a_n^2 \geq 0.$
(b) $< p, p >= 0$ if and only if $a_0^2 + a_1^2 + \cdots + a_n^2 = 0$ and this is equivalent to $a_0 = a_1 = \cdots = a_n = 0.$
(c) $< p, q >= a_0b_0 + a_1b_1 + \cdots + a_nb_n = b_0a_0 + b_1a_1 + \cdots + b_na_n =< q, p > .$
(d) $< \alpha p, q >= \alpha a_0b_0 + \alpha a_1b_1 + \cdots + \alpha a_nb_n = \alpha(a_0b_0 + a_1b_1 + \cdots + a_nb_n) = \alpha < p, q > .$
(e) $< p, q + r >= a_0(b_0 + c_0) + a_1(b_1 + c_1) + \cdots + a_n(b_n + c_n) = a_0b_0 + a_0c_0 + a_1b_1 + a_1c_1 + \cdots + a_nb_n + a_nc_n =< p, q > + < p, r > .$
Therefore, P_n is an inner product space ■

4.3 Linear Independence

The concepts of linear combination, spanning set, and basis for a vector space play a major role in the investigation of the structure of any vector space. In this section we introduce and discuss the first two concepts and the third one will be treated in the next section.
The concept of linear combination will allow us to generate vector spaces from a given set of vectors in a vector space .
Let V be a vector space and $v_1, v_2, \cdots, v_n$ be vectors in V. A vector $w \in V$ is called a **linear combination** of the vectors $v_1, v_2, \ldots, v_n$ if it can be written in the form

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

for some scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Exercise 245
Show that the vector $\vec{w} = (9, 2, 7)$ *is a linear combination of the vectors* $\vec{u} = (1, 2, -1)$ *and* $\vec{v} = (6, 4, 2)$ *whereas the vector* $\vec{w'} = (4, -1, 8)$ *is not.*

Solution.
We must find numbers s and t such that

$$(9, 2, 7) = s(1, 2, -1) + t(6, 4, 2)$$

This leads to the system

$$\left\{\begin{array}{ccccc} s & + & 6t & = & 9 \\ 2s & + & 4t & = & 2 \\ -s & + & 2t & = & 7 \end{array}\right.$$

Solving the first two equations one finds $s = -3$ and $t = 2$ both values satisfy the third equation.

Turning to $(4, -1, 8)$, the question is whether s and t can be found such that $(4, -1, 8) = s(1, 2, -1) + t(6, 4, 2)$. Equating components gives

$$\left\{\begin{array}{rcrcr} s & + & 6t & = & 4 \\ 2s & + & 4t & = & -1 \\ -s & + & 2t & = & 8 \end{array}\right.$$

Solving the first two equations one finds $s = -\frac{11}{4}$ and $t = \frac{9}{8}$ and these values do not satisfy the third equation. That is the system is inconsistent ■

The process of forming linear combinations leads to a method of constructing subspaces, as follows.

Theorem 48
Let $W = \{v_1, v_2, \ldots, v_n\}$ be a subset of a vector space V. Let $span(W)$ be the collection of all linear combinations of elements of W. Then $span(W)$ is a subspace of V.

Proof.
Let $u, v \in W$. Then there exist scalar $\alpha_1, \alpha_2, \cdots, \alpha_n$ and $\beta_1, \beta_2, \cdots, \beta_n$ such that $u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ and $v = \beta_1 v_1 + \beta_2 v_2 + \cdots \beta_n v_n$. Thus, $u + v = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \cdots + (\alpha_n + \beta_n)v_n \in span(W)$ and $\alpha u = (\alpha\alpha_1)v_1 + (\alpha\alpha_2)v_2 + \cdots + (\alpha\alpha_n)v_n \in Span(W)$. Thus, $span(W)$ is a subspace of V. ■

Exercise 246
Show that $P_n = span\{1, x, x^2, \cdots, x^n\}$.

Solution.
If $p(x) \in P_n$ then there are scalars $a_0, a_1 \cdots, a_n$ such that $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in span\{1, x, \cdots, x^n\}$ ■

Exercise 247
Show that $\mathbb{R}^n = span\{e_1, e_2, \cdots, e_n\}$ where e_i is the vector with 1 in the ith component and 0 otherwise.

Solution.
We must show that if $u \in \mathbb{R}^n$ then u is a linear combination of the $e_i's$. Indeed, if $u = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ then

$$u = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

Hence u lies in $span\{e_1, e_2, \cdots, e_n\}$ ■

If every element of V can be written as a linear combination of elements of W then we have $V = span(W)$ and in this case we say that W is a **span** of V or W **generates** V.

Exercise 248
(a)Determine whether $\vec{v_1} = (1,1,2), \vec{v_2} = (1,0,1)$ *and* $\vec{v_3} = (2,1,3)$ *span* $\mathbb{R}^3$.
(b) Show that the vectors $\vec{i} = (1,0,0), \vec{j} = (0,1,0)$, *and* $\vec{k} = (0,0,1)$ *span* $\mathbb{R}^3$.

Solution.
(a) We must show that an arbitrary vector $\vec{v} = (a,b,c)$ in $\mathbb{R}^3$ is a linear combination of the vectors $\vec{v_1}, \vec{v_2}$, and $\vec{v_3}$. That is $\vec{v} = s\vec{v_1} + t\vec{v_2} + w\vec{v_3}$. Expressing this equation in terms of components gives

$$\left\{\begin{array}{ccccccc} s & + & t & + & 2w & = & a \\ s & + & & + & w & = & b \\ 2s & + & t & + & 3w & = & c \end{array}\right.$$

The problem is reduced of showing that the above system is consistent. This system will be consistent if and only if the coefficient matrix A

$$A = \left(\begin{array}{ccc} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{array}\right)$$

is invertible. Since $|A| = 0$ then the system is inconsistent and therefore $\mathbb{R}^3 \neq span\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$.
(b) See Exercise 247 ■

As you have noticed by now, to show that n vectors in $\mathbb{R}^n$ span $\mathbb{R}^n$ it suffices to show that the matrix whose columns are the given vectors is invertible.

Exercise 249
Show that **P**, *the vector space of all polynomials, cannot be spanned by a finite set of polynomials.*

Solution.
Suppose that $P = span\{1, x, x^2, \cdots, x^n\}$ for some positive integer n. But then the polynomial $p(x) = 1 + x + x^2 + \cdots + x^n + x^{n+1}$ is in P but not in $span\{1, x, x^2, \cdots, x^n\}$, a contradiction. Thus, P cannot be spanned by a finite set of polynomials ■

Exercise 250
Show that $span\{0, v_1, v_2, \cdots, v_n\} = span\{v_1, v_2, \cdots, v_n\}$.

Solution.
If $v \in span\{0, v_1, v_2, \cdots, v_n\}$ then $v = \alpha_0 0 + \alpha_1 v_1 + \cdots + \alpha_n v_n = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in span\{v_1, v_2, \cdots, v_n\}$. Conversely, if $v \in span\{v_1, v_2, \cdots, v_n\}$ then $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 1(0) + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in span\{0, v_1, v_2, \cdots, v_n\}$ ■

Next, we introduce a concept which guarantees that any vector in the span of a set S has only one representation as a linear combination of vectors in S.

Spanning sets with this property play a fundamental role in the study of vector spaces as we shall see in the next section.
If $v_1, v_2, \ldots, v_n$ are vectors in a vector space with the property that

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

holds only for $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ then the vectors are said to be **linearly independent.** If there are scalars not all 0 such that the above equation holds then the vectors are called **linearly dependent.**

Exercise 251
Show that the set $S = \{1, x, x^2, \cdots, x^n\}$ *is a linearly independent set in* P_n.

Solution.
Suppose that $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$ for all $x \in \mathbb{R}$. By the Fundamental Theorem of Algebra, a polynomial of degree n has at most n roots. But by the above equation, every real number is a root of the equation. This forces the numbers $a_0, a_1, \cdots, a_n$ to be 0 ■

Exercise 252
Let u *be a nonzero vector. Show that* $\{u\}$ *is linearly independent.*

Solution.
Suppose that $\alpha u = 0$. If $\alpha \neq 0$ then we can multiply both sides by α^{-1} and obtain $u = 0$. But this contradicts the fact that u is a nonzero vector ■

Exercise 253
(a) Show that the vectors $\vec{v_1} = (1, 0, 1, 2), \vec{v_2} = (0, 1, 1, 2)$, *and* $\vec{v_3} = (1, 1, 1, 3)$ *are linearly independent.*
(b) Show that the vectors $\vec{v_1} = (1, 2, -1), \vec{v_2} = (1, 2, -1)$, *and* $\vec{v_3} = (1, -2, 1)$ *are linearly dependent.*

Solution.
(a) Suppose that s, t, and w are real numbers such that $s\vec{v_1} = t\vec{v_2} + w\vec{v_3} = \mathbf{0}$. Then equating components gives

$$\left\{\begin{array}{ccccccc} s & & & + & w & = & 0 \\ & & t & + & w & = 0 & \\ s & + & t & + & w & = & 0 \\ 2s & + & 2t & + & 3w & = & 0 \end{array}\right.$$

The second and third equation leads to $s = 0$. The first equation gives $w = 0$ and the second equation gives $t = 0$. Thus, the given vectors are linearly independent.
(b) These vectors are linearly dependent since $\vec{v_1} + \vec{v_2} - 2\vec{v_3} = \mathbf{0}$ ■

Exercise 254
Show that the polynomials $p_1(x) = 1 - x, p_2(x) = 5 + 3x - 2x^2$, *and* $p_3(x) = 1 + 3x - x^2$ *are linearly dependent vectors in* P_2.

Solution.
Indeed, $3p_1(x) - p_2(x) + 2p_3(x) = 0$ ■

Exercise 255
Show that the unit vectors $e_1, e_2, \cdots, e_n$ in $\mathbb{R}^n$ are linearly independent.

Solution.
Suppose that $x_1e_1 + x_2e_2 + \cdots + x_ne_n = (0, 0, \cdots, 0)$. Then $(x_1, x_2, \cdots, x_n) = (0, 0, \cdots, 0)$ and this leads to $x_1 = x_2 = \cdots = x_n = 0$. Hence the vectors $e_1, e_2, \cdots, e_n$ are linearly independent ■

The following theorem provides us with a criterion for deciding whether a set of vectors is linearly dependent.

Theorem 49
The vectors $v_1, v_2, \ldots, v_n$ are linearly dependent if and only if there is at least one vector that is a linear combination of the remaining vectors.

Proof.
Suppose that $v_1, v_2, \cdots, v_n$ are linearly dependent then there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$ not all zeros such that $\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n = 0$. Suppose that $\alpha_i \neq 0$. Dividing through by this scalar we get $v_i = -\frac{\alpha_1}{\alpha_i}v_1 - \cdots - \frac{\alpha_{i-1}}{\alpha_i}v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}v_{i+1} - \frac{\alpha_n}{\alpha_i}v_n$.
Conversely, suppose that one of the vectors $v_1, v_2, \cdots, v_n$ is a linear combination of the remaining vectors. Say, $v_i = \alpha_1v_1 + \cdots + \alpha_{i-1}v_{i-1} + \alpha_{i+1}v_{i+1} + \cdots + \alpha_nv_n$. Then $\alpha_1v_1 + \cdots + \alpha_{i-1}v_{i-1} + (-1)v_i + \cdots + \alpha_nv_n = 0$ with the coefficient of v_i being -1. Thus, the vectors $v_1, v_2, \cdots, v_n$ are linearly dependent. This ends a proof of the theorem ■

Exercise 256
Let $S = \{v_1, v_2, \ldots, v_n\}$ be a set of vectors in a vector space V. Show that if one of the vectors is zero, then the set is linearly dependent.

Solution.
Suppose for the sake of argument that $v_1 = 0$. Then $1v_1 + 0v_2 + \cdots + 0v_n = 0$ with coefficients of $v_1, v_2, \cdots, v_n$ not all 0. Thus, the set $\{v_1, v_2, \cdots, v_n\}$ is linearly dependent ■

A criterion for linear independence is established in the following theorem.

Theorem 50
Nonzero vectors $v_1, v_2, \cdots, v_n$ are linearly independent if and only if one of them, say v_i, is a linear combination of the preceding vectors:

$$v_i = \alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_{i-1}v_{i-1}. \tag{4.1}$$

Proof.
Suppose that (4.1) holds. Then $\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_{i-1}v_{i-1} + (-1)v_i + 0v_{i+1} + \cdots + 0v_n = 0$. Thus, $v_1, v_2, \cdots, v_n$ are linearly dependent.
Conversely, suppose that $v_1, v_2, \cdots, v_n$ are linearly dependent. Then there exist scalars not all zero such that $\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n = 0$. Let i be the largest index such that $\alpha_i \neq 0$. Then, $\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_iv_i = 0$. Thus, $v_i = -\sum_{k=1}^{i-1} \frac{\alpha_k}{\alpha_i} v_k$ ■

We conclude this section with a theorem which will found to be very useful in later sections.

Theorem 51
Let S_1 and S_2 be finite subsets of a vector space V such that S_1 is a subset of S_2.

(a) If S_1 is linearly dependent then so is S_2.
(b) If S_2 is linearly independent then so is S_1.

Proof.
Without loss of generality we may assume that $S_1 = \{v_1, v_2, \cdots, v_n\}$ and $S_2 = \{v_1, v_2, \cdots, v_n, v_{n+1}\}$.
(a) Suppose that S_1 is a linearly dependent set. We want to show that S_2 is also linearly dependent. By the assumption on S_1 there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$ not all 0 such that
$$\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n = 0.$$
But this implies that
$$\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n + 0v_{n+1} = 0$$
with $\alpha_i \neq 0$ for some i. That is, S_2 is linearly dependent.
(b) Now, suppose that S_2 is a linearly independent set. We want to show that S_1 is also linearly independent. Indeed, if
$$\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n + = 0$$
then
$$\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n + 0v_{n+1} = 0.$$
Since S_2 is linearly independent then $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. But this shows that S_1 is linearly independent. ■

Exercise 257
Let A be an $n \times m$ matrix in reduced row-echelon form. Prove that the nonzero rows of A, viewed as vectors in $\mathbb{R}^m$, form a linearly independent set of vectors.

Solution.
Let $S_1 = \{v_1, v_2, \cdots, v_k\}$ be the set of nonzero rows of a reduced row-echelon $n \times m$ matrix A. Then $S_1 \subset \{e_1, e_2 \cdots, e_m\}$. Since the set $\{e_1, e_2, \cdots, e_m\}$ is linearly independent then by Theorem 51 (b), S_1 is linearly independent ■

4.4 Basis and Dimension

In this section we continue the study of the structure of a vector space by determining a set of vectors that completely describes the vector space.
Let $S = \{v_1, v_2, \ldots, v_n\}$ be a subset of a vector space V. We say that S is a **basis** for V if

(i) S is linearly independent set.
(ii) $V = span(S)$.

Exercise 258
Let e_i be the vector of $\mathbb{R}^n$ whose ith component is 1 and zero otherwise. Show that the set $S = \{e_1, e_2, \ldots, e_n\}$ is a basis for $\mathbb{R}^n$. This is called the **standard basis** *of $\mathbb{R}^n$.*

Solution.
By Exercise 247, we have $\mathbb{R}^n = span\{e_1, e_2, \cdots, e_n\}$. By Exercise 255, the vectors $e_1, e_2, \cdots, e_n$ are linearly independent. Thus $\{e_1, e_2, \cdots, e_n\}$ is a basis of $\mathbb{R}^3$ ■

Exercise 259
Show that $\{1, x, x^2, \cdots, x^n\}$ is a basis of P_n.

Solution.
By Exercise 246, $P_n = span\{1, x, x^2, \cdots, x^n\}$ and by Exercise 251, the set $S = \{1, x, x^2, \cdots, x^n\}$ is linearly independent. Thus, S is a basis of P_n ■

If $S = \{v_1, v_2, \ldots, v_n\}$ is a basis for V then we say that V is a **finite dimensional space** of dimension n. We write $dim(V) = n$. A vector space which is not finite dimensional is said to be **infinite dimensional** vector space. We define the zero vector space to have dimension zero. The vector spaces $M_{mn}, \mathbb{R}^n$, and P_n are finite-dimensional spaces whereas the space P of all polynomials and the vector space of all real-valued functions defined on $\mathbb{R}$ are inifinite dimensional vector spaces.
Unless otherwise specified, the term vector space shall always mean a finite-dimensional vector space.

Exercise 260
Determine a basis and the dimension for the solution space of the homogeneous system

$$\left\{\begin{array}{ccccccccccc} 2x_1 & + & 2x_2 & - & x_3 & + & & + & x_5 & = & 0 \\ -x_1 & - & x_2 & + & 2x_3 & - & 3x_4 & + & x_5 & = & 0 \\ x_1 & + & x_2 & - & 2x_3 & & & - & x_5 & = & 0 \\ & & & & x_3 & + & x_4 & + & x_5 & = & 0 \end{array}\right.$$

Solution.
By Exercise 38, we found that $x_1 = -s - t, x_2 = s, x_3 = -t, x_4 = 0, x_5 = t$. So if S is the vector space of the solutions to the given system then $S = \{(-s-t, s, -t, 0, t) : s, t \in \mathbb{R}\} = \{s(-1, 1, 0, 0, 0)+t(-1, 0, -1, 0, 1) : s, t \in \mathbb{R}\} = span\{(-1, 1, 0, 0, 0), (-1, 0, -1, 0, 1)\}$. Moreover, if $s(-1, 1, 0, 0, 0)+t(-1, 0, -1, 0, 1) = (0, 0, 0, 0, 0)$ then $s = t = 0$. Thus the set $\{(-1, 1, 0, 0, 0), (-1, 0, -1, 0, 1)\}$ is a basis for the solution space of the homogeneous system ■

The following theorem will indicate the importance of the concept of a basis in investigating the structure of vector spaces. In fact, a basis for a vector space V determines the representation of each vector in V in terms of the vectors in that basis.

Theorem 52
If $S = \{v_1, v_2, \ldots, v_n\}$ is a basis for V then any element of V can be written in one and only one way as a linear combination of the vectors in S.

Proof.
Suppose $v \in V$ has the following two representations $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ and $v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$. We want to show that $\alpha_i = \beta_i$ for $1 \le i \le n$. But this follows from $(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n = 0$ and the fact that S is a linearly independent set. ■

We wish to emphasize that if a spanning set for a vector space is not a basis then a vector may have different representations in terms of the vectors in the spanning set. For example let

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$$

then the vector

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

has the following two different representations

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -\frac{4}{3}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

The following theorem indicates how large a linearly independent set can be in a finite-dimensional vector space.

Theorem 53
If $S = \{v_1, v_2, \ldots, v_n\}$ is a basis for a vector space V then every set with more than n vectors of V is linearly dependent.

Proof.
Let $S' = \{w_1, w_2, \ldots, w_m\}$ be any subset of V with $m > n$ vectors. Since S is a basis then we can write

$$w_i = a_{i1}v_1 + a_{i2}v_2 + \ldots + a_{in}v_n,$$

where $1 \leq i \leq m$. Now, suppose that

$$\alpha_1 w_1 + \alpha_2 w_2 + \ldots + \alpha_m w_m = 0.$$

Then this implies

$$\begin{gathered}(\alpha_1 a_{11} + \alpha_2 a_{21} + \cdots + \alpha_m a_{m1})v_1 + (\alpha_1 a_{12} + \alpha_2 a_{22} + \cdots + \alpha_m a_{m2})v_2 \\ + \cdots + (\alpha_1 a_{1n} + \alpha_2 a_{2n} + \ldots + \alpha_m a_{mn})v_n = 0.\end{gathered}$$

Since S is linearly independent then we have

$$\begin{array}{ccccccccc} \alpha_1 a_{11} & + & \alpha_2 a_{21} & + & \ldots & + & \alpha_m a_{m1} & = & 0 \\ \alpha_1 a_{12} & + & \alpha_2 a_{22} & + & \ldots & + & \alpha_m a_{m2} & = & 0 \\ \vdots & & & & & & & & \\ \alpha_1 a_{1n} & + & \alpha_2 a_{2n} & + & \ldots & + & \alpha_m a_{mn} & = & 0 \end{array}$$

But the above homogeneous system has more unknowns then equations and therefore it has nontrivial solutions (Theorem 6). Thus, S' is linearly dependent $\blacksquare$

An immediate consequence of the above theorem is the following

Theorem 54
Suppose that V is a vector space of dimension n. If S is a subset of V of m linearly independent vectors then $m \leq n$.

Proof.
If not, i.e. $m > n$ then by Theorem 53 S is linearly dependent which is a contradiction. $\blacksquare$

We wish to emphasize here that a basis (when it exists) needs not be unique.

Exercise 261
Show that

$$S = \left\{ \left(\begin{array}{c} 1 \\ -1 \end{array} \right), \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \right\}$$

is a basis of $\mathbb{R}^2$.

Solution.
Suppose that

$$s \left(\begin{array}{c} 1 \\ -1 \end{array} \right) + t \left(\begin{array}{c} 3 \\ 2 \end{array} \right) = \left(\begin{array}{c} 0 \\ 0 \end{array} \right)$$

Then this leads to the equations $s+3t=0$ and $-s+2t=0$. Solving this system we find $s=t=0$. Thus, we have shown that S is linearly independent. Next, we show that every vector in $\mathbb{R}^2$ lies in the span of S. Indeed, if $v=(a,b)^T \in \mathbb{R}^2$ we must show that there exist scalars s and t such that

$$s\begin{pmatrix} 1 \\ -1 \end{pmatrix} + t\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

That is the coefficient matrix of this system

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

is invertible. Since $|A| = 5$ then the given system is consistent. This shows that $\mathbb{R}^2 = spanS$. Hence, S is a basis of $\mathbb{R}^2$. Thus, we have found a basis for $\mathbb{R}^2$ which is different than the standard basis ■

A vector space can have different bases; however, all of them have the same number of elements as indicated by the following theorem.

Theorem 55
Let V be a finite dimensionsl space. If $S = \{v_1, v_2, \cdots, v_n\}$ and $S' = \{w_1, \cdots, w_m\}$ are two bases of V then $n = m$. That is, any two bases for a finite dimensional vector space have the same number of vectors.

Proof.
Since S is a basis of V and S' is a linearly independent susbet of V then by Theorem 53 we must have $m \le n$. Now interchange the roles of S and S' to obtain $n \le m$. Hence, $m = n$ ■

Next, we consider the following problem: If S is a span of a finite dimensional vector space then is it possible to find a subset of S that is a basis of V? The answer to this problem is given by the following theorem.

Theorem 56
If $S = \{v_1, v_2, \cdots, v_n\}$ spans a vector space V then S contains a basis of V and $dim(V) \le n$.

Proof.
Either S is linearly independent or linearly dependent. If S is linearly independent then we are finished. So assume that S is linearly dependent. Then there is a vector v_i which is a linear combination of the preceding vectors in S (Theorem 50). Thus, $V = spanS = spanS_1$ where $S_1 = \{v_1, v_2, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\}$.
Now we repeat the argument. If S_1 is linearly independent then S_1 is a basis of V. Otherwise delete a vector of S_1 which is a linear combination of the preceding vectors and obtain a subset S_2 of S_1 that spans V. Continue this process. If a basis is encountered at some stage, we are finished. If not, we ultimately reach $V = span\{v_l\}$ for some l. Since $\{v_l\}$ is also linearly independent then $\{v_l\}$ is a basis of V. ■

Exercise 262
Let $S = \{(1,2),(2,4),(2,1),(3,3),(4,5)\}$

(a) Show that $\mathbb{R}^2 = spanS$.
(b) Find a subset S' *of* S *such that* S' *is a basis of* $\mathbb{R}^2$.

Solution.
(a) Note that $(2,4) = (1,2)$ and $(3,3) = (2,1)+(1,2)$ and $(4,5) = (2,1)+2(1,2)$. Thus, $spanS = span\{(1,2),(2,1)\}$. Now, for any vector $(x,y) \in \mathbb{R}^2$ we must show the existence of two numbers a and b such that $(x,y) = a(1,2) + b(2,1)$. This is equivalent to showing that the system

$$\left\{\begin{array}{ccccc} a & + & 2b & = & x \\ 2a & + & b & = & y \end{array}\right.$$

is consistent. The coefficient matrix of the system is

$$A = \left(\begin{array}{cc} 1 & 2 \\ 2 & 1 \end{array}\right)$$

This matrix is invertible since $|A| = -3$. Thus the system is consistent and therefore $\mathbb{R}^2 = spanS$.
(b) $S' = \{(1,2),(2,1)\}$ ■

We now prove a theorem that we shall have occasion to use several times in constructing a basis containing a given set of linearly independent vectors.

Theorem 57
If $S = \{v_1, v_2, \ldots, v_r\}$ *is a set of* r *linearly independent vectors in a* n*-dimensional vector space* V *and* $r < n$ *then there exist* $n - r$ *vectors* $v_{r+1}, \cdots, v_n$ *such that the enlarged set* $S' = S \cup \{v_{r+1}, \cdots, v_n\}$ *a basis for* V.

Proof.
Since V is of dimension n and S is a linearly independent set then $r \leq n$ by Theorem 57. If $span(S) = V$ then S is a basis of V and by Theorem 55, $r = n$. So suppose $span(S) \neq V$, i.e. $r < n$. Then,, there is a vector $v_{r+1} \in V$ which is not a linear combination of the vectors of S. In this case, the set $\{v_1, v_2, \cdots, v_r, v_{r+1}\}$ is linearly independent. As long as $span\{v_1, v_2, \cdots, v_r, v_{r+1}\} \neq V$ we continue the expansion as above. But the number of linearly independent vectors in the expansion of S can not exceed n according to Theorem 58. That is, the process will terminate in a finite number of steps ■

It follows from the above theorem that a finite-dimensional nonzero vector space always has a finite basis.

Exercise 263
Construct a basis for $\mathbb{R}^3$ *that contains the vector* $(1,2,3)$.

Solution.
Since the vector $(1,0,0)$ is not a multiple of the vector $(1,2,3)$ then the set $\{(1,2,3),(1,0,0)\}$ is linearly independent. Since $dim(\mathbb{R}^3) = 3$ then this set is not a basis of $\mathbb{R}^3$ and therefore cannot span $\mathbb{R}^3$. Next, we see that the vector $(0,1,0)$ is not a linear combination of $\{(1,2,3),(1,0,0)\}$. That is, $(0,1,0) = s(1,2,3) + t(1,0,0)$ for some scalars $s,t \in \mathbb{R}$. This leads to the system

$$\left\{\begin{array}{ccccc} s & + & t & = & 0 \\ 2s & & & = & 1 \\ 3s & & & = & 0 \end{array}\right.$$

Because of the last two equations this system is inconsistent. Thus the set $S' = \{(1,2,3),(1,0,0),(0,1,0)\}$ is linearly independent. By Theorem 58 below, S' is a basis of $\mathbb{R}^3$ ■

In general, to show that a set of vectors S is a basis for a vector space V we must show that S spans V and S is linearly independent. However, if we happen to know that V has dimension equals to the number of elements in S then it suffices to check that either linear independence or spanning- the remaining condition will hold automatically. This is the content of the following two theorems.

Theorem 58
If $S = \{v_1, v_2, \ldots, v_n\}$ is a set of n linearly independent vectors in a n-dimensional vector space V, then S is a basis for V.

Proof.
Since S is a linearly independent set then by Theorem 57 we can extend S to a basis W of V. By Theorem 55, W must have n vectors. Thus, $S = W$ and so S is a basis of V. ■

Theorem 59
If $S = \{v_1, v_2, \ldots, v_n\}$ is a set of n vectors that spans an n-dimensional vector space V, then S is a basis for V.

Proof.
Suppose that V is of dimension n and S is a span of V. If S is linearly independent then S is a basis and we are done. So, suppose that S is linearly dependent. Then by Theorem 49 there is a vector in S which can be written as a linear combination of the remaining vectors. By rearrangement, we can assume that v_n is a linear combination of $v_1, v_2, \cdots, v_{n-1}$. Hence, $span\{v_1, v_2, \cdots, v_n\} = span\{v_1, v_2, \cdots, v_{n-1}\}$. Now, we repeat this process until the spanning set is linearly independent. But in this case the basis will have a number of elements $< n$ and this contradicts Theorem 55 ■

Exercise 264
Show that if V is a finite dimensional vector space and W is a subspace of V then $dim(W) \leq dim(V)$.

Solution.
If $W = \{0\}$ then $dim(W) = 0 \leq dim(V)$. So suppose that $W \neq \{0\}$ and let $0 \neq u_1 \in W$. If $W = span\{u_1\}$ then $dim(W) = 1$ If $W \neq span\{u_1\}$ then there exists a vector $u_2 \in W$ and $u \notin span\{u_1\}$. In this case, $\{u_1, u_2\}$ is linearly independent. If $W = span\{u_1, u_2\}$ then $dim(W) = 2$. If not, repeat the process to find $u_3 \in W$ and $u_3 \notin span\{u_1, u_2\}$. Then $\{u_1, u_2, u_3\}$ is linearly independent. Continue in this way. The process must terminate because the vector space cannot have more than $dim(V)$ linearly independent vectors. Hence W has a basis of at most $dim(V)$ vectors ■

Exercise 265
Show that if W is a subspace of a finite dimensional vector space V and $dim(W) = dim(V)$ then $W = V$.

Solution.
Suppose $dim(W) = dim(V) = m$. Then any basis $\{u_1, w_2, \cdots, u_m\}$ of W is an independent set of m vectors in V and so is a basis of V by Theorem 59. In particular $V = span\{u_1, u_2, \cdots, u_m\} = W$ ■

Next, we discuss an example of an infinite dimensional vector space. We have stated previously that the vector space $F(\mathbb{R})$ is an infinite dimensional vector space. We are now in a position to prove this statement. We must show that every finite subset of $F(\mathbb{R})$ fails to be a spanning set of $F(\mathbb{R})$. We prove this by contradiction.
Suppose that $S = \{f_1, f_2, \cdots, f_n\}$ spans $F(\mathbb{R})$. Without loss of generality we may assume that $0 \notin S$. By Theorem 56 there exists a subset S' of S such that S' is a basis of $F(\mathbb{R})$. Hence, $dim(S') \leq n$. The set $\{1, x, \cdots, x^n\}$ is contained in $F(\mathbb{R})$ and is linearly independent and contains $n+1$ elements. This violates Theorem 53.

In the remainder of this section we consider inner product spaces.
A set of vectors $\{v_1, v_2, \ldots, v_n\}$ of an inner product space is said to be **orthonormal** if the following two conditions are met

(i) $< v_i, v_j >= 0$, for $i \neq j$.
(ii) $< v_i, v_i >= 1$ for $1 \leq i \leq n$.

Exercise 266
(a) Show that the vectors $\vec{v_1} = (0, 1, 0), \vec{v_2} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), \vec{v_3} = (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ form an orthonormal set in $\mathbb{R}^3$ with the Euclidean inner product.
(b) If $S = \{v_1, v_2, \ldots, v_n\}$ is a basis of an inner product space and S is orthonormal then S is called an **orthonormal basis.** *Show that if S is an orthonormal basis and u is any vector then*

$$u =< u, v_1 > v_1+ < u, v_2 > v_2 + \ldots + < u, v_n > v_n.$$

(c) Show that if $S = \{v_1, v_2, \ldots, v_n\}$ is an orthogonal set of nonzero vectors in an inner product space then S is linearly independent.

(d) Use part (b) to show that if $S = \{v_1, v_2, \cdots, v_n\}$ is an orthonormal basis of an inner product space V then every vector u can be written in the form

$$u = w_1 + w_2$$

where w_1 belongs to the span of S and w_2 is orthogonal to the span of S. w_1 is called the **orthogonal projection of u** *on the span of S and is denoted by $proj_{span(S)}u = w_1$.*

Solution.
(a) It is easy to check that $<\vec{v_1}, \vec{v_1}> = <\vec{v_2}, \vec{v_2}> = <\vec{v_3}, \vec{v_3}> = 0$ and $<\vec{v_1}, \vec{v_2}> = <\vec{v_1}, \vec{v_3}> = <\vec{v_2}, \vec{v_3}> = 0$. Thus, $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ is an orthonormal set.
(b) Since S is a basis of an inner product V. For $u \in V$ we have $u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. For $1 \leq i \leq n$ we have $<u, v_i> = \alpha_i$ since S is orthonormal.
(c) We must show that if $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$ then $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Then for $1 \leq i \leq n$ we have $0 = <v_i, 0> = <v_i, \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n> = \alpha_1 <v_1, v_i> + \cdots + \alpha_{i-1} <v_i, v_{i-1}> + \alpha_i <v_i, v_i> + \cdots + \alpha_n <v_i, v_n> = \alpha_i ||v_i||^2$. Since $v_i \neq 0$ then $||v_i|| > 0$. Hence. $\alpha_i = 0$.
(d) Indeed, if $u \in V$ then $u = w_1 + w_2$ where $w_1 = <u, v_1> v_1 + <u, v_2> v_2 + \cdots + <u, v_n> v_n \in span(S)$ and $w_2 = u - w_1$. Moreover, $<w_2, v_i> = <u - w_1, v_i> = <u, v_i> - <w_1, v_i> = <u, v_i> - <u, v_i> = 0, 1 \leq i \leq n$. Thus w_2 is orthogonal to $span(S)$ ■

Exercise 267 *(Gram-Schmidt)*
The purpose of this exercise is to show that every nonzero finite dimensional inner product space with basis $\{u_1, u_2, \ldots, u_n\}$ has an orthonormal basis.

(a) Let $v_1 = \frac{u_1}{||u_1||}$. Show that $||v_1|| = 1$.

(b) Let $W_1 = span\{v_1\}$. Recall that $proj_{W_1} u_2 = <u_2, v_1> v_1$. Let $v_2 = \frac{u_2 - <u_2, v_1> v_1}{||u_2 - <u_2, v_1> v_1||}$. Show that $||v_2|| = 1$ and $<v_1, v_2> = 0$.

(c) Let $W_2 = span\{v_1, v_2\}$ and $proj_{W_2} u_3 = <u_3, v_1> v_1 + <u_3, v_2> v_2$. Let

$$v_3 = \frac{u_3 - <u_3, v_1> v_1 - <u_3, v_2> v_2}{||u_3 - <u_3, v_1> v_1 - <u_3, v_2> v_2||}.$$

Show that $||v_3|| = 1$ and $<v_1, v_3> = <v_2, v_3> = 0$.

Solution.
(a) $||\frac{u_1}{||u_1||}||^2 = <\frac{u_1}{||u_1||}, \frac{u_1}{||u_1||}> = \frac{<u_1, u_1>}{||u_1||^2} = \frac{||u_1||^2}{||u_1||^2} = 1$.
(b) The proof that $||v_2|| = 1$ is similar to (a). On the other hand, $<v_1, v_2> = << u_2, v_1> v_1, \frac{u_2 - <u_2, v_1> v_1}{||u_2 - <u_2, v_1> v_1||}> = \frac{<u_2, v_1>^2}{u_2 - <u_2, v_1> v_1} - \frac{<u_2, v_1>^2}{u_2 - <u_2, v_1> v_1} = 0$.
(c) Similar to (c) ■

Continuing the process of the above exercise we obtain an orthonormal set of vectors $\{v_1, v_2, \dots, v_n\}$. By the previous exercise this set is linearly independent. By Theorem 58, $\{v_1, v_2, \dots, v_n\}$ is a basis of V and consequently an othonormal basis.

4.5 Transition Matrices and Change of Basis

In this section we introduce the concept of coordinate vector in a given basis and then find the relationship between two coordinate vectors of the same vector with respect to two different bases.
From Theorem 52, it follows that if $S = \{v_1, v_2, \dots, v_n\}$ is a basis for a vector space V and u is a vector in V then u can be written uniquely in the form

$$u = \alpha v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

The scalars $\alpha_1, \alpha_2, \dots, \alpha_n$ are called the **coordinates of u relative to the basis S** and we call the matrix

$$[u]_S = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

the **coordinate matrix of u** relative to S.

Exercise 268
Let u and v be two vectors in V and $S = \{v_1, v_2, \cdots, v_n\}$ be a basis of V. Show that
(a) $[u + v]_S = [u]_S + [v]_S$.
(b) $[\alpha u]_S = \alpha[u]_S$, where α is a scalar.

Solution.
(a) Let $u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ and $v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$. Then $u + v = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \cdots + (\alpha_n + \beta_n)v_n$. It follows that

$$[u+v]_S = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = [u]_S + [v]_S$$

(b) Since $\alpha u = \alpha\alpha_1 v_1 + \alpha\alpha_2 v_2 + \cdots + \alpha\alpha_n v_n$ then

$$[\alpha u]_S = \begin{pmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \\ \vdots \\ \alpha\alpha_n \end{pmatrix} = \alpha \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha[u]_S \blacksquare$$

Exercise 269
Let $S = \{v_1, v_2, \cdots, v_n\}$ be a basis of V. Define the function $T : V \longrightarrow \mathbb{R}^n$ given

by $Tv = [v]_S$.
(a) Show that $T(\alpha u + v) = \alpha Tu + Tv$.
(b) Show that if $Tu = Tv$ then $u = v$.
(c) Show that for any $w \in \mathbb{R}^n$ there exists a $v \in V$ such that $Tv = w$.

Solution.
(a) $T(\alpha u + v) = [\alpha u + v]_S = [\alpha u]_S + [v]_S = \alpha[u]_S + [v]_S = \alpha Tu + Tv$.
(b) Suppose that $u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ and $v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$. If $Tu = Tv$ then $\alpha_i = \beta_i$ for $1 \le i \le n$. Hence, $u = v$.
(c) If $w \in \mathbb{R}^n$ then $w = (\alpha_1, \alpha_2, \cdots, \alpha_n)$. Let $u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in V$ and $Tu = w$ ■

If $S' = \{u_1, u_2, \ldots, u_n\}$ is another basis for V then we can write

$$\begin{array}{ccccccccc} u_1 & = & c_{11}v_1 & + & c_{12}v_2 & + & \ldots & c_{1n}v_n \\ u_2 & = & c_{21}v_1 & + & c_{22}v_2 & + & \ldots & c_{2n}v_n \\ \vdots & & & & & & & \vdots \\ u_n & = & c_{n1}v_1 & + & c_{n2}v_2 & + & \ldots & c_{nn}v_n. \end{array}$$

The matrix

$$Q = \begin{pmatrix} c_{11} & c_{21} & \ldots & c_{n1} \\ c_{12} & c_{22} & \ldots & c_{n2} \\ \vdots & & & \vdots \\ c_{1n} & c_{2n} & \ldots & c_{nn} \end{pmatrix}$$

is called the **transition matrix** from S' to S. Note that the jth column of Q is just $[u_j]_S$.

Exercise 270
Let $\vec{v_1} = (0,1,0)^T, \vec{v_2} = (-\frac{4}{5}, 0, \frac{3}{5})^T, \vec{v_3} = (\frac{3}{5}, 0, \frac{4}{5})^T$.
(a) Show that $S = \{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ is a basis of $\mathbb{R}^3$.
(b) Find the coordinate matrix of the vector $\vec{u} = (1,1,1)^T$ relative to the basis S.

Solution.
By Theorem 58, it suffices to show that S is linearly independent. This can be accomplished by showing that the determinant of the matrix with columns $\vec{v_1}, \vec{v_2}, \vec{v_3}$ is invertible. Indeed,

$$\begin{vmatrix} 0 & -\frac{4}{5} & \frac{3}{5} \\ 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4}{5} \end{vmatrix} = 1$$

(b) We have $\vec{u} = \vec{v_1} - \frac{1}{5}\vec{v_2} + \frac{7}{5}\vec{v_3}$. Therefore,

$$[\vec{u}]_S = \begin{pmatrix} 1 \\ -\frac{1}{5} \\ \frac{7}{5} \end{pmatrix} ■$$

Exercise 271
Consider the following two bases of P_2: $S = \{x-1, x+1\}$ and $S' = \{x, x-2\}$. Find the transition matrix from S to S'.

Solution.
Writing the elements of S in terms of the elements of S' we find $x - 1 = \frac{x}{2} + \frac{1}{2}(x-2)$ and $x + 1 = \frac{3}{2}x - \frac{1}{2}(x-2)$. Hence, the transition matrix from S to S' is the matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \blacksquare$$

The next theorem provides a relationship between the coordinate matrices of a vector with respect to two different bases.

Theorem 60
With the above notation we have $[u]_S = Q[u]_{S'}$

Proof.
Suppose

$$[u]_{S'} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

That is, $u = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n$. Now, writting the u_i in terms of the v_i we find that $u = (\beta_1 c_{11} + \beta_2 c_{21} + \cdots + \beta_n c_{n1})v_1 + (\beta_1 c_{12} + \beta_2 c_{22} + \cdots + \beta_n c_{n2})v_2 + \cdots + (\beta_1 c_{1n} + \beta_2 c_{2n} + \cdots c_{nn})v_n$. That is, $[u]_S = Q[u]_{S'}$.$\blacksquare$

Exercise 272
Consider the vectors $\vec{v_1} = (1,0)^T, \vec{v_2} = (0,1)^T, \vec{u_1} = (1,1)^T$, and $\vec{u_2} = (2,1)^T$.
(a) Show that $S = \{\vec{v_1}, \vec{v_2}\}$ and $S' = \{\vec{u_1}, \vec{u_2}\}$ are bases of $\mathbb{R}^2$.
(b) Find the transition matrix Q from S' to S.
(c) Find $[\vec{v}]_S$ given that

$$[\vec{v}]_{S'} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}$$

Solution.
(a) Since $dim(\mathbb{R}^2) = 2$ then it suffices to show that S and S' are linearly independent. Indeed, finding the determinant of the matrix with columns $\vec{v_1}$ and $\vec{v_2}$ we get

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Thus, S is linearly independent. Similar proof for S'.
(b) It is easy to check that $\vec{u_1} = \vec{v_1} + \vec{v_2}$ and $\vec{u_2} = 2\vec{v_1} + \vec{v_2}$. Thus, the transition matrix Q from S' to S is the matrix

$$Q = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$$

(c) By Theorem 60 we have $[\vec{v}]_S = Q[\vec{v}]_{S'} = \left(\begin{array}{cc} 1 & 1 \\ 2 & 1 \end{array}\right)\left(\begin{array}{c} -3 \\ 5 \end{array}\right) = \left(\begin{array}{c} 7 \\ 2 \end{array}\right)$ ■

We next show that Q is nonsingular and its inverse is the transition P matrix from S to S'.

Theorem 61
The matrix Q is invertible and its inverse is the transition matrix from S to S'.

Proof.
Let P be the transition matrix from S to S'. Then by Theorem 60 we have $[u]_{S'} = P[u]_S$. Since $[u]_S = Q[u]_{S'}$ then $[u]_S = QP[u]_S$. Let $x \in \mathbb{R}^n$. Then $x = [u]_S$ for some vector $u \in V$. Thus, $QPx = x$, i.e. $(QP - I_n)x = 0$. Since this is true for arbitrary x then by Exercise 84 we must have $QP = I_n$. By Theorem 20, Q is invertible and $Q^{-1} = P$. ■

Exercise 273
Consider the two bases $S = \{\vec{v_1}, \vec{v_2}\}$ and $S' = \{\vec{u_1}, \vec{u_2}\}$ where $\vec{v_1} = \left(\begin{array}{c} 1 \\ 0 \end{array}\right)$, $\vec{v_2} = \left(\begin{array}{c} 0 \\ 1 \end{array}\right)$, $\vec{u_1} = \left(\begin{array}{c} 1 \\ 1 \end{array}\right)$, and $\vec{u_2} = \left(\begin{array}{c} 2 \\ 1 \end{array}\right)$.
(a) Find the transition matrix P from S to S'.
(b) Find the transition matrix Q from S' to S.

Solution.
(a) Expressing the vectors of S in terms of the vectors of S' we find $\vec{v_1} = -\vec{u_1} + \vec{u_2}$ and $\vec{v_2} = 2\vec{u_1} - \vec{u_2}$. Thus the transition matrix from S to S' is the matrix

$$P = \left(\begin{array}{cc} -1 & 2 \\ 2 & -1 \end{array}\right)$$

(b) We have already found the matrix Q in the previous exercise ■

The next theorem shows that the transition matrix from one orthonormal basis to another has the property $P^T = P^{-1}$. That is, $PP^T = P^TP = I_n$. In this case we call P an **orthogonal** matrix.

Theorem 62
If P is the transition matrix from one orthonormal basis S to another orthonormal basis S' for an inner product space V then $P^{-1} = P^T$.

Proof.
Suppose $P = (b_{ij})$ is the transition matrix from the orthonormal basis $S = \{u_1, u_2, \cdots, u_n\}$ to the orthonormal basis $S' = \{v_1, v_2, \cdots, v_n\}$. We will show that $PP^T = I_n$. Indeed, for $1 \le i \le n$ we have

$$v_i = b_{i1}u_1 + b_{i2}u_2 + \cdots + b_{in}u_n.$$

Thus, $< v_i, v_j >=< b_{i1}u_1 + b_{i2}u_2 + \cdots + b_{in}u_n, b_{j1}u_1 + b_{j2}u_2 + \cdots + b_{jn}u_n >= b_{i1}b_{j1} + b_{i2}b_{j2} + \cdots + b_{in}b_{jn}$ since S is an orthonormal set. Thus, $< v_i, v_j >=< c_i, c_j >$ where $c_1, c_2, \cdots, c_n$ are the columns of P. On the other hand, if $PP^T = (d_{ij})$ then $d_{ij} =< v_i, v_j >$. Since $v_1, v_2, \cdots, v_n$ are orthonormal then $PP^T = I_n$. That is P is orthogonal ■

The following result provides a tool for proving that a matrix is orthogonal.

Theorem 63
Let A be an $n \times n$ matrix. Then the following are equivalent:
(a) A is orthogonal.
(b) The rows of A form an orthonormal set.
(c) The columns of A form an orthonormal set.

Proof.
Let $\{r_1, r_2, \cdots, r_n\}, \{c_1, c_2, \cdots, c_n\}$ denote the rows ond columns of A respectively.
(a) ⇔ (b): We will show that if $AA^T = (b_{ij})$ then $b_{ij} =< r_i, r_j >$. Indeed, from the definition of matrix multiplication we have

$$\begin{aligned} b_{ij} &= (ith\ row\ of\ A)(jth\ column\ of\ A^T) \\ &= a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{in}a_{jn} \\ &= < r_i, r_j > \end{aligned}$$

$AA^T = I_n$ if and only if $b_{ij} = 0$ if $i \neq j$ and $b_{ii} = 1$. That is, $AA^T = I_n$ if and only if $< r_i, r_j >= 0$ for $i \neq j$ and $< r_i, r_i >= 1$.
(a) ⇔ (c): If $A^TA = (d_{ij})$ then one can easily show that $d_{ij} =< c_j, c_i >$. Thus, $A^TA = I_n$ if and only if $d_{ij} = 0$ for $i \neq j$ and $d_{ii} = 1$. Hence, $A^TA = I_n$ if and only if $< c_j, c_i >= 0$ for $i \neq j$ and $< c_i, c_i >= 1$. ■

Exercise 274
Show that the matrix

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

is orthogonal.

Solution.
We will show that this matrix is orthogonal by using the definition. Indeed, using matrix multiplication we find

$$AA^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = I_2 ■$$

Exercise 275
(a) Show that if A is orthogonal then A^T is also orthogonal.
(b) Show that if A is orthogonal then $|A| = \pm 1$.

Solution.
(a)Suppose that A is orthogonal. Then $A^T = A^{-1}$. Taking the transpose of both sides of this equality we find $(A^T)^T = (A^{-1})T = (A^T)^{-1}$. That is, A^T is orthogonal.
(b) Since A is orthogonal then $AA^T = I_n$. Taking the determinant of both sides of this equality to obtain $|AA^T| = |A||A^T| = |A|^2 = 1$ since $|A^T| = |A|$ and $|I_n| = 1$. Hence, $|A| = \pm 1$ ■

4.6 The Rank of a matrix

In this section we study certain vector spaces associated with matrices. We shall also attach a unique number to a matrix that gives us information about the solvability of linear systems and the invertibility of matrices.
Let $A = (a_{ij})$ be an $m \times n$ matrix. For $1 \leq i \leq m$, the ith row of A is the $1 \times n$ matrix

$$r_i = (a_{i1}, a_{i2}, \ldots, a_{in})$$

and for $1 \leq j \leq n$ the jth column of A is the $m \times 1$ matrix

$$c_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

The subspace $span\{r_1, r_2, \ldots, r_m\}$ of $\mathbb{R}^n$ is called the **row space** of A and the subspace $span\{c_1, c_2, \ldots, c_n\}$ of $\mathbb{R}^m$ is called the **column space** of A.
The following theorem will lead to techniques for computing bases for the row and column spaces of a matrix.

Theorem 64
Elementary row operations do not change the row space of a matrix. That is, the row spaces of two equivalent matrices are equal.

Proof.
It suffices to show that if B is the matrix obtained from A by applying a row operation then

$$span\{r_1, r_2, \ldots, r_m\} = span\{r'_1, r'_2, \ldots, r'_m\}$$

where $r'_1, r'_2, \ldots, r'_m$ are the rows of B.
If the row operation is a row interchange then A and B have the same rows and consequently the above equality holds.
If the row operation is multiplication of a row by a nonzero scalar, for instance, $r'_i = \alpha r_i$ then we get $r'_j = r_j$ for $j \neq i$. Hence, r'_j is an element of $span\{r_1, r_2, \ldots, r_m\}$ for $j \neq i$. . Moreover, $r'_i = 0r_1 + \ldots 0r_{i-1} + \alpha r_i + 0r_{i+1} + \ldots + 0r_m$ is also an element of $span\{r_1, r_2, \ldots, r_m\}$. Hence, $span\{r'_1, r'_2, \ldots, r'_m\}$

is a subset of $span\{r_1, r_2, \ldots, r_m\}$. On the other hand, for $j \neq i$ we have $r_j = r'_j$ and therefore $r_j \in span\{r'_1, r'_2, \ldots, r'_m\}$. But $r_i = 0r'_1 + 0r'_2 + \ldots + 0r'_{i-1} + \frac{1}{\alpha}r'_i + 0r'_{i+1} + \ldots + 0r'_m$ which is an element of $span\{r'_1, r'_2, \ldots, r'_m\}$. Hence, $span\{r_1, r_2, \ldots, r_m\}$ is contained in $span\{r'_1, r'_2, \ldots, r'_m\}$.
Similar argument holds when the elementary row operation is an addition of a multiple of one row to another ■

It follows from the above theorem that the row space of a matrix A is not changed by reducing the matrix to a row-echelon form B. It turns out that the nonzero rows of B form a basis of the row space of A as indicated by the following theorem.

Theorem 65
The nonzero row vectors in a row-echelon form of a matrix form a basis for the row space of A.

Proof.
If B is the matrix in echelon form that is equivalent to A. Then by Theorem 64, $span\{r_1, r_2, \cdots, r_m\} = span\{r'_1, r'_2, \cdots, r'_p\}$ where $r'_1, r'_2, \cdots, r'_p$ are the nonzero rows of B with $p \leq m$. We will show that $S = \{r'_1, r'_2, \cdots, r'_p\}$ is a linearly independent set. Suppose not, i.e. $S = \{r'_p, r'_{p-1}, \cdots, r'_1\}$ is linearly dependent. Then by Theorem 50, one of the rows, say r'_i is a linear combination of the preceding rows:

$$r'_i = \alpha_{i+1}r'_{i+1} + \alpha_{i+2}r'_{i+2} + \cdots + \alpha_1 r'_p. \quad (4.2)$$

Suppose that the leading 1 of r'_i occurs at the jth column. Since the matrix is in echelon form, the jth component of $r'_{i+1}r'_{i+2}, \cdots, r'_i$ are all 0 and so the jth component of (4.2) is $\alpha_{i+1}0 + \alpha_{i+2}0 + \cdots + \alpha_p 0 = 0$. But this contradicts the assumption that the jth component of r'_i is 1. Thus, S must be linearly independent ■

Exercise 276
Find a basis for the space spanned by the vectors $v_1 = (1, -2, 0, 0, 3), v_2 = (2, -5, -3, -2, 6), v_3 = (0, 5, 15, 10, 0),$ *and* $v_4 = (2, 6, 18, 8, 6)$.

Solution.
The space spanned by the given vectors is the row space of the matrix

$$\left(\begin{array}{rrrrr} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{array}\right)$$

The reduction of this matrix to row-echelon form is as follows.

Step 1: $r_2 \leftarrow r_2 - 2r_1$ and $r_4 \leftarrow r_4 - 2r_1$

$$\left(\begin{array}{rrrrr} 1 & -2 & 0 & 0 & 3 \\ 0 & -1 & -3 & -2 & 0 \\ 0 & 5 & 15 & 10 & 0 \\ 0 & 10 & 18 & 8 & 0 \end{array}\right)$$

Step 2: $r_3 \leftarrow r_3 + 5r_2$ and $r_4 \leftarrow r_4 + 10r_2$

$$\begin{pmatrix} 1 & -2 & 0 & 0 & 3 \\ 0 & -1 & -3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -12 & -12 & 0 \end{pmatrix}$$

Step 3: $r_2 \leftarrow -r_2$ and $r_3 \leftrightarrow -\frac{1}{12}r_4$

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

Thus, the vectors $(1, -2, 0, 0, 3), (0, 1, 3, 2, 0)$, and $(0, 0, 1, 1, 0)$ form a basis of the vector space spanned by the given vectors ■

Exercise 277
Find a basis for the row space of the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 2 \\ -1 & -2 & 1 & 3 & 0 & -1 \\ 2 & 4 & 1 & 3 & 1 & 9 \\ 1 & 2 & 1 & 3 & 0 & 3 \end{pmatrix}$$

Solution.
The reduction of this matrix to row-echelon form is as follows.

Step 1: $r_2 \leftarrow r_2 + r_1$, $r_3 \leftarrow r_3 - 2r_1$ and $r_4 \leftarrow r_4 - r_1$

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

Step 2: $r_3 \leftarrow r_3 - r_2$ and $r_4 \leftarrow r_4 - r_2$

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

Thus a basis for the vector spanned by the rows of the given matrix consists of the vectors $(1, 2, 0, 0, 0, 2), (0, 0, 1, 3, 0, 1)$, and $(0, 0, 0, 0, 1, 4)$ ■

In the next theorem we list some of the properties of the row and column spaces.

Theorem 66
Let A be an $m \times n$ matrix, U a $p \times m$ matrix, and V an $n \times q$ matrix. Then
(a) the row space of UA is a subset of the row space of A. The two spaces are equal whenever U is nonsingular.
(b) The column space of AV is contained in the solumn space of A. Equality holds when V is nonsingular.

Proof.
(a) Let $U_i = (u_{i1}, u_{i2}, \cdots, u_{im})$ denote the ith row of U. If $r_1, r_2, \cdots, r_m$ denote the rows of A then the ith row of UA is

$$U_i A = (u_{i1}, u_{i2}, \cdots, u_{im}) \begin{pmatrix} r_1 \\ r_2 \\ \cdots \\ r_m \end{pmatrix} = u_{i1}r_1 + u_{i2}r_2 + \cdots + u_{im}r_m$$

This shows that each row of UA is in the row space of A. Hence, the row space of UA is contained in the row space of A. If U is invertible then the row space of A is equal to the row space of $U^{-1}(UA)$ which in turn is a subset of the row space of UA. This ends a proof of (a).
(b) Similar to (a) ■

Next, we will show that a matrix A is row equivalent to a unique reduced row-echelon matrix. First, we prove the following

Theorem 67
If A and B are two reduced row-echelon matrices with the same row space then the leading 1 of A and B occur in the same position.

Proof.
Suppose that the leading 1 of the first row of A occurs in column j_1 and that of B in column k_1. We will show that $j_1 = k_1$. Suppose $j_1 < k_1$. Then the j_1th column of B is zero. Since the first row of A is in the span of the rows of B then the $(1, j_1)$th entry of A is a linear combination of the entries of the j_1th column of B. Hence, $1 = \alpha_1 b_{1j_1} + \alpha_2 b_{2j_1} + \cdots + \alpha_m b_{mj_1} = 0$ since $b_{1j_1} = b_{2j_2} = \cdots = b_{mj_1} = 0$. (Remember that the j_1th column of B is zero). But this is a contradiction. Hence, $j_1 \geq k_1$. Interchanging the roles of A and B to obtain $k_1 \geq j_1$. Hence, $j_1 = k_1$.
Next, let A' be the matrix obtained by deleting the first row of A and B' the matrix obtained by deleting the first row of B. Then clearly A' and B' are in reduced echelon form. We will show that A' and B' have the same row space. Let $R_1, R_2, \cdots, R_m$ denote the rows of B. Let $r = (a_1, a_2, \cdots, a_n)$ be any row of A'. Since r is also a row of A then we can find scalars $d_1, d_2, \cdots, d_m$ such that $r = d_1R_1 + d_2R_2 + \cdots + d_mR_m$. Since A is in reduced row echelon form and r is not the first row of A then $a_{j_1} = a_{k_1} = 0$. Furthermore, since B is in reduced row echelon form all the entries of the k_1th column of B are zero except $b_{1k_1} = 1$. Thus,

$$0 = a_{k_1} = d_1 b_{1k_1} + d_2 b_{2k_1} + \cdots + d_m b_{mk_1}.$$

That is, $d_1 = 0$. Hence, r is in the row space of B'. Since r was arbitrary then the row space of A' is contained in the row space of B'. Interchanging the roles of A' and B' to obtain that the row space of B' is contained in the row space of A'. Hence, A' and B' have the same row space. Repeating the argument earlier, we see that the leading 1 in the first row of A' and the leading 1 in the first row of B' have the same position. Now, the theorem follows by induction ■

Theorem 68
Reduced row-echelon matrices have the same row space if and only if they have the same nonzero rows. Hence, every matrix is row equivalent to a unique reduced row-echelon matrix.

Proof.
Let $A = (a_{ij})$ and $B = (b_{ij})$ be two reduced row-echelon matrices. If A and B have the same nonzero rows then they have the same row space.
Conversely, suppose that A and B have the same row space. Let $r_1, r_2, \cdots, r_s$ be the nonzero rows of B. Let R_i be the ith nonzero row of A. Then there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_s$ such that

$$R_i = \alpha_1 r_1 + \alpha_2 r_2 + \cdots + \alpha_s r_s. \tag{4.3}$$

We will show that $R_i = r_i$ or equivalently $\alpha_i = 1$ and $\alpha_k = 0$ for $k \neq i$. Suppose that the leading 1 of R_i occurs at the j_i column. Then by (4.3) we have

$$1 = \alpha_1 b_{1j_i} + \alpha_2 b_{2j_i} + \cdots + \alpha_s b_{sj_i}. \tag{4.4}$$

By Theorem 67, $b_{ij_i} = 1$ and $b_{kj_i} = 0$ for $k \neq i$. Hence, $\alpha_i = 1$.
It remains to show that $\alpha_k = 0$ for $k \neq i$. So suppose that $k \neq i$. Suppose the leading 1 of R_k occurs in the j_kth column. By (4.3) we have

$$a_{ij_k} = \alpha_1 b_{1j_k} + \alpha_2 b_{2j_k} + \cdots + \alpha_s b_{sj_k}. \tag{4.5}$$

Since B is in reduced row-echelon form then $b_{kj_k} = 1$ and $b_{ij_k} = 0$ for $i \neq k$. Hence, $\alpha_k = a_{ij_k}$. Since, the leading 1 of R_k occurs at the j_k column then by Theorem 67 we have $a_{kj_k} = 1$ and $a_{ij_k} = 0$ for $i \neq k$. Hence, $\alpha_k = 0$, for $k \neq i$. ■

Remark
A procedure for finding a basis for the column space of a matrix A is described as follows: From A we construct the matrix A^T. Since the rows of A^T are the columns of A, then the row space of A^T and the column space of A are identical. Thus, a basis for the row space of A^T will yield a basis for the column space of A. We transform A^T into a row-reduced echelon matrix B. The nonzero rows of B form a basis for the row space of A^T. Finally, the nonzero columns of B^T form a basis for the column space of A.

Exercise 278
Find a basis for the column space of the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 2 \\ -1 & -2 & 1 & 3 & 0 & -1 \\ 2 & 4 & 1 & 3 & 1 & 9 \\ 1 & 2 & 1 & 3 & 0 & 3 \end{pmatrix}$$

Solution.
The columns of A are the rows of the transpose matrix.

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & -2 & 4 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 3 & 3 \\ 0 & 0 & 1 & 0 \\ 2 & -1 & 9 & 3 \end{pmatrix}$$

The reduction of this matrix to row-echelon form is as follows.

Step 1: $r_2 \leftarrow r_2 - 2r_1$ and $r_6 \leftarrow r_6 - 2r_1$

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

Step 2: $r_2 \leftrightarrow r_6$

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

Step 3: $r_3 \leftarrow r_3 - r_2$ and $r_4 \leftarrow r_4 - 3r_2$

$$\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & 5 & 1 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & -12 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Step 4: $r_4 \leftarrow r_4 + 12r_5$ and $r_3 \leftarrow -\frac{1}{4}$

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

Step 5: $r_5 \leftarrow r_5 - r_3$

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

thusS $\{(1,-1,2,1),(0,1,5,1),(0,0,1,0)\}$ is a basis for the row space of A^T, and

$$\left\{\left(\begin{array}{c} 1 \\ -1 \\ 2 \\ 1 \end{array}\right),\left(\begin{array}{c} 0 \\ 1 \\ 5 \\ 1 \end{array}\right),\left(\begin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array}\right)\right\}$$

is a basis for the column space of A ■

Next, we shall establish the relationship between the dimensions of the row space and the column space of a matrix.

Theorem 69
If A is any $m \times n$ matrix, then the row space and the column space of A have the same dimension.

Proof.
Let $r_1, r_2, \ldots, r_m$ be the row vectors of A where

$$r_i = (a_{i1}, a_{i2}, \ldots, a_{in}).$$

Let r be the dimension of the row space. Suppose that $\{v_1, v_2, \ldots, v_r\}$ is a basis for the row space of A where

$$v_i = (b_{i1}, b_{i2}, \ldots, b_{in}).$$

By definition of a basis we have

$$\begin{array}{ccccccc} r_1 & = & c_{11}v_1 & + & c_{12}v_2 & + & \ldots & c_{1r}v_r \\ r_2 & = & c_{21}v_1 & + & c_{22}v_2 & + & \ldots & c_{2r}v_r \\ \vdots & & & & & & & \vdots \\ r_m & = & c_{m1}v_1 & + & c_{m2}v_2 & + & \ldots & c_{mr}v_r. \end{array}$$

Equating entries of these vector equations to obtain

$$\begin{array}{ccccccc} a_{1j} & = & c_{11}b_{1j} & + & c_{12}b_{2j} & + & \dots & c_{1r}b_{rj} \\ a_{2j} & = & c_{21}b_{ij} & + & c_{22}b_{2j} & + & \dots & c_{2r}b_{rj} \\ \vdots & & & & & & & \vdots \\ a_{mj} & = & c_{m1}b_{1j} & + & c_{m2}b_{2j} & + & \dots & c_{mr}b_{rj}. \end{array}$$

or equivalently

$$\left(\begin{array}{c} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{array}\right) = b_{1j}\left(\begin{array}{c} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{array}\right) + b_{2j}\left(\begin{array}{c} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{array}\right) + \dots + b_{rj}\left(\begin{array}{c} c_{1r} \\ c_{2r} \\ \vdots \\ c_{mr} \end{array}\right)$$

where $1 \le j \le n$. It follows that the columns of A are linear combinations of r vectors and consequently the dimension of the column space is less than or equal to the dimension of the row space. Now interchanging row and column we get that the dimension of the row space is less than or equal to the dimension of the column space. ■

The dimension of the row space or the column space of a matrix A is called the **rank** of A and is denoted by $rank(A)$. It follows from Theorem 69 that $rank(A) = rank(A^T)$. Also, by Theorem 65 the rank of A is equal to the number of nonzero rows in the row-echelon matrix equivalent to A.

Remark
Recall that in Section 1.6 it was asserted that no matter how a matrix is reduced (by row operations) to a matrix R in row-echelon form, the number of nonzero rows of R is always the same (and was called the rank of A.) Theorem 69 shows that this assertion is true and that the two notions of rank agree.

Exercise 279
Find the rank of the matrix

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

Solution.
The transpose of the matrix A is the matrix

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

The reduction of this matrix to row-echelon form is as follows.

Step 1: $r_3 \leftarrow r_3 - r_1$ and $r_4 \leftarrow r_4 - r_1$

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

Step 2: $r_4 \leftarrow r_4 + r_2$ and $r_3 \leftarrow r_3 - r_2$

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

Step 3: $r_2 \leftarrow \frac{1}{2}r_2$

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

Thus $\{(1,3,0),(0,1,2)\}$ form a basis for the row space of A^T or equivalently

$$\left\{ \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\}$$

form a basis for the column space of A and hence $rank(A) = 2$ ■

Theorem 70
If A is an $m \times n$ matrix U and V invertible matrices of size $m \times m$ and $n \times n$ respectively. Then $rank(A) = rank(UA) = rank(AV)$.

Proof.
Follows from the definition of rank and Theorem 66 ■

The following theorem shades information about the invertibility of a square matrix as well as the question of existence of solutions to linear systems.

Theorem 71
If A is an $n \times n$ matrix then the following statements are equivalent

(a) A is row equivalent to I_n.
(b) $rank(A) = n$.
(c) The rows of A are linearly independent.
(d) The columns of A are linearly independent.

Proof.
$(a) \Rightarrow (b)$: Since A is row equivalent to I_n and I_n is a matrix in reduced echelon

form then the rows of I_n are linearly independent . Since the row space of A is the span of n vectors then the row space is n-dimensional vector space and consequently $rank(A) = n$.

$(b) \Rightarrow (c)$: Since $rank(A) = n$ then the dimension of the row space is n. Hence, by Theorem 58 the rows of A are linearly independent.

$(c) \Rightarrow (d)$: Assume that the row vectors of A are linearly independent then the dimension of row space is n and consequently the dimension of the column space is n. By Theorem 58, the columns of A are linearly independent.

$(d) \Rightarrow (a)$: Assume the columns of A are linearly independent. Then the dimension of the row space of A is n. Thus, the reduced row echelon form of A has n non-zero rows. Since A is a square matrix the reduced echelon matrix must be the identity matrix. Hence A is row equivalent to I_n. ■

Now consider a system of linear equations $Ax = b$ or equivalently

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

This implies

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

or

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Exercise 280
Show that a system of linear equations $Ax = b$ is consistent if and only if b is in the column space of A.

Solution.
If b is in the column space of A then there exist scalars $x_1, x_2, \cdots x_n$ such that

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

But this is equivalent to $Ax = b$. Hence, the system $Ax = b$ is consistent. The converse is similar ■

Let A be an $n \times n$ matrix. The set $ker(A) = \{x \in \mathbb{R}^n : Ax = 0\}$ is a subspace of $\mathbb{R}^n$ called the **nullspace.** The dimension of this vector space is called the **nullity** of A.

Exercise 281
Find a basis for the nullspace of

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

Solution.
By Exercise 38 we found that

$$ker(A) = span\{(-1, 1, 0, 0, 0), (-1, 0, -1, 0, 1)\}$$

Thus, $nullity(A) = 2$ ■

The following theorem establishes an important relationship between the rank and nullity of a matrix.

Theorem 72
For any $m \times n$ matrix we have

$$rank(A) + nullity(A) = n. \tag{4.6}$$

Proof.
Since A has n columns then the system $Ax = 0$ has n unknowns. These fall into two categories: the leading (or dependent) variables and the independent variables. Thus,

$$[number\ of\ leading\ variables] + [number\ of\ independent\ variables] = n.$$

But the number of leading variables is the same as the number of nonzero rows of the reduced row-echelon form of A, and this is just the rank of A. On the other hand, the number of independent variables is the same as the number of parameters in the general solution which is the same as the dimension of $ker(A)$. Now (4.6) follows ■

Exercise 282
Find the number of parameters in the solution set of the system $Ax = 0$ if A is a 5×7 matrix and $rank(A) = 3$.

Solution.
The proof of the foregoing theorem asserts that $rank(A)$ gives the number of leading variables that occur in solving the system $AX = \mathbf{0}$ whereas $nullity(A)$ gives the number of free variables. In our case, $rank(A) = 3$ so by the previous theorem $nullity(A) = 7 - rank(A) = 4$ ■

The following theorem ties together all the major topics we have studied so far.

Theorem 73
If A is an $n \times n$ matrix, then all the following are equivalent.

(a) A is invertible.
(b) $|A| \neq 0$.
(c) $Ax = 0$ has only the trivial solution.
(d) $nullity(A) = 0$.
(e) A is row equivalent to I_n.
(f) $Ax = b$ is consistent for all $n \times 1$ matrix b.
(g) $rank(A) = n$.
(h) The rows of A are linearly independent.
(i) The columns of A are linearly independent.

4.7 Review Problems

Exercise 283
Show that the midpoint of $P(x_1, y_1, z_1)$ and $Q(x_1, y_2, z_2)$ is the point

$$M(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2}).$$

Exercise 284
(a) Let $\vec{n} = (a, b, c)$ be a vector orthogonal to a plane P. Suppose $P_0(x_0, y_0, z_0)$ is a point in the plane. Write the equation of the plane.
(b) Find an equation of the plane passing through the point $(3, -1, 7)$ and perpendicular to the vector $\vec{n} = (4, 2, -5)$.

Exercise 285
Find the equation of the plane through the points $P_1(1, 2, -1), P_2(2, 3, 1)$ and $P_3(3, -1, 2)$.

Exercise 286
Find the parametric equations of the line passing through a point $P_0(x_0, y_0, z_0)$ and parallel to a vector $\vec{v} = (a, b, c)$.

Exercise 287
Compute $< \vec{u}, \vec{v} >$ when $\vec{u} = (2, -1, 3)$ and $\vec{v} = (1, 4, -1)$.

Exercise 288
Compute the angle between the vectors $\vec{u} = (-1, 1, 2)$ *and* $\vec{v} = (2, 1, -1)$.

Exercise 289
For any vectors $\vec{u}$ *and* $\vec{v}$ *we have*

$$||\vec{u} \times \vec{v}||^2 = ||\vec{u}||^2||\vec{v}||^2 - <\vec{u}, \vec{v}>^2 .$$

Exercise 290
Show that $||\vec{u} \times \vec{v}||$ *is the area of the parallelogram with sides* $\vec{u}$ *and* $\vec{v}$.

Exercise 291
Let $\mathbf{P}$ *be the collection of polynomials in the indeterminate* x. *Let* $p(x) = a_0 + a_x + a_2x^2 + \cdots$ *and* $q(x) = b_0 + b_1x + b_2x^2 + c\dots$ *be two polynomials in* $\mathbf{P}$. *Define the operations:*
(a) Addition: $p(x) + q(x) = a_0 + b_0 + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$
(b) Multiplication by a scalar: $\alpha p(x) = \alpha a_0 + (\alpha a_1)x + (\alpha a_2)x^2 + \cdots$.
Show that $\mathbf{P}$ *is a vector space.*

Exercise 292
Define on $\mathbb{R}^2$ *the following operations:*
(i) $(x, y) + (x', y') = (x + x', y + y,);$
(ii) $\alpha(x, y) = (\alpha y, \alpha x)$.
Show that $\mathbb{R}^2$ *with the above operations is not a vector space.*

Exercise 293
Let $U = \{p(x) \in \mathbf{P} : p(3) = 0\}$. *Show that* U *is a subspace of* $\mathbf{P}$.

Exercise 294
Let P_n *denote the collection of all polynomials of degree* n. *Show that* P_n *is a subspace of* $\mathbf{P}$.

Exercise 295
Show that $< f, g >= \int_a^b f(x)g(x)dx$ *is an inner product on the space* $C([a, b])$ *of continuous functions.*

Exercise 296
Show that if u *and* v *are two orthogonal vectors of an inner product space, i.e.* $< u, v >= 0$, *then*

$$||u + v||^2 = ||u||^2 + ||v||^2.$$

Exercise 297
(a) Prove that a line through the origin in $\mathbb{R}^3$ *is a subspace of* $\mathbb{R}^3$ *under the standard operations.*
(b) Prove that a line not through the origin in $\mathbb{R}^3$ *in not a subspace of* $\mathbb{R}^3$.

Exercise 298
Show that the set $S = \{(x, y) : x \leq 0\}$ *is not a vector space of* $\mathbb{R}^2$ *under the usual operations of* $\mathbb{R}^2$.

Exercise 299
Show that the collection $C([a,b])$ of all continuous functions on $[a,b]$ with the operations:

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ (\alpha f)(x) &= \alpha f(x)\end{aligned}$$

is a vector space.

Exercise 300
Let $S = \{(a,b,a+b) : a,b \in \mathbb{R}\}$. Show that S is a subspace of $\mathbb{R}^3$ under the usual operations.

Exercise 301
Let V be a vector space. Show that if $u,v,w \in V$ are such that $u+v = u+w$ then $v = w$.

Exercise 302
Let H and K be subspaces of a vector space V.
(a) The **intersection** *of H and K, denoted by $H \cap K$, is the subset of V that consists of elements that belong to both H and K. Show that $H \cap V$ is a subspace of V.*
(b) The **union** *of H and K, denoted by $H \cup K$, is the susbet of V that consists of all elements that belong to either H or K. Give, an example of two subspaces of V such that $H \cup K$ is not a subspace.*
(c) Show that if $H \subset K$ or $K \subset H$ then $H \cup K$ is a subspace of V.

Exercise 303
Let u and v be vectors in an inner product vector space. Show that

$$||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2).$$

Exercise 304
Let u and v be vectors in an inner product vector space. Show that

$$||u+v||^2 - ||u-v||^2 = 4 < u,v > .$$

Exercise 305
(a) Use Cauchy -Schwarz's inequality to show that if $u = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ and $v = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$ then

$$(x_1y_1 + x_2y_2 + \cdots + x_ny_n)^2 \le (x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_n^2).$$

(b) Use (a) to show

$$n^2 \le (a_1 + a_2 + \cdots + a_n)(\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n})$$

where $a_i > 0$ for $1 \le i \le n$.

Exercise 306
Let $C([0,1])$ be the vector space of continuous functions on $[0,1]$. Define

$$< f, g >= \int_0^1 f(x)g(x)dx.$$

(a) Show that $< .,. >$ is an inner product on $C([0,1])$.
(b) Show that

$$\left[\int_0^1 f(x)g(x)dx\right]^2 \leq \left[\int_0^1 f(x)^2 dx\right]\left[\int_0^1 g(x)^2 dx\right].$$

This inequality is known as **Holder's inequality**.
(c) Show that

$$\left[\int_0^1 (f(x)+g(x))^2 dx\right]^{\frac{1}{2}} \leq \left[\int_0^1 f(x)^2 dx\right]^{\frac{1}{2}} + \left[\int_0^1 g(x)^2 dx\right]^{\frac{1}{2}}.$$

This inequality is known as **Minkowski's inequality**.

Exercise 307
Let $W = span\{v_1, v_2, \cdots, v_n\}$, where $v_1, v_2, \cdots, v_n$ are vectors in V. Show that any subspace U of V containing the vectors $v_1, v_2, \cdots, v_n$ must contain W, i.e. $W \subset U$. That is, W is the smallest subspace of V containing $v_1, v_2, \cdots, v_n$.

Exercise 308
Express the vector $\vec{u} = (-9, -7, -15)$ as a linear combination of the vectors $\vec{v_1} = (2,1,4), \vec{v_2} = (1,-1,3), \vec{v_3} = (3,2,5)$.

Exercise 309
(a) Show that the vectors $\vec{v_1} = (2,2,2), \vec{v_2} = (0,0,3)$, and $\vec{v_3} = (0,1,1)$ span $\mathbb{R}^3$.
(b) Show that the vectors $\vec{v_1} = (2,-1,3), \vec{v_2} = (4,1,2)$, and $\vec{v_3} = (8,-1,8)$ do not span $\mathbb{R}^3$.

Exercise 310
Show that

$$M_{22} = span\{\left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right), \left(\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right), \left(\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right), \left(\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right)\}$$

Exercise 311
(a) Show that the vectors $\vec{v_1} = (2,-1,0,3), \vec{v_2} = (1,2,5,-1)$, and $\vec{v_3} = (7,-1,5,8)$ are linearly dependent.
(b) Show that the vectors $\vec{v_1} = (4,-1,2)$ and $\vec{v_2} = (-4,10,2)$ are linearly independent.

Exercise 312
Show that the $\{u, v\}$ is linearly dependent if and only if one is a scalar multiple of the other.

Exercise 313
Let V be the vector of all real-valued functions with domain $\mathbb{R}$. If f, g, h are twice differentiable functions then we define $w(x)$ by the determinant

$$w(x) = \begin{vmatrix} f(x) & g(x) & h(x) \\ f'(x) & g'(x) & h'(x) \\ f''(x) & g''(x) & h''(x) \end{vmatrix}$$

We call $w(x)$ the **Wronskian** *of f, g, and h. Prove that f, g, and h are linearly independent if and only if $w(x) \neq 0$.*

Exercise 314
Use the Wronskian to show that the functions e^x, xe^x, x^2e^x are linearly independent.

Exercise 315
Show that $\{1+x, 3x+x^2, 2+x-x^2\}$ is linearly independent in P_2.

Exercise 316
Show that $\{1+x, 3x+x^2, 2+x-x^2\}$ is a basis for P_2.

Exercise 317
Find a basis of P_3 containing the linearly independent set $\{1+x, 1+x^2\}$.

Exercise 318
Let $\vec{v_1} = (1,2,1), \vec{v_2} = (2,9,0), \vec{v_3} = (3,3,4)$. Show that $S = \{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ is a basis for $\mathbb{R}^3$.

Exercise 319
Let S be a subset of $\mathbb{R}^n$ with $n+1$ vectors. Is S linearly independent or linearly dependent?

Exercise 320
Find a basis for the vector space M_{22} of 2×2 matrices.

Exercise 321
(a) Let U, W be subspaces of a vector space V. Show that the set $U + W = \{u + w : u \in U \text{ and } w \in W\}$ is a subspace of V.
(b) Let M_{22} be the collection of 2×2 matrices. Let U be the collection of matrices in M_{22} whose second row is zero, and W be the collection of matrices in M_{22} whose second column is zero. Find $U + W$.

Exercise 322
Let $S = \{(1,1)^T, (2,3)^T\}$ and $S' = \{(1,2)^T, (0,1)^T\}$ be two bases of $\mathbb{R}^2$. Let $\vec{u} = (1,5)^T$ and $\vec{v} = (5,4)^T$.

(a) Find the coordinate vectors of $\vec{u}$ and $\vec{v}$ with respect to the basis S.

(b) What is the transition matrix P from S to S'?
(c) Find the coordinate vectors of $\vec{u}$ and $\vec{v}$ with respect to S' using P.
(d) Find the coordinate vectors of $\vec{u}$ and $\vec{v}$ with respect to S' directly.
(e) What is the transition matrix Q from S' to S?
(f) Find the coordinate vectors of $\vec{u}$ and $\vec{v}$ with respect to S using Q.

Exercise 323
Suppose that $S' = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ is a basis of $\mathbb{R}^3$, where $\vec{u_1} = (1,0,1), \vec{u_2} = (1,1,0), \vec{u_3} = (0,0,1)$. Let $S = \{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$. Suppose that the transition matrix from S to S' is

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

Determine S.

Exercise 324
Show that if A is not a square matrix then either the row vectors of A or the column vectors of A are linearly dependent.

Exercise 325
Prove that the row vectors of an $n \times n$ invertible matrix A form a basis for $\mathbb{R}^n$.

Exercise 326
Compute the rank of the matrix

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

and find a basis for the row space of A.

Exercise 327
Let U and W be subspaces of a vector space V. We say that V is the **direct sum** *of U and W if and only if $V = U + W$ and $U \cap W = \{0\}$. We write $V = U \oplus W$. Show that V is the direct sum of U and W if and only if every vector v in V can be written uniquely in the form $v = u + w$.*

Exercise 328
Let V be an inner product space and W is a subspace of V. Let $W^{\perp} = \{u \in V : < u, w >= 0, \forall w \in W\}$.
(a) Show that $W^{\perp}$ is a subspace of V.
(b) Show that $V = W \oplus W^{\perp}$.

Exercise 329
Let U and W be subspaces of a vector space V. Show that $dim(U + W) = dim(U) + dim(W) - dim(U \cap W)$.

Exercise 330
Show that $\cos 2x \in span\{\cos^2 x, \sin^2 x\}$.

Exercise 331
If X *and* Y *are subsets of a vector space* V *and if* $X \subset Y$*, show that* $spanX \subset spanY$.

Exercise 332
Show that the vector space $F(\mathbb{R})$ *of all real-valued functions defined on* $\mathbb{R}$ *is an infinite-dimensional vector space.*

Exercise 333
Show that $\{\sin x, \cos x\}$ *is linearly independent in the vector space* $F([0, 2\pi])$.

Exercise 334
Find a basis for the vector space V *of all* 2×2 *symmetric matrices.*

Exercise 335
Find the dimension of the vector space M_{mn} *of all* $m \times n$ *matrices.*

Exercise 336
Let A *be an* $n \times n$ *matrix. Show that there exist scalars* $a_0, a_1, a_2, \cdots, a_{n^2}$ *not all 0 such that*

$$a_0 I_n + a_1 A + a_2 A^2 + \cdots + a_{n^2} A^{n^2} = \mathbf{0}. \tag{4.7}$$

Exercise 337
Show that if A *is an* $n \times n$ *invertible skew-symmetric matrix then* n *must be even.*

Chapter 5

Eigenvalues and Diagonalization

Eigenvalues and eigenvectors arise in many physical applications such as the study of vibrations, electrical systems, genetics, chemical reactions, quantum mechanics, economics, etc. In this chapter we introduce these two concepts and we show how to find them. Eigenvalues and eigenvectors are used in a diagonalization process of a square matrix that we discuss in Section 5.2

5.1 Eigenvalues and Eigenvectors of a Matrix

If A is an $n \times n$ matrix and x is a nonzero vector in $\mathbb{R}^n$ such that $Ax = \lambda x$ for some real number λ then we call x an **eigenvector** corresponding to the **eigenvalue** λ.

Exercise 338
Show that $x = \left(\begin{array}{c} 1 \\ 2 \end{array}\right)$ *is an eigenvector of the matrix*

$$A = \left(\begin{array}{cc} 3 & 0 \\ 8 & -1 \end{array}\right)$$

corresponding to the eigenvalue $\lambda = 3$.

Solution.
The value $\lambda = 4$ is an eigenvalue of A with eigenvector x since

$$Ax = \left(\begin{array}{cc} 3 & 0 \\ 8 & -1 \end{array}\right)\left(\begin{array}{c} 1 \\ 2 \end{array}\right) = \left(\begin{array}{c} 3 \\ 6 \end{array}\right) = 3x\blacksquare$$

Eigenvalues can be either real numbers or complex numbers. In this book an eigenvalue is always assumed to be a real number.

To find the eigenvalues of a square matrix A we rewrite the equation $Ax = \lambda x$ as

$$Ax = \lambda I_n x$$

or equivalently

$$(\lambda I_n - A)x = 0.$$

For λ to be an eigenvalue, there must be a nonzero solution to the above homogeneous system. But, the above system has a nontrivial solution if and only if the coefficient matrix $(\lambda I_n - A)$ is singular (Exercise 172), that is , if and only if

$$|\lambda I_n - A| = 0.$$

This equation is called the **characteristic equation** of A.

Exercise 339
Find the characteristic equation of the matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{pmatrix}$$

Solution.
The characteristic equation of A is the equation

$$\begin{vmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{vmatrix} = 0$$

That is, the equation: $\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$ ■

Theorem 74
$p(\lambda) = |\lambda I_n - A|$ is a polynomial function in λ of degree n and leading coefficient 1. This is called the **characteristic polynomial** *of A.*

Proof.
One of the elementary product will contain all the entries on the main diagonal,i.e. will be the product $(\lambda - a_{11})((\lambda - a_{22}) \cdots (\lambda - a_{nn})$. This is a polynomial of degree n and leading coefficient 1. Any other elementary product will contain at most $n - 2$ factors of the form $\lambda - a_{ii}$. Thus,

$$p(\lambda) = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \textit{ terms of lower degree} \quad (5.1)$$

This ends a proof of the theorem ■

Exercise 340
Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{pmatrix}$$

Solution.
The characteristic polynomial of A is

$$p(\lambda) = \begin{vmatrix} \lambda - 5 & -8 & -16 \\ -4 & \lambda - 1 & -8 \\ 4 & 4 & \lambda + 11 \end{vmatrix}$$

Expanding this determinant we obtain $p(\lambda) = (\lambda+3)(\lambda^2+2\lambda-3) = (\lambda+3)^2(\lambda-1)$ ■

Exercise 341
Show that the coefficient of λ^{n-1} is the negative of the trace of A.

Solution.
This follows from (5.1) ■

Exercise 342
Show that the constant term in the characteristic polynomial of a matrix A is $(-1)^n|A|$.

Solution.
The constant term of the polynomial $p(\lambda)$ corresponds to $p(0)$. It follows that $p(0) = \textit{constant term} = |-A| = (-1)^n|A|$ ■

Exercise 343
Find the eigenvalues of the matrices

(a)

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

(b)

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

Solution.
(a) The characteristic equation of A is given by

$$\begin{vmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{vmatrix} = 0$$

Expanding the determinant and simplifying, we obtain

$$\lambda^2 - 3\lambda + 2 = 0$$

or

$$(\lambda - 1)(\lambda - 2) = 0.$$

Thus, the eigenvalues of A are $\lambda = 2$ and $\lambda = 1$.

(b) The characteristic equation of the matrix B is

$$\left|\begin{array}{cc} \lambda+2 & 1 \\ -5 & \lambda-2 \end{array}\right| = 0$$

Expanding the determinant and simplifying, we obtain

$$\lambda^2 - 9 = 0$$

and the eigenvalues are $\lambda = \pm 3$ ■

Exercise 344
Find the eigenvalues of the matrix

$$A = \left(\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{array}\right)$$

Solution.
According to Exercise 339 the characteristic equation of A is $\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$. Using the rational root test we find that $\lambda = 4$ is a solution to this equation. Using synthetic division of polynomials we find

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0.$$

The eigenvalues of the matrix A are the solutions to this equation, namely, $\lambda = 4, \lambda = 2 + \sqrt{3}$, and $\lambda = 2 - \sqrt{3}$ ■

Next, we turn to the problem of finding the eigenvectors of a square matrix. Recall that an eigenvector is a nontrivial solution to the matrix equation $(\lambda I_n - A)x = 0$.

Theorem 75
The set of eigenvectors of a matrix corresponding to an eigenvalue λ, together with the zero vector, is a subspace of $\mathbb{R}^n$. This subspace is called the **eigenspace** *of A corresponding to λ and will be denoted by V_λ.*

Proof.
Let $V_\lambda = \{x \in \mathbb{R}^n : Ax = \lambda x\}$. We will show that $V_\lambda \cup \{0\}$ is a subspace of $\mathbb{R}^n$. Let $v, w \in V_\lambda$ and $\alpha \in \mathbb{R}$. Then $A(v + w) = Av + Aw = \lambda v + \lambda w = \lambda(v + w)$. That is $v + w \in V_\lambda$. Also, $A(\alpha v) = \alpha A v = \lambda(\alpha v)$ so $\alpha v \in V_\lambda$. Hence, V_λ is a subspace of $\mathbb{R}^n$. ■

By the above theorem, determining the eigenspaces of a square matrix is reduced to two problems: First find the eigenvalues of the matrix, and then find the corresponding eigenvectors which are solutions to linear homogeneous systems.

Exercise 345
Find the eigenspaces of the matrix

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

Solution.
The characteristic equation of the matrix A is

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

Expanding the determinant and simplifying we obtain

$$(\lambda-5)^2(\lambda-1)=0.$$

The eigenvalues of A are $\lambda = 5$ and $\lambda = 1$.
A vector $x = (x_1, x_2, x_3)^T$ is an eigenvector corresponding to an eigenvalue λ if and only if x is a solution to the homogeneous system

$$\left\{\begin{array}{ccccccc} (\lambda-3)x_1 & + & 2x_2 & & & = & 0 \\ 2x_1 & + & (\lambda-3)x_2 & & & = & 0 \\ & & & & (\lambda-5)x_3 & = & 0 \end{array}\right. \tag{5.2}$$

If $\lambda = 1$, then (5.2) becomes

$$\left\{\begin{array}{ccccccc} -2x_1 & + & 2x_2 & & & = & 0 \\ 2x_1 & - & 2x_2 & & & = & 0 \\ & & & - & 4x_3 & = & 0 \end{array}\right. \tag{5.3}$$

Solving this system yields

$$x_1 = s, x_2 = s, x_3 = 0$$

The eigenspace corresponding to $\lambda = 1$ is

$$V_1 = \left\{ \begin{pmatrix} s \\ s \\ 0 \end{pmatrix} : s \in \mathbb{R} \right\} = span \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

If $\lambda = 5$, then (5.2) becomes

$$\left\{\begin{array}{ccccc} 2x_1 & + & 2x_2 & & = 0 \\ 2x_1 & + & 2x_2 & & = 0 \\ & & & 0x_3 & = 0 \end{array}\right. \tag{5.4}$$

Solving this system yields

$$x_1 = -t, x_2 = t, x_3 = s$$

The eigenspace corresponding to $\lambda = 5$ is

$$\begin{aligned} V_5 &= \left\{ \begin{pmatrix} -t \\ t \\ s \end{pmatrix} : s \in \mathbb{R} \right\} \\ &= \left\{ t \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} : s, t \in \mathbb{R} \right\} \\ &= span \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \blacksquare \end{aligned}$$

Exercise 346
Find bases for the eigenspaces of the matrix

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

Solution.
The characteristic equation of the matrix A is

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

Expanding the determinant and simplifying we obtain

$$(\lambda - 2)^2(\lambda - 1) = 0.$$

The eigenvalues of A are $\lambda = 2$ and $\lambda = 1$.
A vector $x = (x_1, x_2, x_3)^T$ is an eigenvector corresponding to an eigenvalue λ if and only if x is a solution to the homogeneous system

$$\left\{ \begin{array}{rcrcrcl} \lambda x_1 & & & + & 2x_3 & = & 0 \\ -x_1 & + & (\lambda - 2)x_2 & - & x_3 & = & 0 \\ -x_1 & & & + & (\lambda - 3)x_3 & = & 0 \end{array} \right. \qquad (5.5)$$

If $\lambda = 1$, then (5.5) becomes

$$\left\{ \begin{array}{rcrcrcl} x_1 & & & + & 2x_3 & = & 0 \\ -x_1 & - & x_2 & - & x_3 & = & 0 \\ -x_1 & & & - & 2x_3 & = & 0 \end{array} \right. \qquad (5.6)$$

Solving this system yields

$$x_1 = -2s, x_2 = s, x_3 = s$$

The eigenspace corresponding to $\lambda = 1$ is

$$V_1 = \left\{ \begin{pmatrix} -2s \\ s \\ s \end{pmatrix} : s \in \mathbb{R} \right\} = span \left\{ \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}$$

and $(-2,1,1)^T$ is a basis for V_1.

If $\lambda = 2$, then (5.5) becomes

$$\left\{\begin{array}{rcrcl} 2x_1 & + & 2x_3 & = & 0 \\ -x_1 & - & x_3 & = & 0 \\ -x_1 & - & x_3 & = & 0 \end{array}\right. \quad (5.7)$$

Solving this system yields

$$x_1 = -s, x_2 = t, x_3 = s$$

The eigenspace corresponding to $\lambda = 2$ is

$$\begin{aligned} V_2 &= \left\{\left(\begin{array}{c} -s \\ t \\ s \end{array}\right) : s \in \mathbb{R}\right\} \\ &= \left\{s\left(\begin{array}{c} -1 \\ 0 \\ 1 \end{array}\right) + t\left(\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right) : s, t \in \mathbb{R}\right\} \\ &= span\left\{\left(\begin{array}{c} -1 \\ 0 \\ 1 \end{array}\right), \left(\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right)\right\} \end{aligned}$$

One can easily check that the vectors $(-1,0,1)^T$ and $(0,1,0)^T$ are linearly independent and therefore these vectors form a basis for V_2 ■

Exercise 347
Show that $\lambda = 0$ is an eigenvalue of a matrix A if and only if A is singular.

Solution.
If $\lambda = 0$ is an eigenvalue of A then it must satisfy $|0I_n - A| = |-A| = 0$. That is $|A| = 0$ and this implies that A is singular. Conversely, if A is singular then $0 = |A| = |0I_n - A|$ and therefore 0 is an eigenvalue of A ■

Exercise 348
(a) Show that the eigenvalues of a triangular matrix are the entries on the main diagonal.
(b) Find the eigenvalues of the matrix

$$A = \left(\begin{array}{ccc} \frac{1}{2} & 0 & 0 \\ -1 & \frac{2}{3} & 0 \\ 5 & -8 & -\frac{1}{4} \end{array}\right)$$

Solution.
(a) Suppose that A is upper triangular $n \times n$ matrix. Then the matrix $\lambda I_n - A$ is also upper triangular with entries on the main diagonal are $\lambda - a_{11}, \lambda -$

$a_{22}, \cdots, \lambda a_{nn}$. Since the determinant of a triangular matrix is just the product of the entries of the main diaginal then the characteristic equation of A is

$$(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) = 0.$$

Hence, the eigenvalues of A are $a_{11}, a_{22}, \cdots, a_{nn}$.
(b) Using (a), the eigenvalues of A are $\lambda = \frac{1}{2}, \lambda = \frac{2}{3}$, and $\lambda = -\frac{1}{4}$ ■

Exercise 349
Show that A and A^T have the same characteristic polynomial and hence the same eigenvalues.

Solution.
We use the fact that a matrix and its transpose have the same determinant (Theorem 28). Hence,

$$|\lambda I_n - A^T| = |(\lambda I_n - A)^T| = |\lambda I_n - A|.$$

Thus, A and A^T have the same characteristic equation and therefore the same eigenvalues ■

The **algebraic multiplicity** of an eigenvalue λ of a matrix A is the multiplicity of λ as a root of the characteristic polynomial, and the dimension of the eigenspace corresponding to λ is called the **geometric multiplicity** of λ.

Exercise 350
Find the algebraic and the geometric multiplicity of the eigenvalues of the matrix

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

Solution.
The characteristic equation of the matrix A is

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

Expanding the determinant and simplifying we obtain

$$(\lambda - 2)^2(\lambda - 1) = 0.$$

The eigenvalues of A are $\lambda = 2$ (of algebraic multiplicity 2) and $\lambda = 1$ (of algebraic multiplicity 1).
A vector $x = (x_1, x_2, x_3)^T$ is an eigenvector corresponding to an eigenvalue λ if and only if x is a solution to the homogeneous system

$$\left\{ \begin{array}{ccccccc} (\lambda - 2)x_1 & - & x_2 & & & = & 0 \\ & & (\lambda - 2)x_2 & & & = & 0 \\ -2x_1 & - & 3x_2 & + & (\lambda - 1)x_3 & = & 0 \end{array} \right. \quad (5.8)$$

If $\lambda = 1$, then (5.8) becomes

$$\left\{\begin{array}{rcrcl} -x_1 & - & x_2 & = & 0 \\ & - & x_2 & = & 0 \\ -2x_1 & - & 3x_2 & = & 0 \end{array}\right. \tag{5.9}$$

Solving this system yields

$$x_1 = 0, x_2 = 0, x_3 = s$$

The eigenspace corresponding to $\lambda = 1$ is

$$V_1 = \left\{ \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} : s \in \mathbb{R} \right\} = span \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

and $(0, 0, 1)^T$ is a basis for V_1. The geometric multiplicity of $\lambda = 1$ is 1.

If $\lambda = 2$, then (5.8) becomes

$$\left\{\begin{array}{rcrcrcl} & - & x_2 & & & = & 0 \\ -2x_1 & - & 3x_2 & + & x_3 & = & 0 \end{array}\right. \tag{5.10}$$

Solving this system yields

$$x_1 = \frac{1}{2}s, x_2 = 0, x_3 = s$$

The eigenspace corresponding to $\lambda = 2$ is

$$\begin{aligned} V_2 &= \left\{ \begin{pmatrix} \frac{1}{2}s \\ 0 \\ s \end{pmatrix} : s \in \mathbb{R} \right\} \\ &= span \left\{ \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right\} \end{aligned}$$

and the vector $(\frac{1}{2}, 0, 1)^T$ is a basis for V_2 so that the geometric multiplicity of $\lambda = 2$ is 1 ■

There are many matrices with real entries but with no real eigenvalues. An example is given in the next exercise.

Exercise 351
Show that the following matrix has no real eigenvalues.

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

Solution.
The characteristic equation of the matrix A is

$$\begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix}$$

Expanding the determinant we obtain

$$\lambda^2 + 1 = 0.$$

The solutions to this equation are the imaginary complex numbers $\lambda = i$ and $\lambda = -i$, and since we are assuming in this chapter that all our scalars are real numbers, A has no real eigenvalues ■

Symmetric matrices with real entries have always real eigenvalues. In order to prove this statement we recall the reader of the following definition. A **complex** number is any number of the form $z = a + bi$ where $i = \sqrt{-1}$ is the **imaginary root.** The number a is called the **real part** of z and b is called the **imaginary part**. Also, recall that $a + bi = a' + b'i$ if and only if $a = a'$ and $b = b'$.

Theorem 76
If $\lambda = a + bi$ is an eigenvalue of a real symmetric $n \times n$ matrix A then $b = 0$, that is λ is real.

Proof.
Since $\lambda = a + bi$ is an eigenvalue of A then λ satisfies the equation

$$[(a + bi)I_n - A](u + vi) = 0 + 0i,$$

where u, v, are vectors in $\mathbb{R}^n$ not both equal to zero and 0 is the zero vector of $\mathbb{R}^n$. The above equation can be rewritten in the following form

$$(aI_nu - Au - bI_nv) + (aI_nv + bI_nu - Av)i = 0 + 0i$$

Setting the real and imaginary parts equal to zero, we have

$$aI_nu - Au - bI_nv = 0$$

and

$$aI_nv + bI_nu - Av = 0.$$

These equations yield the following equalities

$$< v, aI_nu - Au - bI_nv >=< v, 0 >= 0$$

and

$$< aI_nv + bI_nu - Av, u >=< 0, u >= 0$$

or equivalently

$$\begin{aligned} a < v, I_n u > - < v, Au > -b < v, I_n v > &= 0 \quad (5.11) \\ a < I_n v, u > - < Av, u > +b < I_n u, u > &= 0. \quad (5.12) \end{aligned}$$

But $< I_n v, u >= u^T(I_n v) = (I_n u)^T v =< v, I_n u >$ and $< Av, u >= u^T(Av) = u^T(A^T v) = (u^T A^t)v = (Au)^T v =< v, Au >$, where $< .,. >$ is the inner product in$\mathbb{R}^n$. Subtracting the two equations in (5.11), we now get

$$-b < v, I_n v > -b < I_n u, u >= 0$$

or

$$-b[< u, u > + < v, v >] = 0.$$

Since u and v not both the zero vector, then either $< u, u >> 0$ or $< v, v >> 0$. From the previous equation we conclude that $b = 0$. ■

We next introduce a concept for square matrices that will be fundamental in the next section. We say that two $n \times n$ matrices A and B are similar if there exists a nonsingular matrix P such that $B = P^{-1}AP$. We write $A \sim B$. The matrix P is not unique. For example, if $A = B = I_n$ then any invertible matrix P will satisfy the definition.

Exercise 352
Show that $\sim$ is an equivalence relation on the collection M_{nn} of square matrices. That is, show the following:
(a) $A \sim A$ ($\sim$ is reflexive).
(b) If $A \sim B$ then $B \sim A$ ($\sim$ is symmetric).
(c) If $A \sim B$ and $B \sim C$ then $A \sim C$ ($\sim$ is transitive).

Solution.
(a) Since $A = I_n^{-1}AI_n$ then $A \sim A$.
(b) Suppose that $A \sim B$. Then there is an invertible matrix P such that $B = P^{-1}AP$. By premultiplying by P and postmultiplying by P^{-1} we find $A = PBP^{-1} = Q^{-1}BQ$, where $Q = P^{-1}$. Hence, $B \sim A$.
(c) Suppose that $A \sim B$ and $B \sim C$ then there exist invertible matric P and Q such that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. It follows that $C = Q^{-1}BQ = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ) = R^{-1}AR$, with $R = PQ$. This says that $A \sim C$ ■

Exercise 353
Let A and B be similar matrices. Show the following:
(a) $|A| = |B|$.
(b) $tr(A) = tr(B)$.
(c) $rank(A) = rank(B)$.
(d) $|\lambda I_n - A| = |\lambda I_n - B|$.

Solution.
Since $A \sim B$ then there exists an invertible matrix P such that $B = P^{-1}AP$.

(a) $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A|$ since $|P^{-1}| = |P|^{-1}$.
(b) $tr(B) = tr(P^{-1}AP) = tr(PP^{-1}A) = tr(A)$ (See Exercise 88 (a)).
(c) By Theorem 70, we have $rank(B) = rank(PA) = rank(A)$.
(d) Indeed, $|\lambda I_n - B| = |\lambda I_n - P^{-1}AP| = |P^{-1}(\lambda I_n)P| = |\lambda I_n - A|$. It follows that two similar matrices have the same eigenvalues ■

Exercise 354
Show that the following matrices are not similar.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Solution.
The eigenvalues of A are $\lambda = 3$ and $\lambda = -1$. The eigenvalues of B are $\lambda = 0$ and $\lambda = 2$. According to Exercise 353 (d), these two matrices cannot be similar ■

Exercise 355
Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ including repetitions. Show the following.

(a) $tr(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
(b) $|A| = \lambda_1\lambda_2\cdots\lambda_n$.

Solution.
Factoring the characteristic polynomial of A we find

$$\begin{array}{rcl} p(\lambda) & = & (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\ & = & \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + \lambda_1\lambda_2\cdots\lambda_n \end{array}$$

(a) By Exercise 341, $tr(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.
(b) $|A| = p(0) = \lambda_1\lambda_2\cdots\lambda_n$ ■

We end this section with the following important result of linear algebra.

Theorem 77 *(Cayley-Hamilton)*
Every square matrix is the zero of its characteristic polynomial.

Proof.
Let $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial corresponding to A. We will show that $p(A) = 0$. The cofactors of A are polynomials in λ of degree at most $n-1$. Thus,

$$adj(\lambda I_n - A) = B_{n-1}\lambda^{n-1} + \cdots + B_1\lambda + B_0$$

where the B_i are $n \times n$ matrices with entries independent of λ. Hence,

$$(\lambda I_n - A)adj(\lambda I_n - A) = |\lambda I_n - A|I_n.$$

That is

$$B_{n-1}\lambda^n + (B_{n-2} - AB_{n-1})\lambda^{n-1} + (B_{n-3} \\ -AB_{n-2})\lambda^{n-2} + \cdots + (B_0 - AB_1)\lambda - AB_0 = \\ I_n\lambda^{n-1} + a_{n-1}I_n\lambda^{n-1} + a_{n-2}I_n\lambda^{n-2} + \cdots a_1I_n\lambda + a_0I_n.$$

Equating coefficients of corresponding powers of λ,

$$\begin{array}{ccc} B_{n-1} & = & I_n \\ B_{n-2} - AB_{n-1} & = & a_{n-1}I_n \\ B_{n-3} - AB_{n-2} & = & a_{n-2}I_n \\ \cdots\cdots\cdots & & \\ B_0 - AB_1 & = & a_1I_n \\ -AB_0 & = & a_0I_n \end{array}$$

Multiplying the above matrix equations by $A^n, A^{n-1}, \cdots, A, I_n$ respectively and adding the resulting matrix equations to obtain

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I_n.$$

In other words, $p(A) = 0$ ■

Exercise 356
Let A be an $n \times n$ matrix whose characteristic polynomial is $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$. Show that if A is invertible then its inverse can be found by means of the formula

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I_n)$$

Solution.
By the Cayley-Hamilton theorem $p(A) = \mathbf{0}$. that is

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I_n = \mathbf{0}.$$

Multiplying both sides by A^{-1} to obtain

$$a_0A^{-1} = -A^{n-1} - a_{n-1}A^{n-2} - \cdots - a_1I_n.$$

Since A is invertible then $|A| = a_0 \neq 0$. Divide both sides of the last equality by a_0 to obtain

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I_n)$$ ■

5.2 Diagonalization of a Matrix

In this section we shall discuss a method for finding a basis of $\mathbb{R}^n$ consisting of eigenvectors of a given $n \times n$ matrix A. It turns out that this is equivalent to finding an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. The latter statement suggests the following terminology.
A square matrix A is called **diagonalizable** if A is similar to a diagonal matrix. That is, there exists an invertible matrix P such that $P^{-1}AP = D$ is a diagonal matrix. The next theorem gives a characterization of diagonalizable matrices and tells how to construct a suitable characterization. In fact, it supports our statement mentioned at the begining of this section that the problem of finding a basis of $\mathbb{R}^n$ consisting of eigenvectors of A is equivalent to diagonalizing A.

Theorem 78
If A is an $n \times n$ square matrix, then the following statements are all equivalent.
(a) A is diagonalizable.
(b) A has n linearly independent eigenvectors.

Proof.
$(a) \Rightarrow (b)$: Suppose A is diagonalizable. Then there are an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. That is

$$AP = PD = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

For $1 \le i \le n$, let

$$p_i = \begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$$

Then the columns of AP are $\lambda_1 p_1, \lambda_2 p_2, \ldots, \lambda_n p_n$. But $AP = [Ap_1, Ap_2, \ldots, Ap_n]$. Hence, $Ap_i = \lambda_i p_i$, for $1 \le i \le n$. Since P is invertible its column vectors are linearly independent and hence are all nonzero vectors (Theorem 71). Thus $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A and $p_1, p_2, \ldots, p_n$ are the corresponding eigenvectors. Hence, A has n linearly independent eigenvectors.

$(b) \Rightarrow (a)$: Suppose that A has n linearly independent eigenvectors $p_1, p_2, \ldots, p_n$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

Then the columns of AP are $Ap_1, Ap_2, \ldots, Ap_n$. But $Ap_1 = \lambda_1 p_1, Ap_2 = \lambda_2 p_2, \ldots, Ap_n = \lambda_n p_n$. Hence

$$\begin{array}{rcl} AP & = & \left(\begin{array}{cccc} \lambda_1 p_{11} & \lambda_2 p_{12} & \ldots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \ldots & \lambda_n p_{2n} \\ \vdots & & & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \ldots & \lambda_n p_{nn} \end{array}\right) \\ & = & \left(\begin{array}{cccc} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \vdots & & & \vdots \\ p_{n1} & p_{n2} & \ldots & p_{nn} \end{array}\right) \left(\begin{array}{ccccc} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{array}\right) = PD \end{array}$$

Since the column vectors of P ar linearly independent, P is invertible. Hence $A = PDP^{-1}$, that is A is similar to a diagonal matrix ∎

From the above proof we obtain the following procedure for diagonalizing a diagonalizable matrix.

Step 1. Find n linearly independent eigenvectors of A, say $p_1, p_2, \cdots, p_n$.
Step 2. Form the matrix P having $p_1, p_2, \cdots, p_n$ as its column vectors.
Step 3. The matrix $P^{-1}AP$ will then be diagonal with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its diagonal entries, where λ_i is the eigenvalue corresponding to p_i, $1 \le i \le n$.

Exercise 357
Find a matrix P that diagonalizes

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

Solution.
From Exercise 345 of the previous section the eigenspaces corresponding to the eigenvalues $\lambda = 1$ and $\lambda = 5$ are

$$V_1 = \left\{ \left(\begin{array}{c} s \\ s \\ 0 \end{array}\right) : s \in \mathbb{R} \right\} = span \left\{ \left(\begin{array}{c} 1 \\ 1 \\ 0 \end{array}\right) \right\}$$

and

$$\begin{array}{rcl} V_5 & = & \left\{ \left(\begin{array}{c} -t \\ t \\ s \end{array}\right) : s \in \mathbb{R} \right\} \\ & = & \left\{ t \left(\begin{array}{c} -1 \\ 1 \\ 0 \end{array}\right) + s \left(\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right) : s, t \in \mathbb{R} \right\} \\ & = & span \left\{ \left(\begin{array}{c} -1 \\ 1 \\ 0 \end{array}\right), \left(\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right) \right\} \end{array}$$

Let $\vec{v_1} = (1, 1, 0)^T, \vec{v_2} = (-1, 1, 0)$, and $\vec{v_3} = (0, 0, 1)^T$. It is easy to verify that these vectors are linearly independent. The matrices

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

and

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

satisfy $AP = PD$ or $D = P^{-1}AP$ ■

Exercise 358
Show that the matrix

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

is not diagonalizable.

Solution.
The characteristic equation of the matrix A is

$$\begin{pmatrix} \lambda + 3 & -2 \\ 2 & \lambda - 1 \end{pmatrix}$$

Expanding the determinant and simplifying we obtain

$$(\lambda + 1)^2 = 0.$$

The only eigenvalue of A is $\lambda = -1$.
A vector $x = (x_1, x_2)^T$ is an eigenvector corresponding to an eigenvalue λ if and only if x is a solution to the homogeneous system

$$\left\{ \begin{array}{ccccc} (\lambda + 3)x_1 & - & 2x_2 & = & 0 \\ 2x_1 & + & (\lambda - 1)x_2 & = & 0 \end{array} \right. \tag{5.13}$$

If $\lambda = -1$, then (5.13) becomes

$$\left\{ \begin{array}{ccccc} 2x_1 & - & 2x_2 & = & 0 \\ 2x_1 & - & 2x_2 & = & 0 \end{array} \right. \tag{5.14}$$

Solving this system yields $x_1 = s, x_2 = s$. Hence the eigenspace corresponding to $\lambda = -1$ is

$$V_{-1} = \left\{ \begin{pmatrix} s \\ s \end{pmatrix} : s \in \mathbb{R} \right\} = span \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

Since $dim(V_{-1}) = 1$, A does not have two linearly independent eigenvectors and is therefore not diagonalizable ■

In many applications one is concerned only with knowing whether a matrix is diagonalizable without the need of finding the matrix P. In order to establish this result we first need the following theorem.

Theorem 79
If $v_1, v_2, \dots, v_n$ are nonzero eigenvectors that correspond to distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ then the set $\{v_1, v_2, \dots, v_n\}$ is linearly indepedent.

Proof.
The proof is by induction on n. If $n = 1$ then $\{v_1\}$ is linearly independent (Exercise 252). So assume that the vectors $\{v_1, v_2, \cdots, v_{n-1}\}$ are linearly independent. Suppose that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_n v_n = 0. \tag{5.15}$$

Apply A to both sides of (5.15) and using the fact that $Av_i = \lambda_i v_i$ to obtain

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots \alpha_n \lambda_n v_n = 0. \tag{5.16}$$

Now, multiplying (5.15) by λ_n and subtracting the resulting equation from (5.16) we obtain

$$\alpha_1(\lambda_1 - \lambda_n)v_1 + \alpha_2(\lambda_2 - \lambda_n)v_2 + \cdots \alpha_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} = 0. \tag{5.17}$$

By the induction hypothesis, all the coefficients must be zero. Since the λ_i are distinct, i.e. $\lambda_i - \lambda_n \neq 0$ for $i \neq n$ then, $\alpha_1 = \alpha_2 = \cdots = \alpha_{n-1} = 0$. Substituting this into (5.15) to obtain $\alpha_n v_n = 0$, and hence $\alpha_n = 0$. This shows that $\{v_1, v_2, \cdots, v_n\}$ is linearly indepedent. ■

As a consquence of the above theorem we have the following useful result.

Theorem 80
If A is an $n \times n$ matrix with n distinct eigenvalues then A is diagonalizable.

Proof.
If $v_1, v_2, \cdots, v_n$ are nonzero eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \cdots \lambda_n$ then by Theorem 79, $v_1, v_2, \cdots, v_n$ are linearly independent vectors. Thus, A is diagonalizable by Theorem 78■

Exercise 359
Show that the following matrix is diagonalizable.

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

Solution.
The characteristic equation of the matrix A is

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

Expanding the determinant and simplifying we obtain

$$(\lambda - 1)(\lambda - 3)(\lambda + 1) = 0$$

The eigenvalues are $1, 3$ and -1, so A is diagonalizable by Theorem 78 ■

The converse of Theorem 80 is false. That is, an $n \times n$ matrix A may be diagonalizable even if it does not have n distinct eigenvalues.

Exercise 360
Show that the matrix

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

is diagonalizable with only one eigenvalue.

Solution.
The characteristic equation of the matrix A is

$$\begin{pmatrix} \lambda - 3 & 0 \\ 0 & \lambda - 3 \end{pmatrix}$$

Expanding the determinant and simplifying we obtain

$$(\lambda - 3)^2 = 0.$$

The only eigenvalue of A is $\lambda = 3$. By letting $P = I_n$ and $D = A$ we see that $D = P^{-1}AP$, i.e. A is diagonalizable ■

Next we turn our attention to the study of symmetric matrices since they arise in many applications. The interesting fact about symmetric matrices is that they are always diagonalizable. This follows from Theorem 76 and Theorem 80.Moreover, the diagonalizing matrix P has some noteworthy properties, namely, $P^T = P^{-1}$, i.e. P is **orthogonal**. More properties are listed in the next theorem.

Theorem 81
The following are all equivalent:
(a) A is diagonalizable by an orthogonal matrix.
(b) A has an orthonormal set of n eigenvectors.
(c) A is symmetric.

Proof.
(a) $\Rightarrow$ (b): Since A is diagonalizable by an orthogonal matrix P then there exist an orthogonal matrix P and a diagonal matrix D such that $P^{-1}AP = D$. As shown in the proof of Theorem 78, the column vectors of P are the n eigenvectors of A. Since P is orthogonal then these vectors form an orthonormal set according to Theorem 63.

(b) $\Rightarrow$ (c): Suppose that A has an orthonormal set of n eigenvectors of A. As shown in the proof of Theorem 78 the matrix P with columns the eigenvectors of A diagonalizes A and $P^{-1}AP = D$. Since the columns of P are orthonormal then $PP^T = I_n$. By Theorem 20 $P^{-1} = P^T$. Thus, $A^T = (PDP^{-1})^T = (P^{-1})^T D^T P^T = (P^T)^{-1} D^T P^T = PDP^{-1} = A$. This says that A is symmetric.
(c) $\Rightarrow$ (a): Suppose that A is symmetric. We proceed by induction on n. If $n = 1$, A is already doagonal. Suppose that (c) implies (a) for all $(n-1) \times (n-1)$ symmetric matrices. Let λ_1 be a real eigenvalue of A and let x_1 be a corresponding eigenvector with $||x_1|| = 1$. Using the Gram-schmidt algorithm we can extend x_1 to a orthonormal basis $\{x_1, x_2, \cdots, x_n\}$ basis of $\mathbb{R}^n$. Let P_1 be the matrix whose columns are $x_1, x_2, \cdots, x_n$ respectively. Then P_1 is orthogonal by Theorem 63. Moreover

$$P_1^T A P_1 = \begin{pmatrix} \lambda_1 & X \\ 0 & A_1 \end{pmatrix}$$

Since A is symmetric then $P_1^T AP$ is also symmetric so that $X = 0$ and A_1 is an $(n-1) \times (n-1)$ symmetric matrix. By the induction hypothesis, there exist an orthogonal matrix Q and a diagonal matrix D_1 such that $Q^T A_1 Q = D_1$. Hence, the matrix

$$P_2 = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$$

is orthogonal and

$$\begin{aligned} (P_2P_1)^T A (P_2P_1) &= P_2^T (P_1 A P_1) P_2 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & Q^T \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & D_1 \end{pmatrix} \end{aligned}$$

is diagonal. Let $T = P_1 P_2$. Then one can easily check that $P_1 P_2$ is orthogonal. This ends a proof of the theorem ■

Exercise 361
Let A be an $m \times n$ matrix. Show that the matrix $A^T A$ has an orthonormal set of n eigenvectors.

Solution.
The matrix $A^T A$ is of size $n \times n$ and is symmetric. By Theorem 81, $A^T A$ has an orthonormal set of n eigenvectors ■

We now turn to the problem of finding an orthogonal matrix P to diagonalize a symmetric matrix. The key is the following theorem.

Theorem 82
If A is a symmetric matrix then the eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof.
Let A be an $n \times n$ symmetric matrix and α, β be two distinct eigenvalues of A. Let $v = (v_1, v_2, \cdots, v_n)$ and $w = (w_1, w_2, \cdots, w_n)$ be two eigenvectors corresponding to α and β. We want to show that $< v, w >= 0$. That is, $v_1w_1 + v_2w_2 + \cdots + v_nw_n = 0$. But this is the same thing as the matrix multiplication $v^T w = 0$. Since v is an eigenvector corresponding to α then $Av = \alpha v$. Taking transpose of both sides to obtain $v^T A^T = \alpha v^T$. Since A is symmetric, i.e. $A^T = A$ then $v^T A = \alpha v^T$. Multiply this equation from the right by w to obtian $v^T Aw = \alpha v^T w$. Hence, $\beta v^T w = \alpha v^T w$. That is, $(\alpha - \beta)v^T w = 0$. By Exercise 79 we conclude that $v^T w = 0$. ■

As a consequence of this theorem we obtain the following procedure for diagonalizing a matrix by an orthogonal matrix P.

Step 1. Find a basis for each eigenspace of A.
Step 2. Apply the Gram-Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.
Step 3. Form the matrix P whose columns are the basis vectors constructed in Step 2. This matrix is orthogonal.

Exercise 362
Consider the matrix

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

(a) Show that A is symmetric.
(b) Find an orthogonal matrix P and a diagonal matrix D such that $P^{-1}AP = D$.

Solution.
(a) One can either check that $A^T = A$ or by noticing that mirror images of entries across the main diagonal are equal.
(b) The characteristic equation of the matrix A is

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

Expanding the determinant and simplifying we obtain

$$\lambda(\lambda - 1)(\lambda - 6) = 0$$

The eigenvalues are $0, 1$ and 6. One can eaily check the following

$$V_0 = span \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$$

$$V_1 = span\left\{\left(\begin{array}{c} 2 \\ 1 \\ 0 \end{array}\right)\right\}$$

$$V_9 = span\left\{\left(\begin{array}{c} -1 \\ 2 \\ 5 \end{array}\right)\right\}$$

Let $\vec{v_1} = \frac{1}{\sqrt{6}}(1,-2,1)^T, \vec{v_2} = \frac{1}{\sqrt{5}}(2,1,0)^T$, and $\vec{v_3} = \frac{1}{\sqrt{30}}(-1,2,5)^T$. One can easily check that these vectors form an orthonormal set so that the matrix P with columns the vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}$ is orthogonal and

$$P^{-1}AP = \left(\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{array}\right) = D\blacksquare$$

5.3 Review Problems

Exercise 363
Show that $\lambda = -3$ is an eigenvalue of the matrix

$$A = \left(\begin{array}{ccc} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{array}\right)$$

and then find the corresponding eigenspace V_{-3}.

Exercise 364
Find the eigenspaces of the matrix

$$A = \left(\begin{array}{cc} 3 & 0 \\ 8 & -1 \end{array}\right)$$

Exercise 365
Find the characteristic polynomial, eigenvalues, and eigenspaces of the matrix

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

Exercise 366
Find the bases of the eigenspaces of the matrix

$$A = \left(\begin{array}{ccc} -2 & 0 & 1 \\ -6 & -2 & 0 \\ 19 & 5 & -4 \end{array}\right)$$

Exercise 367
Show that if λ *is a nonzero eigenvalue of an invertible matrix* A *then* $\frac{1}{\lambda}$ *is an eigenvalue of* A^{-1}.

Exercise 368
Show that if λ *is an eigenvalue of a matrix* A *then* λ^m *is an eigenvalue of* A^m *for any positive integer* m.

Exercise 369
(a) Show that if D *is a diagonal matrix then* D^k, *where* k *is a positive integer, is a diagonal matrix whose entries are the entries of* D *raised to the power* k.
(b) Show that if A *is similar to a diagonal matrix* D *then* A^k *is similar to* D^k.

Exercise 370
Show that the identity matrix I_n *has exactly one eigenvalue. Find the corresponding eigenspace.*

Exercise 371
Show that if $A \sim B$ *then*
(a) $A^T \sim B^T$.
(b) $A^{-1} \sim B^{-1}$.

Exercise 372
If A *is invertible show that* $AB \sim BA$ *for all* B.

Exercise 373
Let A *be an* $n \times n$ **nilpotent** *matrix, i.e.* $A^k = \mathbf{0}$ *for some positive ineteger* k.

(a) Show that $\lambda = 0$ *is the only eigenvalue of* A.
(b) Show that $p(\lambda) = \lambda^n$.

Exercise 374
Suppose that A *and* B *are* $n \times n$ *similar matrices and* $B = P^{-1}AP$. *Show that if* λ *is an eigenvalue of* A *with corresponding eigenvector* x *then* λ *is an eigenvalue of* B *with corresponding eigenvector* $P^{-1}x$.

Exercise 375
Let A *be an* $n \times n$ *matrix with* n *odd. Show that* A *has at least one real eigenvalue.*

Exercise 376
Consider the following $n \times n$ *matrix*

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \end{pmatrix}$$

Show that the characterisitc polynomial of A *is given by* $p(\lambda) = \lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$. *Hence, every monic polynomial (i.e. the coefficient of the highest power of* λ *is 1) is the characteristic polynomial of some matrix.* A *is called the* **companion matrix** *of* $p(\lambda)$.

Exercise 377
Find a matrix P that diagonalizes

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

Exercise 378
Show that the matrix A is diagonalizable.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{pmatrix}$$

Exercise 379
Show that the matrix A is not diagonalizable.

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

Exercise 380
Show that if A is diagonalizable then the rank of A is the number of nonzero eigenvalues of A.

Exercise 381
Show that A is diagonalizable if and only if A^T is diagonalizable.

Exercise 382
Show that if A and B are similar then A is diagonalizable if and only if B is diagonalizable.

Exercise 383
Give an example of two diagonalizable matrices A and B such that $A+B$ is not diagonalizable.

Exercise 384
Show that the following are equivalent for a symmetric matrix A.
(a) A is orthogonal.
(b) $A^2 = I_n$.
(c) All eigenvalues of A are ± 1.

Exercise 385
A matrix that we obtain from the identity matrix by writing its rows in a different order is called **permutation matrix**. *Show that every permutation matrix is orthogonal.*

Exercise 386
Let A be an $n \times n$ skew symmetric matrix. Show that
(a) $I_n + A$ is nonsingular.
(b) $P = (I_n - A)(I_n + A)^{-1}$ is orthogonal.

Exercise 387
We call square matrix E a **projection matrix** *if $E^2 = E = E^T$.*

(a) If E is a projection matrix, show that $P = I_n - 2E$ is orthogonal and symmetric.
(b) If P is orthogonal and symmetric, show that $E = \frac{1}{2}(I_n - P)$ is a projection matrix.

Chapter 6

Linear Transformations

In this chapter we shall discuss a special class of functions whose domains and ranges are vector spaces. Such functions are referred to as linear transformations, a concept to be defined in Section 6.1. Linear transformations play an important role in many areas of mathematics, the physical and social sciences, engineering, and economics.

6.1 Definition and Elementary Properties

A **linear transformation** T from a vector space V to a vector space W is a function $T : V \to W$ that satisfies the following two conditions

(i) $T(u + v) = T(u) + T(v)$, for all u, v in V.
(ii) $T(\alpha u) = \alpha T(u)$ for all u in V and scalar α.

If $W = \mathbb{R}$ then we call T a **linear functional** on V.
It is important to keep in mind that the addition in $u + v$ refers to the addition operation in V whereas that in $T(u) + T(v)$ refers to the addition operation in W. Similar remark for the scalar multiplication.

Exercise 388
Show that $T : \mathbb{R}^2 \to \mathbb{R}^3$ *defined by*

$$T(x, y) = (x, x + y, x - y)$$

is a linear transformation.

Solution.
We verify the two conditions of the definition. Given (x_1, y_1) and (x_2, y_2) in

$\mathbb{R}^2$, compute

$$\begin{aligned} T((x_1,y_1)+(x_2,y_2)) &= T(x_1+x_2,y_1+y_2) \\ &= (x_1+x_2, x_1+x_2+y_1+y_2, x_1+x_2-y_1-y_2) \\ &= (x_1,x_1+y_1,x_1-y_1)+(x_2,x_2+y_2,x_2-y_2) \\ &= T(x_1,y_1)+T(x_2,y_2) \end{aligned}$$

This proves the first condition. For the second condition, we let $\alpha \in \mathbb{R}$ and compute

$$T(\alpha(x,y)) = T(\alpha x,\alpha y) = (\alpha x, \alpha x+\alpha y, \alpha x-\alpha y) = \alpha(x,x+y,x-y) = \alpha T(x,y)$$

Hence T is a linear transformation ■

Exercise 389
Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be given by $T(x,y) = (x,y,1)$. Show that T is not linear.

Solution.
We show that the first condition of the definition is violated. Indeed, for any two vectors (x_1,y_1) and (x_2,y_2) we have

$$\begin{aligned} T((x_1,y_1)+(x_2,y_2)) &= T(x_1+x_2,y_1+y_2) = (x_1+x_2,y_1+y_2,1) \\ &\neq (x_1,y_1,1)+(x_2,y_2,1) = T(x_1,y_1)+T(x_2,y_2) \end{aligned}$$

Hence the given transformation is not linear ■

Exercise 390
Show that an $m \times n$ matrix defines a linear transforamtion from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Solution.
Given x and y in $\mathbb{R}^m$ and $\alpha \in \mathbb{R}$, matrix arithmetic yields $T(x+y) = A(x+y) = Ax + Ay = Tx + Ty$ and $T(\alpha x) = A(\alpha x) = \alpha Ax = \alpha Tx$. Thus, T is linear ■

Exercise 391
Consider the matrices

$$E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, F = \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix}, G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

(a) Show that the transformation $T_E(x,y) = (y,x)$ is linear. This transformation is a **reflection** *in the line $y = x$.*
(b) Show that $T_F(x,y) = (\alpha x, y)$ is linear. Such a transformation is called an **expansion** *if $\alpha > 1$ and a* **compression** *if $\alpha < 1$.*
(c) Show that $T_G(x,y) = (x+y,y)$ is linear. This transformation is called a **shear**

Solution.
(a) Given (x_1,y_1) and (x_2,y_2) is $\mathbb{R}^2$ and $\alpha \in \mathbb{R}$ we find

$$\begin{aligned} T_E((x_1,y_1)+(x_2,y_2)) &= T_E(x_1+x_2,y_1+y_2) = (y_1+y_2,x_1+x_2) \\ &= (y_1,x_1)+(y_2,x_2) = T_E(x_1,y_1)+T_E(x_2,y_2) \end{aligned}$$

and

$$T_E(\alpha(x,y)) = T_E(\alpha x, \alpha y) = (\alpha y, \alpha x) = \alpha(y,x) = \alpha T_E(x,y)$$

Hence, T_E is linear.
(b) Given (x_1, y_1) and (x_2, y_2) is $\mathbb{R}^2$ and $\beta \in \mathbb{R}$ we find

$$\begin{array}{rcl} T_F((x_1,y_1)+(x_2,y_2)) &=& T_F(x_1+x_2, y_1+y_2) = (\alpha(x_1+x_2), y_1+y_2) \\ &=& (\alpha x_1, y_1) + (\alpha x_2, y_2) = T_F(x_1,y_1) + T_F(x_2,y_2) \end{array}$$

and

$$T_F(\beta(x,y)) = T_F(\beta x, \beta y) = (\beta\alpha x, \beta y) = \beta(\alpha x, y) = \beta T_F(x,y)$$

Hence, T_F is linear.
(c) Given (x_1, y_1) and (x_2, y_2) is $\mathbb{R}^2$ and $\alpha \in \mathbb{R}$ we find

$$\begin{array}{rcl} T_G((x_1,y_1)+(x_2,y_2)) &=& T_G(x_1+x_2, y_1+y_2) = (x_1+x_2+y_1+y_2, y_1+y_2) \\ &=& (x_1+y_1, y_1) + (x_2+y_2, 2_2) = T_G(x_1,y_1) + T_G(x_2,y_2) \end{array}$$

and

$$T_G(\alpha(x,y)) = T_G(\alpha x, \alpha y) = (\alpha(x+y), \alpha y) = \alpha(x+y, y) = \alpha T_G(x,y)$$

Hence, T_G is linear. ■

Exercise 392
(a) Show that the identity transformation defined by $I(v) = v$ for all $v \in V$ is a linear transformation.
(b) Show that the zero transformation is linear.

Solution.
(a) For all $u, v \in V$ and $\alpha \in \mathbb{R}$ we have $I(u+v) = u+v = Iu + Iv$ and $I(\alpha u) = \alpha u = \alpha I u$. So I is linear.
(b) For all $u, v \in V$ and $\alpha \in \mathbb{R}$ we have $\mathbf{0}(u+v) = \mathbf{0} = \mathbf{0}u + \mathbf{0}v$ and $\mathbf{0}(\alpha u) = \mathbf{0} = \alpha 0 u$. So $\mathbf{0}$ is linear ■

The next theorem collects four useful properties of all linear transformations.

Theorem 83
If $T : V \to W$ is a linear transformation then

(a) $T(0) = 0$
(b) $T(-u) = -T(u)$
(c) $T(u-w) = T(u) - T(w)$
(d) $T(\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n) = \alpha_1 T(u_1) + \alpha_2 T(u_2) + \ldots + \alpha_n T(u_n)$.

Proof.
(a) $T(0) = T(2 \times 0) = 2T(0)$. Thus, $T(0) = 0$.
(b) $T(-u) = T((-1)u) = (-1)T(u) = -T(u)$.
(c) $T(u-w) = T(u+(-w)) = T(u) + T(-w) = T(u) + (-T(w)) = T(u) - T(w)$.

(d) Use induction ■

The following theorem provides a criterion for showing that a transformation is linear.

Theorem 84
A function $T : V \to W$ is linear if and only if $T(\alpha u + v) = \alpha T(u) + T(v)$ for all $u, v \in V$ and $\alpha \in \mathbb{R}$.

Proof.
Suppose first that T is linear. Let $u, v \in V$ and $\alpha \in \mathbb{R}$. Then $\alpha u \in V$. Since T is linear we have $T(\alpha u + v) = T(\alpha u) + T(v) = \alpha T(u) + T(v)$.
Conversely, suppose that $T(\alpha u + v) = \alpha T(u) + T(v)$ for all $u, v \in V$ and $\alpha \in \mathbb{R}$. In particular, letting $\alpha = 1$ we see that $T(u + v) = T(u) + T(v)$ for all $u, v \in V$. Now, letting $v = 0$ we see that $T(\alpha u) = \alpha T(u)$. Thus, T is linear ■

Exercise 393
Let M_{mn} denote the vector space of all $m \times n$ matrices.
(a) Show that $T : M_{mn} \to M_{nm}$ defined by $T(A) = A^T$ is a linear transformation.
(b) Show that $T : M_{nn} \to \mathbb{R}$ defined by $T(A) = tr(A)$ is a linear functional.

Solution.
(a) For any $A, B \in M_{mn}$ and $\alpha \in \mathbb{R}$ we find $T(\alpha A + B) = (\alpha A + B)^T = \alpha A^T + B^T = \alpha T(A) + T(B)$. Hence, T is a linear transformation.
(b) For any $A, B \in M_{nn}$ and $\alpha \in \mathbb{R}$ we have $T(\alpha A + b) = tr(\alpha A + B) = \alpha tr(A) + tr(B) = \alpha T(A) + T(B)$ so T is a linear functional ■

Exercise 394
Let V be an inner product space and v_0 be any fixed vector in V. Let $T : V \to \mathbb{R}$ be the transformation $T(v) =< v, v_0 >$. Show that T is linear functional.

Solution
Indeed, for $u, v \in V$ and $\alpha \in \mathbb{R}$ we find $T(\alpha u + v) =< \alpha u + v, v_0 >=< \alpha u, v_0 > + < v, v_0 >= \alpha < u, v_0 > + < v, v_0 >= \alpha T(u) + T(v)$. Hence, T is a linear functional ■

Exercise 395
Let $\{v_1, v_2, \ldots, v_n\}$ be a basis for a vector space V and let $T : V \to W$ be a linear transformation. Show that if $T(v_1) = T(v_2) = \ldots = T(v_n) = 0$ then $T(v) = 0$ for any vector v in V.

Solution.
Let $v \in V$. Then there exist scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Since T is linear then $T(v) = \alpha T v_1 + \alpha T v_2 + \cdots + \alpha_n T v_n = 0$ ■

Exercise 396
Let $S : V \to W$ and $T : V \to W$ be two linear transformations. Show the following:
(a) $S + T$ and $S - T$ are linear transformations.
(b) αT is a linear transformation where α denotes a scalar.

Solution.
(a) Let $u, v \in V$ and $\alpha \in \mathbb{R}$ then

$$\begin{array}{rcl}(S \pm T)(\alpha u + v) & = & S(\alpha u + v) \pm T(\alpha u + v)\\ & = & \alpha S(u) + S(v) \pm (\alpha T(u) + T(v))\\ & = & \alpha(S(u) \pm T(u)) + (S(v) \pm T(v))\\ & = & \alpha(S \pm T)(u) + (S \pm T)(v)\end{array}$$

(b) Let $u, v \in V$ and $\beta \in \mathbb{R}$ then

$$\begin{array}{rcl}(\alpha T)(\beta u + v) & = & (\alpha T)(\beta u) + (\alpha T)(v)\\ & = & \alpha\beta T(u) + \alpha T(v)\\ & = & \beta(\alpha T(u)) + \alpha T(v)\\ & = & \beta(\alpha T)(u) + (\alpha T)(v)\end{array}$$

Hence, αT is a linear transformation ■

The following theorem shows that two linear transformations defined on V are equal whenever they have the same effect on a basis of the vector space V.

Theorem 85
Let $V = span\{v_1, v_2, \cdots, v_n\}$. If T and S are two linear transformations from V into a vector space W such that $T(v_i) = S(v_i)$ for each i then $T = S$.

Proof.
Let $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in V$. Then

$$\begin{array}{rcl}T(v) & = & T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n)\\ & = & \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n)\\ & = & \alpha_1 S(v_1) + \alpha_2 S(v_2) + \cdots + \alpha_n S(v_n)\\ & = & S(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = S(v)\end{array}$$

Since this is true for any $v \in V$ then $T = S$. ■

The following very useful theorem tells us that once we say what a linear transformation does to a basis for V, then we have completely specified T.

Theorem 86
Let V be an n-dimensional vector space with basis $\{v_1, v_2, \cdots, v_n\}$. If $T : V \longrightarrow W$ is a linear transformation then for any $v \in V$, Tv is completely determined by $\{Tv_1, Tv_2, \cdots, Tv_n\}$.

Proof.
For $v \in V$ we can find scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Since T is linear then $Tv = \alpha_1 Tv_1 + \alpha_2 Tv_2 + \cdots + \alpha_n Tv_n$. This says that Tv is completely determined by the vectors $Tv_1, Tv_2, \cdots, Tv_n$ ■

Theorem 87
Let V and W be two vector spaces and $\{e_1, e_2, \cdots, e_n\}$ be a basis of V. Given any vectors $w_1, w_2, \cdots, w_n$ in W, there exists a unique linear transformation $T : V \longrightarrow W$ such that $T(e_i) = w_i$ for each i.

Proof.
Let $v \in V$. Then v can be represented uniquely in the form $v = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n$. Define $T : V \longrightarrow W$ by $T(v) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n$. Clearly, $T(e_i) = w_i$ for each i. One can easily show that T is linear. Now the uniqueness of T follows from Theorem 85 ■

Exercise 397
Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Write vectors in $\mathbb{R}^n$ as columns. Show that there exists an $m \times n$ matrix A such that $T(x) = Ax$ for all $x \in \mathbb{R}^n$. The matrix A is called the **standard matrix** *of T.*

Solution.
Consider the standard basis of $\mathbb{R}^n$, $\{e_1, e_2, \cdots, e_n\}$. Let $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$. Then $x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$. Thus,

$$T(x) = x_1 T(e_1) + x_2 T(e_2) + \cdots + x_n T(e_n) = Ax$$

where $A = [\ T(e_1) \quad T(e_2) \quad \cdots \quad T(e_n)\ ]$ ■

Exercise 398
Find the standard matrix of $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - 2y + z \\ x - z \end{pmatrix}$$

Solution.
Indeed, by simple inspection one finds that

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} ■$$

6.2 Kernel and Range of a Linear Transformation

In this section we discuss two important subspaces associated with a linear transformation T, namely the kernel of T and the range of T. Also, we discuss some further properties of T as a function such as, the concepts of one-one, onto and the inverse of T. Let $T: V \to W$ be a linear transformation. The **kernel** of T (denoted by $ker(T)$) and the **range** of T (denoted by $R(T)$) are defined by

$$\begin{array}{rcl} ker(T) & = & \{x \in V : T(v) = 0\} \\ R(T) & = & \{w \in W : T(v) = w, v \in V\} \end{array}$$

The following theorem asserts that $ker(T)$ and $R(T)$ are subspaces.

Theorem 88
Let $T: V \to W$ be a linear transformation. Then
(a) $ker(T)$ is a subspace of V.
(b) $R(T)$ is a subspace of W.

Proof.
(a) Let $v_1, v_2 \in ker(T)$ and $\alpha \in \mathbb{R}$. Then $T(\alpha v_1 + v_2) = \alpha T v_1 + T v_2 = 0$. That is, $\alpha v_1 + v_2 \in ker(T)$. This proves that $ker(T)$ is a subspace of V.
(b) Let $w_1, w_2 \in R(T)$. Then there exist $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$. Let $\alpha \in \mathbb{R}$. Then $T(\alpha v_1 + v_2) = \alpha T v_1 + T v_2 = \alpha w_1 + w_2$. Hence, $\alpha w_1 + w_2 \in R(T)$. This shows that $R(T)$ is a subspace of W ■

Exercise 399
If $T: \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x, y, z) = (x - y, z, y - x)$, find $ker(T)$ and $R(T)$.

Solution.
If $(x, y, z) \in ker(T)$ then $(0, 0, 0) = T(x, y, z) = (x - y, z, y - x)$. This leads to the system

$$\left\{ \begin{array}{rcrcrcl} x & - & y & & & = & 0 \\ -x & + & y & & & = & 0 \\ & & & & z & = & 0 \end{array} \right.$$

The general solution is given by $(s, s, 0)$ and therefore $ker(T) = span\{(1, 1, 0)\}$. Now, let $(u, v, w) \in R(T)$ be given. Then there is a vector $(x, y, z) \in \mathbb{R}^3$ such that $T(x, y, z) = (u, v, w)$. This yields the following system

$$\left\{ \begin{array}{rcrcrcl} x & - & y & & & = & u \\ -x & + & y & & & = & w \\ & & & & z & = & v \end{array} \right.$$

and the solution is given by $(u, v, -u)$. Hence,

$$R(T) = span\{(1, 0, -1), (0, 1, 0)\} ■$$

Exercise 400
Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be given by $Tx = Ax$. Find $ker(T)$ and $R(T)$.

Solution.
We have

$$ker(T) = \{x \in \mathbb{R}^n : Ax = \mathbf{0}\} = null(A)$$

and

$$R(T) = \{Ax : x \in \mathbb{R}^n\} \blacksquare$$

Exercise 401
Let V be any vector space and α be a scalar. Let $T : V \to V$ be the transformation defined by $T(v) = \alpha v$.

(a) Show that T is linear.
(b) What is the kernel of T?
(c) What is the range of T?

Solution.
(a) Let $u, v \in V$ and $\beta \in \mathbb{R}$. Then $T(\beta u + v) = \alpha(\beta u + v) = \alpha\beta u + \alpha v = \beta T(u) + T(v)$. Hence, T is linear
(b) If $v \in ker(T)$ then $0 = T(v) = \alpha v$. If $\alpha = 0$ then $ker(T) = V$. If $\alpha \neq 0$ then $ker(T) = \{0\}$.
(c) If $\alpha = 0$ then $R(T) = \{0\}$. If $\alpha \neq 0$ then $R(T) = V$ since $T(\frac{1}{\alpha}v) = v$ for all $v \in V$ $\blacksquare$

Since the kernel and the range of a linear transformation are subspaces of given vector spaces, we may speak of their dimensions. The dimension of the kernel is called the **nullity** of T (denoted $nullity(T)$) and the dimension of the range of T is called the **rank** of T (denoted $rank(T)$).

Exercise 402
Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be given by $T(x, y) = (x, x + y, y)$.

(a) Show that T is linear.
(b) Find $nullity(T)$ and $rank(T)$.

Solution.
(a) Let (x_1, y_1) and (x_2, y_2) be two vectors in $\mathbb{R}^2$. Then for any $\alpha \in \mathbb{R}$ we have

$$\begin{aligned} T(\alpha(x_1, y_1) + (x_2, y_2)) &= T(\alpha x_1 + x_2, \alpha y_1 + y_2) \\ &= (\alpha x_1 + x_2, \alpha x_1 + x_2 + \alpha y_1 + y_2, \alpha y_1 + y_2) \\ &= (\alpha x_1, \alpha(x_1 + y_1), \alpha y_1) + (x_2, x_2 + y_2, y_2) \\ &= \alpha T(x_1, y_1) + T(x_2, y_2) \end{aligned}$$

(b) Let $(x, y) \in ker(T)$. Then $(0, 0, 0) = T(x, y) = (x, x + y, y)$ and this leads to $ker(T) = \{(0, 0)\}$. Hence, $nullity(T) = 0$. Now, let $(u, v, w) \in R(T)$. Then there exists $(x, y) \in \mathbb{R}^2$ such that $(x, x + y, y) = T(x, y) = (u, v, w)$. Hence, $R(T) = \{(x, x + y, y) : x, y \in \mathbb{R}\} = span\{(1, 1, 0), (0, 1, 1)\}$. Thus, $rank(T) = 2$ $\blacksquare$

Since linear transformations are functions then it makes sense to talk about one-one and onto functions. We say that a linear transformation $T : V \to W$ is **one-one** if $Tv = Tw$ implies $v = w$. We say that T is **onto** if $R(T) = W$. If T is both one-one and onto we say that T is an **isomorphism** and the vector spaces V and W are said to be **isomorphic** and we write $V \cong W$. The identity transformation is an isomorphism of any vector space onto itself. That is, if V is a vector space then $V \cong V$.

The following theorem is used as a criterion for proving that a linear transformation is one-one.

Theorem 89
Let $T : V \to W$ be a linear transformation. Then T is one-one if and only if $ker(T) = \{0\}$.

Proof.
Suppose first that T is one-one. Let $v \in ker(T)$. Then $Tv = 0 = T0$. Since T is one-one then $v = 0$. Hence, $ker(T) = \{0\}$.
Conversely, suppose that $ker(T) = \{0\}$. Let $u, v \in V$ be such that $Tu = Tv$, i.e $T(u - v) = 0$. This says that $u - v \in ker(T)$, which implies that $u - v = 0$. Thus, T is one-one ■

Another criterion of showing that a linear transformation is one-one is provided by the following theorem.

Theorem 90
Let $T : V \to W$ be a linear transformation. Then the following are equivalent:
(a) T is one-one.
(b) If S is linearly independent set of vectors then $T(S)$ is also linearly independent.

Proof.
(a) $\Rightarrow$ (b): Follows from Exercise 405 below.
(b) $\Rightarrow$ (a): Suppose that $T(S)$ is linearly independent for any linearly independent set S. Let v be a nonzero vector of V. Since $\{v\}$ is linearly independent then $\{Tv\}$ is linearly independent. That is, $Tv \neq 0$. Hence, T is one-one ■

Exercise 403
consider the transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by $T(x, y, z) = (x + y, x - y)$
(a) Show that T is linear.
(b) Show that T is onto but not one-one.

Solution.
(a) Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be two vectors in $\mathbb{R}^3$ and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} T(\alpha(x_1, y_1, z_1) + (x_2, y_2, z_2)) &= T(\alpha x_1 + x_2, \alpha y_1 + y_2, \alpha z_1 + z_2) \\ &= (\alpha x_1 + x_2 + \alpha y_1 + y_2, \alpha x_1 + x_2 - \alpha y_1 - y_2) \\ &= (\alpha(x_1 + y_1), \alpha(x_1 - y_1)) + (x_2 + y_2, x_2 - y_2) \\ &= \alpha T(x_1, y_1, z_1) + T(x_2, y_2, z_2) \end{aligned}$$

(b) Since $(0,0,1) \in ker(T)$ then by Theorem 89 T is not one-one. Now, let $(u,v,w) \in \mathbb{R}^3$ be such that $T(u,v,w) = (x,y)$. In this case, $x = \frac{1}{2}(u+v)$ and $y = \frac{1}{2}(u-v)$. Hence, $R(T) = \mathbb{R}^3$ so that T is onto ■

Exercise 404
Consider the transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ given by $T(x,y) = (x+y, x-y, x)$
(a) Show that T is linear.
(b) Show that T is one-one but not onto.

Solution.
(a) Let (x_1,y_1) and (x_2,y_2) be two vectors in $\mathbb{R}^2$. Then for any $\alpha \in \mathbb{R}$ we have

$$\begin{aligned} T(\alpha(x_1,y_1)+(x_2,y_2)) &= T(\alpha x_1 + x_2, \alpha y_1 + y_2) \\ &= (\alpha x_1 + x_2 + \alpha y_1 + y_2, \alpha x_1 + x_2 - \alpha y_1 - y_2, \alpha x_1 + x_2) \\ &= (\alpha(x_1+y_1), \alpha(x_1-y_1), \alpha x_1) + (x_2+y_2, x_2-y_2, x_2) \\ &= \alpha T(x_1,y_1) + T(x_2,y_2) \end{aligned}$$

Hence, T is linear.
(b) If $(x,y) \in ker(T)$ then $(0,0,0) = T(x,y) = (x+y, x-y, x)$ and this leads to $(x,y) = (0,0)$. Hence, $ker(T) = \{(0,0)\}$ so that T is one-one. To show that T is not onto, take the vector $(0,0,1) \in \mathbb{R}^3$. Suppose that $(x,y) \in \mathbb{R}^2$ is such that $T(x,y) = (0,0,1)$. This leads to $x = 1$ and $x = 0$ which is impossible. Thus, T is not onto ■

Exercise 405
Let $T : V \to W$ be a one-one linear transformation. Show that if $\{v_1, v_2, \ldots, v_n\}$ is a basis for V then $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for $R(T)$.

Solution.
The fact that $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ is linearly independent follows from Theorem 90. It remains to show that $R(T) = span\{T(v_1), T(v_2), \cdots, T(v_n)\}$. Indeed, let $w \in R(T)$. Then there exists $v \in V$ such that $T(v) = w$. Since $\{v_1, v_2, \cdots, v_n\}$ is a basis of V then v can be written uniquely in the form $v = \alpha v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Hence, $w = T(v) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n)$. That is, $w \in span\{T(v_1), T(v_2), \cdots, T(v_n)\}$. We conclude that $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ is a basis of $R(T)$ ■

Exercise 406
Show that if $T : V \to W$ is a linear transformation such that $dim(ker(T)) = dimV$ then $ker(T) = V$.

Solution.
This follows from Exercise 265 ■

The following important result is called the **dimension theorem**.

Theorem 91
If $T : V \to W$ is a linear transformation with $dim(V) = n$, then

$$nullity(T) + rank(T) = n.$$

Proof.
Let $k = dim(ker(T))$. If $k = 0$ then $ker(T) = \{0\}$ and consequently $ker(T)$ has no basis. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis for V. Then $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for $R(T)$ (Exercise 405). Thus the conclusion holds. Now, if $k = n$ then $ker(T) = V$ (Exercise 406) and consequently $T(v) = 0$ for all v in V. Hence, $R(T) = \{0\}$ and the conclusion holds. So we assume that $0 < k < n$. Let $\{v_1, v_2, \ldots, v_k\}$ be a basis for $ker(T)$. Extend this basis to a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V. We show that $\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$ is a basis for $R(T)$.
Clearly, $span\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$ is contained in $R(T)$. On the other hand, let w be an element of $R(T)$. Then $w = T(v)$ for some v in V. But then $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ and consequently

$$\begin{aligned} w &= \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \\ &= \alpha_{k+1} T(v_{k+1}) + \alpha_{k+2} T(v_{k+2}) + \ldots \alpha_n + T(v_n) \end{aligned}$$

since $T(v_1) = T(v_2) = \ldots = T(v_k) = 0$. It follows that

$$R(T) = span\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}.$$

Now, suppose that

$$\alpha_{k+1} T(v_{k+1}) + \alpha_{k+2} T(v_{k+2}) + \ldots \alpha_n T(v_n) = 0$$

then

$$T(\alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \ldots + \alpha_n v_n) = 0.$$

This implies that the vector $\alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \ldots + \alpha_n v_n$ is in $ker(T)$. Consequently,

$$\alpha_{k+1} v_{k+1} + \alpha_{k+2} v_{k+2} + \ldots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_k v_k.$$

Hence, $\beta_1 = \beta_2 = \ldots = \beta_k = \alpha_{k+1} = \ldots = \alpha_n = 0$. It follows that the vectors $T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)$ are linearly independent and form a basis for $R(T)$. This shows that the conclusion of the theorem holds ■

Remark
According to the above theorem, if $T : V \to W$ is a linear transformation and $\{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ is a basis of V such that $\{v_{k+1}, \cdots, v_n\}$ is a basis of $ker(T)$ then $\{T(v_1), \cdots, T(v_k)\}$ is a basis of $R(T)$.

We have seen that a linear transformation $T : V \to W$ can be one-one and onto, one-one but not onto, and onto but not one-one. The foregoing theorem shows that each of these properties implies the other if the vector spaces V and W have the same dimension.

Theorem 92
Let $T : V \to W$ be a linear transformation such that $dim(V) = dim(W) = n$.

Then
(a) if T is one - one, then T is onto;
(b) if T is onto, then T is one-one.

Proof.
(a) If T is one-one then $ker(T) = \{0\}$. Thus, $dim(ker(T)) = 0$. By Theorem 91 we have $dim(R(T)) = n$. Hence, $R(T) = W$. That is, T is onto.
(b) If T is onto then $dim(R(T)) = n$. By Theorem 91, $dim(ker(T)) = 0$. Hence, $ker(T) = \{0\}$, i.e. T is one-one ■

The following theorem says that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a multiplication by an $m \times n$ matrix.

Theorem 93
Let $T : \mathbb{R}^n \to \mathbb{R}^m$. Then there exists an $m \times n$ matrix A such that

$$Tx = Ax$$

for all x in $\mathbb{R}^n$.

Proof.
We have

$$\begin{array}{rcl} x = I_n x & = & [e_1, e_2, \ldots, e_n]x \\ & = & x_1 e_1 + x_2 e_2 + \ldots + x_n e_n \end{array}$$

Now using the linearity of T to obtain

$$\begin{array}{rcl} Tx & = & T(x_1 e_1 + x_2 e_2 + \ldots + x_n e_n) \\ & = & x_1 T(e_1) + x_2 T(e_2) + \ldots + x_n T(e_n) \\ & = & [T(e_1), T(e_2), \ldots, T(e_n)]x \\ & = & Ax. \end{array}$$

This ends a proof of the theorem ■

The matrix A is called the **standard matrix for the linear transformation T.**

Exercise 407
Find the standard matrix A for the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(x) = 3x$.

Solution.
Since $T(e_i) = 3e_i$ then the standard matrix is the scalar matrix $3I_n$ ■

Theorem 94
Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $T(x) = Ax$, where A is an $m \times n$ matrix. Then
(a) $ker(T)$ is the nullspace of A.
(b) $R(T)$ is the column space of A.

Proof.
(a) $ker(T) = \{x \in \mathbb{R}^n : Ax = 0\} = null(A)$.
(b) If $b \in R(T)$ then there exists $x \in \mathbb{R}^n$ such that $Ax = b$. This is equivalent to $b = x_1c_1 + x_2c_2 + \cdots + x_nc_n$, where $x_1, x_2, \cdots, x_n$ are the components of x and $c_1, c_2, \cdots, c_n$ are the columns of A. Hence, $b \in span\{c_1, c_2, \cdots, c_n\}$ =column space of A. Hence, $R(T)$ is a subset of the column space of A. The converse is similar ■

It follows from the above theorem that $rank(T) = rank(A)$. Hence, by Theorem 91, $dim(ker(A)) = n - rank(A)$. Thus, if $rank(A) < n$ then the homogeneous system $Ax = 0$ has a nontrivial solution and if $rank(A) \geq n$ then the system has only the trivial solution.

A linear transformation $T : V \to W$ is said to be **invertible** if and only if there exists a unique function $T^{-1} : W \to V$ such that $T \circ T^{-1} = id_W$ and $T^{-1} \circ T = id_V$.

Theorem 95
Let $T : V \to W$ be an invertible linear transformation. Then
(a) T^{-1} is linear.
(b) $(T^{-1})^{-1} = T$.

Proof.
(a) Suppose $T^{-1}(w_1) = v_1, T^{-1}(w_2) = v_2$ and $\alpha \in \mathbb{R}$. Then $\alpha w_1 + w_2 = \alpha T(v_1) + T(v_2) = T(\alpha v_1 + v_2$. That is, $T^{-1}(\alpha w_1 + w_2) = \alpha v_1 + v_2 = \alpha T^{-1}(w_1) + T^{-1}(w_2)$.
(b) Follows from the definition of invertible functions ■

The following theorem provides us with an important link between invertible matrices and one-one linear transformations.

Theorem 96
A linear transformation $T : V \to W$ is invertible if and only if $ker(T) = \{0\}$ and $R(T) = W$.

Proof.
Suppose that T is such that $ker(T) = \{0\}$ and $R(T) = W$. That is, T is one-one and onto. Define, $S : W \to V$ by $S(w) = v$ where $T(v) = w$. Since T is one-one and onto then S is well-defined. Moreover, if $w \in W$ then $(T \circ S)(w) = T(S(w)) = T(v) = w$, i.e. $T \circ S = id_W$. Similarly, one shows that $S \circ T = id_V$. It remains to show that S is unique. Indeed, if S' is a linear transformation from W into V such that $S' \circ T = id_V$ and $T \circ S' = id_W$ then $S'(w) = S'(T(S(w))) = (S' \circ T)(S(w)) = S(w)$ for all $w \in W$. Hence, $S' = S$ and so T is invertible.
Conversely, suppose that T is invertible. If $T(v_1) = T(v_2)$ then $T^{-1}(T(v_1)) = T^{-1}(T(v_2))$. This implies that $v_1 = v_2$, i.e. T is one-one. Now, if $w \in W$ then $T(T^{-1}(w)) = w$. Let $v = T^{-1}(w)$ to obtain $T(v) = w$. That is T is onto. ■

Exercise 408
Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by $T(x) = Ax$ where A is the matrix

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

(a) Prove that T is invertible.
(b) What is $T^{-1}(x)$?

Solution.
(a) We must show that T is one-one and onto. Let $x = (x_1, y_1, z_1) \in ker(T)$. Then $Tx = Ax = (0, 0, 0)$. Since $|A| = -1 \neq 0$ then A is invertible and therefore $x = (0, 0, 0)$. Hence, $ker(T) = \{(0, 0, 0)\}$. Now since A is invertible the system $Ax = b$ is always solvable. This shows that $R(T) = \mathbb{R}^3$. Hence, by the above theorem, T is invertible.
(b) $T^{-1}x = A^{-1}x$ ■

The following theorem gives us a very useful characterization of isomorphism: They are linear transformations that preserve bases.

Theorem 97
Let $T : V \to W$ be a linear transformation with $V \neq \{0\}$. Then the following are all equivalent:
(a) T is an isomorphism.
(b) If $\{v_1, v_2, \cdots, v_n\}$ is a basis of V then $\{T(v_1, T(v_2), \cdots, T(v_n)\}$ is a basis of W.
(c) There exists a basis $\{v_1, v_2, \cdots, v_n\}$ of V such that $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ is a basis of W.

Proof.
(a) ⇒ (b): Suppose that T is an isomorphism. Let $\{v_1, v_2, \cdots, v_n\}$ be a basis of V. If $\alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = 0$ then $T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = 0$. Since T is one-one then $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$. But the vectors $v_1, v_2, \cdots, v_n$ are linearly independent. Hence, $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Next, we show that $span\{T(v_1), T(v_2), \cdots, T(v_n)\} = W$. Indeed, if $w \in W$ then there is a $v \in V$ such that $T(v) = w$ (recall that T is onto). But then $w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Thus, $T(w) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n)$. That is $w \in span\{T(v_1), T(v_2), \cdots, T(v_n)\}$.
(b) ⇒ (c): Since $V \neq \{0\}$ then V has a basis, say $\{v_1, v_2, \cdots, v_n\}$. By (b), $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ is a basis of W.
(c) ⇒ (a): Suppose that $T(v) = 0$. Since $v \in V$ then $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Hence, $\alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = 0$. Since $T(v_1), T(v_2), \cdots, T(v_n)$ are linearly independent then $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. This implies that $v = 0$. Hence $ker(T) = \{0\}$ and T is one-one. To show that T is onto, let $w \in W$. Then $w = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = T(v)$ where $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \in V$. ■

Finally, we end this section with the following theorem.

Theorem 98
Two finite dimensional vector spaces V and W are isomorphic if and only if $dim(V) = dim(W)$.

Proof.
Suppose that V and W are isomorphic. Then there is an isomorphism $T : V \to W$. Let $\{v_1, v_2, \cdots, v_n\}$ be a basis of V. Then $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ is a basis of W (Theorem 97). Thus, $dim(W) = n = dim(V)$.
Conversely, suppose that $dim(V) = dim(W)$. Let $\{v_1, v_2, \cdots, v_n\}$ be a basis of V and $\{w_1, w_2, \cdots, w_n\}$ be a basis of W. Then there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = w_i$ (Theorem 87). Hence, $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ is a basis of W. By Theorem 97, T is an isomorphism
■

6.3 The Matrix Representation of a Linear Transformation

Let V and W be two vector spaces such that $dim(V) = n$ and $dim(W) = m$. Let $T : V \to W$ be a linear transformation. The purpose of this section is to represent T as a matrix multiplication. The basic idea is to work with coordinate matrices rather than the vectors themselves.

Theorem 99
Let V and W be as above. Let $S = \{v_1, v_2, \cdots, v_n\}$ and $S' = \{u_1, u_2, \cdots, u_m\}$ be ordered bases for V and W respectively. Let A be the $m \times n$ matrix whose jth column is the coordinate vector of $T(v_j)$ with respect to the basis S'. Then A is the only matrix with the property that if $w = T(v)$ for some $v \in V$ then $[w]_{S'} = A[v]_S$.

Proof.
Since $T(v_j) \in W$ and S' is a basis of W then there exist unique scalars $a_{1j}, a_{2j}, \cdots, a_{mj}$ such that $T(v_j) = a_{1j}u_1 + a_{2j}u_2 + \cdots + a_{mj}u_m$. Hence, the coordinate of $T(v_i)$ with respect to S' is the vector

$$[T(v_j)]_{S'} = \left(\begin{array}{c} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{array}\right)$$

Let A be the $m \times n$ matrix whose jth column is $[T(v_j)]_{S'}$. We next show that A satisfies the property stated in the theorem. Indeed, let $v \in V$ and $w = T(v)$.

Then $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Take T of both sides to obtain

$$\begin{array}{rcc} T(v) & = & \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) \\ & = & \alpha_1(a_{11}u_1 + a_{21}u_2 + \cdots + a_{m1}u_m) \\ & + & \alpha_2(a_{12}u_1 + a_{22}u_2 + \cdots + a_{m2}u_m) \\ & + & \cdots + \alpha_n(a_{1n}u_1 + a_{2n}u_2 + \cdots + a_{mn}u_m) \\ & = & (a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n)u_1 \\ & + & (a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n)u_2 \\ & + & \cdots + (a_{m1}\alpha_1 + a_{m2}\alpha_2 + \cdots + a_{mn}\alpha_n)u_m \end{array}$$

Now, since $w \in W$ then $w = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_m u_m$. This and the above equalities lead to the following system of m equations in n unknowns:

$$\begin{array}{ccccccccc} a_{11}\alpha_1 & + & a_{12}\alpha_2 & + & \cdots & + & a_{1n}\alpha_n & = & \beta_1 \\ a_{21}\alpha_1 & + & a_{22}\alpha_2 & + & \cdots & + & a_{2n}\alpha_n & = & \beta_2 \\ \vdots & & & & & & & & \\ a_{m1}\alpha_1 & + & a_{m2}\alpha_2 & + & \cdots & + & a_{mn}\alpha_n & = & \beta_m \end{array}$$

In matrix notation, we have $A[v]_S = [w]_{S'}$.
It remains to show that A defined as above is the only $m \times n$ matrix with the property $A[v]_S = [w]_{S'}$. Let B be another matrix with the property $B[v]_S = [w]_{S'}$ and $B \neq A$. Write $A = (a_{ij})$ and $B = (b_{ij})$. Since A and B are assumed to be different then the kth columns of these two matrices are unequal. The coordinate vector of v_k is the vector

$$[v_k]_S = \left(\begin{array}{c} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{array}\right)$$

where the 1 is in the kth row of $[v_k]_S$. Thus,

$$[T(v_k)]_{S'} = A[v_k]_S = \left(\begin{array}{c} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{array}\right)$$

The right-hand side is just the kth column of A. On the other hand,

$$[T(v_k)]_{S'} = B[v_k]_S = \left(\begin{array}{c} b_{1k} \\ b_{2k} \\ \vdots \\ b_{mk} \end{array}\right)$$

which is the kth column of B. What we have shown here is that $T(v_k)$ has two different coordinate vectors with respect to the same basis S', which is impossible. Hence, $A = B$. ■

The matrix A is called the **representation of T with respect to the ordered bases S and S'**. In case $V = W$ then $S = S'$ and we call S the **matrix representation of T with respect to S.**

Exercise 409
Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by the formula

$$T\left(\begin{array}{c} x \\ y \end{array}\right) = \left(\begin{array}{c} x+2y \\ 2x-y \end{array}\right).$$

(a) Let $S = \{\vec{e_1}, \vec{e_2}\}$ be the standard basis of $\mathbb{R}^2$. Find the matrix representation of T with respect to S.
(b) Let

$$S' = \{\left(\begin{array}{c} -1 \\ 2 \end{array}\right), \left(\begin{array}{c} 2 \\ 0 \end{array}\right)\}.$$

Find the matrix representation of T with respect to S and S'.

Solution.
(a) We have the following computation

$$T\left(\begin{array}{c} 1 \\ 0 \end{array}\right) = \left(\begin{array}{c} 1 \\ 2 \end{array}\right)$$
$$T\left(\begin{array}{c} 0 \\ 1 \end{array}\right) = \left(\begin{array}{c} 2 \\ -1 \end{array}\right)$$

Thus, the matrix representation of T with respect to S is

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

(b) Let A be the matrix representation of T with respect to S and S'. Then the columns of A are determined as follows.

$$\begin{array}{rcl} T\left(\begin{array}{c} 1 \\ 0 \end{array}\right) & = & \left(\begin{array}{c} 1 \\ 2 \end{array}\right) \\ & = & \left(\begin{array}{c} -1 \\ 2 \end{array}\right) + \left(\begin{array}{c} 2 \\ 0 \end{array}\right) \\ T\left(\begin{array}{c} 0 \\ 1 \end{array}\right) & = & \left(\begin{array}{c} 2 \\ -1 \end{array}\right) \\ & = & -\frac{1}{2}\left(\begin{array}{c} -1 \\ 2 \end{array}\right) + \frac{3}{4}\left(\begin{array}{c} 2 \\ 0 \end{array}\right) \end{array}$$

Thus,

$$A = \left(\begin{array}{cc} 1 & -\frac{1}{2} \\ 1 & \frac{3}{4} \end{array}\right) ■$$

Exercise 410
Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\\0\end{pmatrix}, T\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}2\\0\\1\end{pmatrix}, T\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}1\\0\\1\end{pmatrix}$$

(a) Find the matrix representation of T with respect to the standard basis S of $\mathbb{R}^3$.
(b) Find

$$T\begin{pmatrix}1\\2\\3\end{pmatrix}$$

using the definition of T and then the matrix obtained in (a).

Solution.
(a) The matrix representation of T with respect to the standard basis S of $\mathbb{R}^3$ is

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

(b) Using the definition of T we have

$$\begin{aligned}T\begin{pmatrix}1\\2\\3\end{pmatrix} &= T\begin{pmatrix}1\\0\\0\end{pmatrix}+2T\begin{pmatrix}0\\1\\0\end{pmatrix}+3T\begin{pmatrix}0\\0\\1\end{pmatrix}\\ &= \begin{pmatrix}1\\1\\0\end{pmatrix}+\begin{pmatrix}4\\0\\2\end{pmatrix}+\begin{pmatrix}3\\0\\3\end{pmatrix}\\ &= \begin{pmatrix}8\\1\\5\end{pmatrix}\end{aligned}$$

Using (a) we find

$$\begin{aligned}T\begin{pmatrix}1\\2\\3\end{pmatrix} &= \begin{pmatrix}1&21&\\1&0&0\\0&1&1\end{pmatrix}\begin{pmatrix}1\\2\\3\end{pmatrix}\\ &= \begin{pmatrix}8\\1\\5\end{pmatrix}\blacksquare\end{aligned}$$

Exercise 411
Let $T : V \to V$ be a linear transformation defined by $T(v) = \alpha v$, where α is a fixed scalar. Prove that the matrix representation of T with respect to any ordered basis for V is a scalar matrix, i.e. a diagonal matrix whose diagonal entries are equal.

Solution.
If $\{v_1, v_2, \cdots, v_n\}$ is a basis for V then $T(v_i) = \alpha v_i$ for all $1 \leq i \leq n$. It follows that the matrix representation of T with respect to any basis is the scalar matrix αI_n ■

Exercise 412
Let V be an n-dimensional vector space. Prove that the matrix representation of the identity transformation on V is the identity matrix I_n.

Solution.
Since $id(v) = v$ for all $v \in V$ then in particular $id(v) = v$ if v is a basis element. It follows that the matrix representation of id is the identity matrix I_n ■

Exercise 413
Let V and W be vector spaces such that $dim(V) = n$ and $dim(W) = m$. Show that the set $L(V, W)$ of all linear transformations from V into W is a vector space under the operations of additions of functions and scalar multiplication.

Solution.
$L(V, W)$ is closed under addition and scalar multiplication according to Exercise 396. It is left for the reader to check the axioms of a vector space ■

Theorem 100
The vector space $L(V, W)$ is isomorphic to the vector space M_{mn} of all $m \times n$ matrices.

Proof.
Let $S = \{v_1, v_2, \cdots, v_n\}$ and $S' = \{u_1, u_2, \cdots, u_m\}$ be ordered bases for V and W respectively. Given $T \in L(V, W)$ there exists a unique matrix $A \in M_{mn}$ such that $[w]_{S'} = A[v]_S$. Thus, the function $f : L(V, W) \to M_{mn}$ given by $f(T) = A$ is well-defined linear transformation (See Exercise 414). We next show that f is one-one. Indeed, let $T_1, T_2 \in L(V, W)$ be such that $T_1 \neq T_2$. Then $T_1(v_k) \neq T_2(v_k)$ for some $1 \leq k \leq n$. This implies that $[T_1(v_k)]_{S'} \neq [T_2(v_k)]_{S'}$ which in turn implies that the matrix representation of T_1 with respect to S and S' is different from the matrix representation of T_2. Hence, $f(T_1) \neq f(T_2)$. It remains ro show that f is onto. Let $A = (a_{ij}) \in M_{mn}$. Define a function $T : V \to W$ by $T(v_k) = a_{1k}u_1 + a_{2k}u_2 + \cdots + a_{mk}u_m$ and $T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n)$. Then T is a linear transformation, i,e $T \in L(V, W)$ and the matrix representing T with respect to S and S' is A. (See Exercise 415) ■

Exercise 414
Show that the function f defined in the previous theorem is a linear transformation from $L(V, W)$ into M_{mn}.

Solution.
Let $S, T \in L(V, W)$ and $\alpha \in \mathbb{R}$. Then there exist matrices $A, B \in M_{mn}$ such

that $f(T) = A$ and $f(S) = B$. Since $\alpha T + S$ is a linear transformation then there exists a matric $C \in M_{mn}$ such that $f(\alpha T + S) = C$. If v_i is a basis vector of V then $(\alpha T(v_i) + S(v_i))$ is the ith column of C. That is, the matrix C looks like $[\ \alpha T(v_1) + S(v_1) \quad \alpha T(v_2) + S(v_2) \quad \cdots \quad \alpha T(v_n) + S(v_n)\ ]$. But this is the same as $\alpha A + B$. Hence, $f(\alpha T + S) = \alpha f(T) + f(S)$ ■

Exercise 415
Show that the function T defined in Theorem 100 is a linear transforamtion from V into W with matrix representation with respect to S and S' equals to A.

Solution.
Let $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ and $w = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$ be two vectors in V and $\alpha \in \mathbb{R}$. Then $T(\alpha v + w) = T((\alpha\alpha_1 + \beta_1)v_1 + (\alpha\alpha_2 + \beta_2)v_2 + \cdots + (\alpha\alpha_n + \beta_n)v_n) = (\alpha\alpha_1 + \beta_1)T(v_1) + (\alpha\alpha_2 + \beta_2)T(v_2) + \cdots + (\alpha\alpha_n + \beta_n)T(v_n) = \alpha T(v) + T(w)$. That is, T is linear. The fact that the matrix representation of T with respect to S and S' is clear ■

Theorem 100 implies that $dim(L(V, W)) = dim(M_{mn}) = mn$. Also, it means that when dealing with finite-dimensional vector spaces, we can always replace all linear transformations by their matrix representations and work only with matrices.

6.4 Review Problems

Exercise 416
Show that the function $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x, y) = (x + y, x - 2y, 3x)$ is a linear transformation.

Exercise 417
Let P_n be the vector space of all polynomials of degree at most n.
(a) Show that $D : P_n \longrightarrow P_{n-1}$ given by $D(p) = p'$ is a linear transformation.
(b) Show that $I : P_n \longrightarrow P_{n+1}$ given by $I(p) = \int_0^x p(t)dt$ is a linear transformation.

Exercise 418
If $T : \mathbb{R}^3 \longrightarrow \mathbb{R}$ is a linear transformation with $T(3, -1, 2) = 5$ and $T(1, 0, 1) = 2$. Find $T(-1, 1, 0)$.

Exercise 419
Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be the transformation $T(x, y, z) = (x, y)$. Show that T is linear. This transformation is called a **projection.**

Exercise 420
Let θ be a given angle. Define $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Show that T is a linear transformation. Geometrically, Tv is the vector that results if v is rotated counterclockwise by an angle θ. We call this transformation the **rotation of $\mathbb{R}^2$ through the angle θ**

Exercise 421
Show that the following transformations are not linear.

(a) $T : M_{nn} \longrightarrow \mathbb{R}$ *given by* $T(A) = |A|$.
(b) $T : M_{mn} \longrightarrow \mathbb{R}$ *given by* $T(A) = rank(A)$.

Exercise 422
If $T_1 : U \longrightarrow V$ *and* $T_2 : V \longrightarrow W$ *are linear transformations, then* $T_2 \circ T_1 : U \longrightarrow W$ *is also a linear transformation.*

Exercise 423
Let $\{v_1, v_2, \cdots, v_n\}$ *be a basis for a vector space V, and let* $T : V \longrightarrow V$ *be a linear transformation. Show that if* $T(v_i) = v_i$, *for* $1 \leq i \leq n$ *then* $T = id_V$, *i.e. T is the identity transformation on V.*

Exercise 424
Let T be a linear transformation on a vector space V such that $T(v - 3v_1) = w$ *and* $T(2v - v_1) = w_1$. *Find* $T(v)$ *and* $T(v_1)$ *in terms of w and* w_1.

Exercise 425
Let $T : M_{mn} \to M_{mn}$ *be given by* $T(X) = AX$ *for all* $X \in M_{mn}$, *where A is an* $m \times m$ *invertible matrix. Show that T is both one-one and onto.*

Exercise 426
Consider the transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *defined by* $T(x) = Ax$, *where* $A \in M_{mn}$.

(a) Show that $R(T) = span\{c_1, c_2, \cdots, c_n\}$, *where* $c_1, c_2, \cdots, c_n$ *are the columns of A.*
(b) Show that T is onto if and only if $rank(A) = m$ *(i.e. the rows of A are linearly independent).*
(c) Show that T is one-one if and only if $rank(A) = n$ *(i.e. the columns of A are linearly independent).*

Exercise 427
Let $T : V \longrightarrow W$ *be a linear transformation. Show that if the vectors*

$$T(v_1), T(v_2), \cdots, T(v_n)$$

are linearly independent then the vectors $v_1, v_2, \cdots, v_n$ *are also linearly independent.*

Exercise 428
Show that the projection transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ *defined by* $T(x, y, z) = (x, y)$ *is not one-one.*

Exercise 429
Let M_{nn} be the vector space of all $n \times n$ matrices. Let $T : M_{nn} \to M_{nn}$ be given by $T(A) = A - A^T$.
(a) Show that T is linear.
(b) Find $ker(T)$ and $R(T)$.

Exercise 430
Let $T : V \to W$. Prove that T is one-one if and only if $dim(R(T)) = dim(V)$.

Exercise 431
Show that the linear transformation $T : M_{nn} \to M_{nn}$ given by $T(A) = A^T$ is an isomorphism.

Exercise 432
Let $T : P_2 \to P_1$ be the linear transformation $Tp = p'$. Consider the standard ordered bases $S = \{1, x, x^2\}$ and $S' = \{1, x\}$. Find the matrix representation of T with respect to the basis S and S'.

Exercise 433
Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$

(a) Find the matrix representation of T with respect to the standard basis S of $\mathbb{R}^2$.
(b) Let

$$S' = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\}$$

Find the matrix representation of T with repspect to the bases S and S'.

Exercise 434
Let V be the vector space of continuous functions on $\mathbb{R}$ with the ordered basis $S = \{\sin t, \cos t\}$. Find the matrix representation of the linear transformation $T : V \to V$ defined by $T(f) = f'$ with respect to S.

Exercise 435
Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation whose matrix representation with the respect to the standard basis of $\mathbb{R}^3$ is given by

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

Find

$$T\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

.