Chapter 3

SUBSPACES

3.1 Introduction

Throughout this chapter, we will be studying F^n, the set of all n–dimensional column vectors with components from a field F. We continue our study of matrices by considering an important class of subsets of F^n called *subspaces*. These arise naturally for example, when we solve a system of m linear homogeneous equations in n unknowns.

We also study the concept of linear dependence of a family of vectors. This was introduced briefly in Chapter 2, Remark 2.5.4. Other topics discussed are the *row space, column space* and *null space* of a matrix over F, the *dimension* of a subspace, particular examples of the latter being the *rank* and *nullity* of a matrix.

3.2 Subspaces of F^n

DEFINITION 3.2.1 A subset S of F^n is called a subspace of F^n if

1. The zero vector belongs to S; (that is, $0 \in S$);
2. If $u \in S$ and $v \in S$, then $u + v \in S$; (S is said to be closed under vector addition);
3. If $u \in S$ and $t \in F$, then $tu \in S$; (S is said to be closed under scalar multiplication).

EXAMPLE 3.2.1 Let $A \in M_{m\times n}(F)$. Then the set of vectors $X \in F^n$ satisfying $AX = 0$ is a subspace of F^n called the *null space* of A and is denoted here by $N(A)$. (It is sometimes called the *solution space* of A.)

Proof. (1) $A0 = 0$, so $0 \in N(A)$; (2) If $X, Y \in N(A)$, then $AX = 0$ and $AY = 0$, so $A(X+Y) = AX + AY = 0 + 0 = 0$ and so $X + Y \in N(A)$; (3) If $X \in N(A)$ and $t \in F$, then $A(tX) = t(AX) = t0 = 0$, so $tX \in N(A)$.

For example, if $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then $N(A) = \{0\}$, the set consisting of just the zero vector. If $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, then $N(A)$ is the set of all scalar multiples of $[-2, 1]^t$.

EXAMPLE 3.2.2 Let $X_1, \ldots, X_m \in F^n$. Then the set consisting of all linear combinations $x_1X_1 + \cdots + x_mX_m$, where $x_1, \ldots, x_m \in F$, is a subspace of F^n. This subspace is called the subspace *spanned* or *generated* by $X_1, \ldots, X_m$ and is denoted here by $\langle X_1, \ldots, X_m\rangle$. We also call $X_1, \ldots, X_m$ a spanning family for $S = \langle X_1, \ldots, X_m\rangle$.

Proof. (1) $0 = 0X_1 + \cdots + 0X_m$, so $0 \in \langle X_1, \ldots, X_m\rangle$; (2) If $X, Y \in \langle X_1, \ldots, X_m\rangle$, then $X = x_1X_1 + \cdots + x_mX_m$ and $Y = y_1X_1 + \cdots + y_mX_m$, so

$$\begin{aligned} X + Y &= (x_1X_1 + \cdots + x_mX_m) + (y_1X_1 + \cdots + y_mX_m) \\ &= (x_1 + y_1)X_1 + \cdots + (x_m + y_m)X_m \in \langle X_1, \ldots, X_m\rangle. \end{aligned}$$

(3) If $X \in \langle X_1, \ldots, X_m\rangle$ and $t \in F$, then

$$\begin{aligned} X &= x_1X_1 + \cdots + x_mX_m \\ tX &= t(x_1X_1 + \cdots + x_mX_m) \\ &= (tx_1)X_1 + \cdots + (tx_m)X_m \in \langle X_1, \ldots, X_m\rangle. \end{aligned}$$

For example, if $A \in M_{m\times n}(F)$, the subspace generated by the columns of A is an important subspace of F^m and is called the *column space* of A. The column space of A is denoted here by $C(A)$. Also the subspace generated by the rows of A is a subspace of F^n and is called the *row space* of A and is denoted by $R(A)$.

EXAMPLE 3.2.3 For example $F^n = \langle E_1, \ldots, E_n\rangle$, where $E_1, \ldots, E_n$ are the n–dimensional unit vectors. For if $X = [x_1, \ldots, x_n]^t \in F^n$, then $X = x_1E_1 + \cdots + x_nE_n$.

EXAMPLE 3.2.4 Find a spanning family for the subspace S of $\mathbb{R}^3$ defined by the equation $2x - 3y + 5z = 0$.

Solution. (S is in fact the null space of $[2,\ -3,\ 5]$, so S is indeed a subspace of $\mathbb{R}^3$.)

If $[x,\ y,\ z]^t \in S$, then $x = \frac{3}{2}y - \frac{5}{2}z$. Then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{3}{2}y - \frac{5}{2}z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} \frac{3}{2} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -\frac{5}{2} \\ 0 \\ 1 \end{bmatrix}$$

and conversely. Hence $[\frac{3}{2},\ 1,\ 0]^t$ and $[-\frac{5}{2},\ 0,\ 1]^t$ form a spanning family for S.

The following result is easy to prove:

LEMMA 3.2.1 Suppose each of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$. Then any linear combination of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$.

As a corollary we have

THEOREM 3.2.1 Subspaces $\langle X_1, \ldots, X_r\rangle$ and $\langle Y_1, \ldots, Y_s\rangle$ are equal if each of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$ and each of $Y_1, \ldots, Y_s$ is a linear combination of $X_1, \ldots, X_r$.

COROLLARY 3.2.1 Subspaces $\langle X_1, \ldots, X_r,\ Z_1, \ldots, Z_t\rangle$ and $\langle X_1, \ldots, X_r\rangle$ are equal if each of $Z_1, \ldots, Z_t$ is a linear combination of $X_1, \ldots, X_r$.

EXAMPLE 3.2.5 If X and Y are vectors in $\mathbb{R}^n$, then

$$\langle X,\ Y\rangle = \langle X + Y,\ X - Y\rangle.$$

Solution. Each of $X + Y$ and $X - Y$ is a linear combination of X and Y. Also

$$X = \frac{1}{2}(X + Y) + \frac{1}{2}(X - Y) \quad \text{and} \quad Y = \frac{1}{2}(X + Y) - \frac{1}{2}(X - Y),$$

so each of X and Y is a linear combination of $X + Y$ and $X - Y$.

There is an important application of Theorem 3.2.1 to row equivalent matrices (see Definition 1.2.4):

THEOREM 3.2.2 If A is row equivalent to B, then $R(A) = R(B)$.

Proof. Suppose that B is obtained from A by a sequence of elementary row operations. Then it is easy to see that each row of B is a linear combination of the rows of A. But A can be obtained from B by a sequence of elementary operations, so each row of A is a linear combination of the rows of B. Hence by Theorem 3.2.1, $R(A) = R(B)$.

REMARK 3.2.1 If A is row equivalent to B, it is not always true that $C(A) = C(B)$.

For example, if $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, then B is in fact the reduced row–echelon form of A. However we see that

$$C(A) = \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle$$

and similarly $C(B) = \left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle$.

Consequently $C(A) \neq C(B)$, as $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \in C(A)$ but $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin C(B)$.

3.3 Linear dependence

We now recall the definition of linear dependence and independence of a family of vectors in F^n given in Chapter 2.

DEFINITION 3.3.1 Vectors $X_1, \ldots, X_m$ in F^n are said to be *linearly dependent* if there exist scalars $x_1, \ldots, x_m$, *not all zero*, such that

$$x_1X_1 + \cdots + x_mX_m = 0.$$

In other words, $X_1, \ldots, X_m$ are linearly dependent if some X_i is expressible as a linear combination of the remaining vectors.

$X_1, \ldots, X_m$ are called *linearly independent* if they are not linearly dependent. Hence $X_1, \ldots, X_m$ are linearly independent if and only if the equation

$$x_1X_1 + \cdots + x_mX_m = 0$$

has only the trivial solution $x_1 = 0, \ldots, x_m = 0$.

EXAMPLE 3.3.1 The following three vectors in $\mathbb{R}^3$

$$X_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ 7 \\ 12 \end{bmatrix}$$

are linearly dependent, as $2X_1 + 3X_2 + (-1)X_3 = 0$.

REMARK 3.3.1 If $X_1, \ldots, X_m$ are linearly independent and

$$x_1X_1 + \cdots + x_mX_m = y_1X_1 + \cdots + y_mX_m,$$

then $x_1 = y_1, \ldots, x_m = y_m$. For the equation can be rewritten as

$$(x_1 - y_1)X_1 + \cdots + (x_m - y_m)X_m = 0$$

and so $x_1 - y_1 = 0, \ldots, x_m - y_m = 0$.

THEOREM 3.3.1 A family of m vectors in F^n will be linearly dependent if $m > n$. Equivalently, any linearly independent family of m vectors in F^n must satisfy $m \leq n$.

Proof. The equation

$$x_1X_1 + \cdots + x_mX_m = 0$$

is equivalent to n homogeneous equations in m unknowns. By Theorem 1.5.1, such a system has a non–trivial solution if $m > n$.

The following theorem is an important generalization of the last result and is left as an exercise for the interested student:

THEOREM 3.3.2 A family of s vectors in $\langle X_1, \ldots, X_r\rangle$ will be linearly dependent if $s > r$. Equivalently, a linearly independent family of s vectors in $\langle X_1, \ldots, X_r\rangle$ must have $s \leq r$.

Here is a useful criterion for linear independence which is sometimes called the *left–to–right test*:

THEOREM 3.3.3 Vectors $X_1, \ldots, X_m$ in F^n are linearly independent if

(a) $X_1 \neq 0$;

(b) For each k with $1 < k \leq m$, X_k is not a linear combination of $X_1, \ldots, X_{k-1}$.

One application of this criterion is the following result:

THEOREM 3.3.4 Every subspace S of F^n can be represented in the form $S = \langle X_1, \ldots, X_m\rangle$, where $m \leq n$.

Proof. If $S = \{0\}$, there is nothing to prove – we take $X_1 = 0$ and $m = 1$.

So we assume S contains a non–zero vector X_1; then $\langle X_1\rangle \subseteq S$ as S is a subspace. If $S = \langle X_1\rangle$, we are finished. If not, S will contain a vector X_2, not a linear combination of X_1; then $\langle X_1, X_2\rangle \subseteq S$ as S is a subspace. If $S = \langle X_1, X_2\rangle$, we are finished. If not, S will contain a vector X_3 which is not a linear combination of X_1 and X_2. This process must eventually stop, for at stage k we have constructed a family of k linearly independent vectors $X_1, \ldots, X_k$, all lying in F^n and hence $k \le n$.

There is an important relationship between the columns of A and B, if A is row–equivalent to B.

THEOREM 3.3.5 Suppose that A is row equivalent to B and let $c_1, \ldots, c_r$ be distinct integers satisfying $1 \le c_i \le n$. Then

(a) Columns $A_{*c_1}, \ldots, A_{*c_r}$ of A are linearly dependent if and only if the corresponding columns of B are linearly dependent; indeed more is true:

$$x_1A_{*c_1} + \cdots + x_rA_{*c_r} = 0 \Leftrightarrow x_1B_{*c_1} + \cdots + x_rB_{*c_r} = 0.$$

(b) Columns $A_{*c_1}, \ldots, A_{*c_r}$ of A are linearly independent if and only if the corresponding columns of B are linearly independent.

(c) If $1 \le c_{r+1} \le n$ and c_{r+1} is distinct from $c_1, \ldots, c_r$, then

$$A_{*c_{r+1}} = z_1A_{*c_1} + \cdots + z_rA_{*c_r} \Leftrightarrow B_{*c_{r+1}} = z_1B_{*c_1} + \cdots + z_rB_{*c_r}.$$

Proof. First observe that if $Y = [y_1, \ldots, y_n]^t$ is an n–dimensional column vector and A is $m \times n$, then

$$AY = y_1A_{*1} + \cdots + y_nA_{*n}.$$

Also $AY = 0 \Leftrightarrow BY = 0$, if B is row equivalent to A. Then (a) follows by taking $y_i = x_{c_j}$ if $i = c_j$ and $y_i = 0$ otherwise.

(b) is logically equivalent to (a), while (c) follows from (a) as

$$\begin{aligned} A_{*c_{r+1}} = z_1A_{*c_1} + \cdots + z_rA_{*c_r} &\Leftrightarrow z_1A_{*c_1} + \cdots + z_rA_{*c_r} + (-1)A_{*c_{r+1}} = 0 \\ &\Leftrightarrow z_1B_{*c_1} + \cdots + z_rB_{*c_r} + (-1)B_{*c_{r+1}} = 0 \\ &\Leftrightarrow B_{*c_{r+1}} = z_1B_{*c_1} + \cdots + z_rB_{*c_r}. \end{aligned}$$

EXAMPLE 3.3.2 The matrix

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

has reduced row–echelon form equal to

$$B = \begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}.$$

We notice that B_{*1}, B_{*2} and B_{*4} are linearly independent and hence so are A_{*1}, A_{*2} and A_{*4}. Also

$$\begin{aligned} B_{*3} &= 2B_{*1} + 3B_{*2} \\ B_{*5} &= (-1)B_{*1} + 2B_{*2} + 3B_{*4}, \end{aligned}$$

so consequently

$$\begin{aligned} A_{*3} &= 2A_{*1} + 3A_{*2} \\ A_{*5} &= (-1)A_{*1} + 2A_{*2} + 3A_{*4}. \end{aligned}$$

3.4 Basis of a subspace

We now come to the important concept of *basis* of a vector subspace.

DEFINITION 3.4.1 Vectors $X_1, \ldots, X_m$ belonging to a subspace S are said to form a basis of S if

(a) Every vector in S is a linear combination of $X_1, \ldots, X_m$;

(b) $X_1, \ldots, X_m$ are linearly independent.

Note that (a) is equivalent to the statement that $S = \langle X_1, \ldots, X_m \rangle$ as we automatically have $\langle X_1, \ldots, X_m \rangle \subseteq S$. Also, in view of Remark 3.3.1 above, (a) and (b) are equivalent to the statement that every vector in S is *uniquely* expressible as a linear combination of $X_1, \ldots, X_m$.

EXAMPLE 3.4.1 The unit vectors $E_1, \ldots, E_n$ form a basis for F^n.

REMARK 3.4.1 The subspace $\{0\}$, consisting of the zero vector alone, does not have a basis. For every vector in a linearly independent family must necessarily be non–zero. (For example, if $X_1 = 0$, then we have the non–trivial linear relation

$$1X_1 + 0X_2 + \cdots + 0X_m = 0$$

and $X_1, \ldots, X_m$ would be linearly independent.)

However if we exclude this case, every other subspace of F^n has a basis:

THEOREM 3.4.1 A subspace of the form $\langle X_1, \ldots, X_m \rangle$, where at least one of $X_1, \ldots, X_m$ is non–zero, has a basis $X_{c_1}, \ldots, X_{c_r}$, where $1 \leq c_1 < \cdots < c_r \leq m$.

Proof. (The *left–to–right algorithm*). Let c_1 be the least index k for which X_k is non–zero. If $c_1 = m$ or if all the vectors X_k with $k > c_1$ are linear combinations of X_{c_1}, terminate the algorithm and let $r = 1$. Otherwise let c_2 be the least integer $k > c_1$ such that X_k is not a linear combination of X_{c_1}.

If $c_2 = m$ or if all the vectors X_k with $k > c_2$ are linear combinations of X_{c_1} and X_{c_2}, terminate the algorithm and let $r = 2$. Eventually the algorithm will terminate at the r–th stage, either because $c_r = m$, or because all vectors X_k with $k > c_r$ are linear combinations of $X_{c_1}, \ldots, X_{c_r}$.

Then it is clear by the construction of $X_{c_1}, \ldots, X_{c_r}$, using Corollary 3.2.1 that

(a) $\langle X_{c_1}, \ldots, X_{c_r} \rangle = \langle X_1, \ldots, X_m \rangle$;

(b) the vectors $X_{c_1}, \ldots, X_{c_r}$ are linearly independent by the left–to–right test.

Consequently $X_{c_1}, \ldots, X_{c_r}$ form a basis (called the *left–to–right basis*) for the subspace $\langle X_1, \ldots, X_m \rangle$.

EXAMPLE 3.4.2 Let X and Y be linearly independent vectors in $\mathbb{R}^n$. Then the subspace $\langle 0,\, 2X,\, X,\, -Y,\, X+Y \rangle$ has left–to–right basis consisting of $2X$, $-Y$.

A subspace S will in general have more than one basis. For example, any permutation of the vectors in a basis will yield another basis. Given one particular basis, one can determine all bases for S using a simple formula. This is left as one of the problems at the end of this chapter.

We settle for the following important fact about bases:

THEOREM 3.4.2 Any two bases for a subspace S must contain the same number of elements.

Proof. For if $X_1, \ldots, X_r$ and $Y_1, \ldots, Y_s$ are bases for S, then $Y_1, \ldots, Y_s$ form a linearly independent family in $S = \langle X_1, \ldots, X_r \rangle$ and hence $s \leq r$ by Theorem 3.3.2. Similarly $r \leq s$ and hence $r = s$.

DEFINITION 3.4.2 This number is called the *dimension* of S and is written $\dim S$. Naturally we define $\dim \{0\} = 0$.

It follows from Theorem 3.3.1 that for any subspace S of F^n, we must have $\dim S \leq n$.

EXAMPLE 3.4.3 If $E_1, \ldots, E_n$ denote the n–dimensional unit vectors in F^n, then $\dim \langle E_1, \ldots, E_i \rangle = i$ for $1 \leq i \leq n$.

The following result gives a useful way of exhibiting a basis.

THEOREM 3.4.3 A linearly independent family of m vectors in a subspace S, with $\dim S = m$, must be a basis for S.

Proof. Let $X_1, \ldots, X_m$ be a linearly independent family of vectors in a subspace S, where $\dim S = m$. We have to show that every vector $X \in S$ is expressible as a linear combination of $X_1, \ldots, X_m$. We consider the following family of vectors in S: $X_1, \ldots, X_m, X$. This family contains $m+1$ elements and is consequently linearly dependent by Theorem 3.3.2. Hence we have

$$x_1X_1 + \cdots + x_mX_m + x_{m+1}X = 0, \tag{3.1}$$

where not all of $x_1, \ldots, x_{m+1}$ are zero. Now if $x_{m+1} = 0$, we would have

$$x_1X_1 + \cdots + x_mX_m = 0,$$

with not all of $x_1, \ldots, x_m$ zero, contradictiong the assumption that $X_1 \ldots, X_m$ are linearly independent. Hence $x_{m+1} \neq 0$ and we can use equation 3.1 to express X as a linear combination of $X_1, \ldots, X_m$:

$$X = \frac{-x_1}{x_{m+1}}X_1 + \cdots + \frac{-x_m}{x_{m+1}}X_m.$$

3.5 Rank and nullity of a matrix

We can now define three important integers associated with a matrix.

DEFINITION 3.5.1 Let $A \in M_{m\times n}(F)$. Then

(a) column rank $A = \dim C(A)$;

(b) row rank $A = \dim R(A)$;

(c) nullity $A = \dim N(A)$.

We will now see that the reduced row–echelon form B of a matrix A allows us to exhibit bases for the row space, column space and null space of A. Moreover, an examination of the number of elements in each of these bases will immediately result in the following theorem:

THEOREM 3.5.1 Let $A \in M_{m\times n}(F)$. Then

(a) column rank $A =$ row rank A;

(b) column rank $A +$ nullity $A = n$.

Finding a basis for $R(A)$: The r non–zero rows of B form a basis for $R(A)$ and hence row rank $A = r$.

For we have seen earlier that $R(A) = R(B)$. Also

$$\begin{aligned} R(B) &= \langle B_{1*}, \ldots, B_{m*}\rangle \\ &= \langle B_{1*}, \ldots, B_{r*},\ 0 \ldots, 0\rangle \\ &= \langle B_{1*}, \ldots, B_{r*}\rangle. \end{aligned}$$

The linear independence of the non–zero rows of B is proved as follows: Let the leading entries of rows $1, \ldots, r$ of B occur in columns $c_1, \ldots, c_r$. Suppose that

$$x_1 B_{1*} + \cdots + x_r B_{r*} = 0.$$

Then equating components $c_1, \ldots, c_r$ of both sides of the last equation, gives $x_1 = 0, \ldots, x_r = 0$, in view of the fact that B is in reduced row– echelon form.

Finding a basis for $C(A)$: The r columns $A_{*c_1}, \ldots, A_{*c_r}$ form a basis for $C(A)$ and hence column rank $A = r$. For it is clear that columns $c_1, \ldots, c_r$ of B form the left–to–right basis for $C(B)$ and consequently from parts (b) and (c) of Theorem 3.3.5, it follows that columns $c_1, \ldots, c_r$ of A form the left–to–right basis for $C(A)$.

Finding a basis for $N(A)$: For notational simplicity, let us suppose that $c_1 = 1, \ldots, c_r = r$. Then B has the form

$$B = \begin{bmatrix} 1 & 0 & \cdots & 0 & b_{1r+1} & \cdots & b_{1n} \\ 0 & 1 & \cdots & 0 & b_{2r+1} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{rr+1} & \cdots & b_{rn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Then $N(B)$ and hence $N(A)$ are determined by the equations

$$\begin{aligned} x_1 &= (-b_{1r+1})x_{r+1} + \cdots + (-b_{1n})x_n \\ &\vdots \\ x_r &= (-b_{rr+1})x_{r+1} + \cdots + (-b_{rn})x_n, \end{aligned}$$

where $x_{r+1}, \ldots, x_n$ are arbitrary elements of F. Hence the general vector X in $N(A)$ is given by

$$\begin{aligned} \begin{bmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{bmatrix} &= x_{r+1} \begin{bmatrix} -b_{1r+1} \\ \vdots \\ -b_{rr+1} \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} -b_n \\ \vdots \\ -b_{rn} \\ 0 \\ \vdots \\ 1 \end{bmatrix} \qquad (3.2) \\ &= x_{r+1}X_1 + \cdots + x_n X_{n-r}. \end{aligned}$$

Hence $N(A)$ is spanned by $X_1, \ldots, X_{n-r}$, as $x_{r+1}, \ldots, x_n$ are arbitrary. Also $X_1, \ldots, X_{n-r}$ are linearly independent. For equating the right hand side of equation 3.2 to 0 and then equating components $r+1, \ldots, n$ of both sides of the resulting equation, gives $x_{r+1} = 0, \ldots, x_n = 0$.

Consequently $X_1, \ldots, X_{n-r}$ form a basis for $N(A)$.

Theorem 3.5.1 now follows. For we have

$$\begin{aligned} \text{row rank}\, A &= \dim R(A) = r \\ \text{column rank}\, A &= \dim C(A) = r. \end{aligned}$$

Hence

$$\text{row rank}\, A = \text{column rank}\, A.$$

Also

$$\text{column rank}\, A + \text{nullity}\, A = r + \dim N(A) = r + (n - r) = n.$$

DEFINITION 3.5.2 The common value of column rank A and row rank A is called the *rank* of A and is denoted by rank A.

EXAMPLE 3.5.1 Given that the reduced row–echelon form of

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

equal to

$$B = \begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix},$$

find bases for $R(A)$, $C(A)$ and $N(A)$.

Solution. $[1, 0, 2, 0, -1]$, $[0, 1, 3, 0, 2]$ and $[0, 0, 0, 1, 3]$ form a basis for $R(A)$. Also

$$A_{*1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; A_{*2} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \; A_{*4} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

form a basis for $C(A)$.

Finally $N(A)$ is given by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2x_3 + x_5 \\ -3x_3 - 2x_5 \\ x_3 \\ -3x_5 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 1 \\ -2 \\ 0 \\ -3 \\ 1 \end{bmatrix} = x_3 X_1 + x_5 X_2,$$

where x_3 and x_5 are arbitrary. Hence X_1 and X_2 form a basis for $N(A)$.

Here rank $A = 3$ and nullity $A = 2$.

EXAMPLE 3.5.2 Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Then $B = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ is the reduced row–echelon form of A.

Hence $[1,\ 2]$ is a basis for $R(A)$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is a basis for $C(A)$. Also $N(A)$ is given by the equation $x_1 = -2x_2$, where x_2 is arbitrary. Then

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

and hence $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ is a basis for $N(A)$.

Here $\operatorname{rank} A = 1$ and $\operatorname{nullity} A = 1$.

EXAMPLE 3.5.3 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is the reduced row–echelon form of A.

Hence $[1,\ 0]$, $[0,\ 1]$ form a basis for $R(A)$ while $[1,\ 3]$, $[2,\ 4]$ form a basis for $C(A)$. Also $N(A) = \{0\}$.

Here $\operatorname{rank} A = 2$ and $\operatorname{nullity} A = 0$.

We conclude this introduction to vector spaces with a result of great theoretical importance.

THEOREM 3.5.2 Every linearly independent family of vectors in a subspace S can be extended to a basis of S.

Proof. Suppose S has basis $X_1, \ldots, X_m$ and that $Y_1, \ldots, Y_r$ is a linearly independent family of vectors in S. Then

$$S = \langle X_1, \ldots, X_m \rangle = \langle Y_1, \ldots, Y_r,\ X_1, \ldots, X_m \rangle,$$

as each of $Y_1, \ldots, Y_r$ is a linear combination of $X_1, \ldots, X_m$.

Then applying the left–to–right algorithm to the second spanning family for S will yield a basis for S which includes $Y_1, \ldots, Y_r$.

3.6 PROBLEMS

1. Which of the following subsets of $\mathbb{R}^2$ are subspaces?

   (a) $[x,\ y]$ satisfying $x = 2y$;

   (b) $[x,\ y]$ satisfying $x = 2y$ and $2x = y$;

   (c) $[x,\ y]$ satisfying $x = 2y + 1$;

   (d) $[x,\ y]$ satisfying $xy = 0$;

(e) $[x, y]$ satisfying $x \geq 0$ and $y \geq 0$.

[Answer: (a) and (b).]

2. If X, Y, Z are vectors in $\mathbb{R}^n$, prove that

$$\langle X, Y, Z\rangle = \langle X+Y, X+Z, Y+Z\rangle.$$

3. Determine if $X_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$ and $X_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \end{bmatrix}$ are linearly independent in $\mathbb{R}^4$.

4. For which real numbers λ are the following vectors linearly independent in $\mathbb{R}^3$?

$$X_1 = \begin{bmatrix} \lambda \\ -1 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ \lambda \\ -1 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ -1 \\ \lambda \end{bmatrix}.$$

5. Find bases for the row, column and null spaces of the following matrix over $\mathbb{Q}$:

$$A = \begin{bmatrix} 1 & 1 & 2 & 0 & 1 \\ 2 & 2 & 5 & 0 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 8 & 11 & 19 & 0 & 11 \end{bmatrix}.$$

6. Find bases for the row, column and null spaces of the following matrix over $\mathbb{Z}_2$:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

7. Find bases for the row, column and null spaces of the following matrix over $\mathbb{Z}_5$:

$$A = \begin{bmatrix} 1 & 1 & 2 & 0 & 1 & 3 \\ 2 & 1 & 4 & 0 & 3 & 2 \\ 0 & 0 & 0 & 1 & 3 & 0 \\ 3 & 0 & 2 & 4 & 3 & 2 \end{bmatrix}.$$

8. Find bases for the row, column and null spaces of the matrix A defined in section 1.6, Problem 17. (Note: In this question, F is a field of four elements.)

9. If $X_1, \ldots, X_m$ form a basis for a subspace S, prove that

$$X_1,\ X_1 + X_2, \ldots, X_1 + \cdots + X_m$$

also form a basis for S.

10. Let $A = \begin{bmatrix} a & b & c \\ 1 & 1 & 1 \end{bmatrix}$. Find conditions on a, b, c such that (a) $\operatorname{rank} A = 1$; (b) $\operatorname{rank} A = 2$.

[Answer: (a) $a = b = c$; (b) at least two of a, b, c are distinct.]

11. Let S be a subspace of F^n with $\dim S = m$. If $X_1, \ldots, X_m$ are vectors in S with the property that $S = \langle X_1, \ldots, X_m \rangle$, prove that $X_1 \ldots, X_m$ form a basis for S.

12. Find a basis for the subspace S of $\mathbb{R}^3$ defined by the equation

$$x + 2y + 3z = 0.$$

Verify that $Y_1 = [-1,\ -1,\ 1]^t \in S$ and find a basis for S which includes Y_1.

13. Let $X_1, \ldots, X_m$ be vectors in F^n. If $X_i = X_j$, where $i < j$, prove that $X_1, \ldots X_m$ are linearly dependent.

14. Let $X_1, \ldots, X_{m+1}$ be vectors in F^n. Prove that

$$\dim \langle X_1, \ldots, X_{m+1} \rangle = \dim \langle X_1, \ldots, X_m \rangle$$

if X_{m+1} is a linear combination of $X_1, \ldots, X_m$, but

$$\dim \langle X_1, \ldots, X_{m+1} \rangle = \dim \langle X_1, \ldots, X_m \rangle + 1$$

if X_{m+1} is not a linear combination of $X_1, \ldots, X_m$.

Deduce that the system of linear equations $AX = B$ is consistent, if and only if

$$\operatorname{rank} [A|B] = \operatorname{rank} A.$$

15. Let $a_1, \ldots, a_n$ be elements of F, not all zero. Prove that the set of vectors $[x_1, \ldots, x_n]^t$ where $x_1, \ldots, x_n$ satisfy

$$a_1 x_1 + \cdots + a_n x_n = 0$$

is a subspace of F^n with dimension equal to $n - 1$.

16. Prove Lemma 3.2.1, Theorem 3.2.1, Corollary 3.2.1 and Theorem 3.3.2.

17. Let R and S be subspaces of F^n, with $R \subseteq S$. Prove that

$$\dim R \leq \dim S$$

and that equality implies $R = S$. (This is a very useful way of proving equality of subspaces.)

18. Let R and S be subspaces of F^n. If $R \cup S$ is a subspace of F^n, prove that $R \subseteq S$ or $S \subseteq R$.

19. Let $X_1, \ldots, X_r$ be a basis for a subspace S. Prove that all bases for S are given by the family $Y_1, \ldots, Y_r$, where

$$Y_i = \sum_{j=1}^{r} a_{ij} X_j,$$

and where $A = [a_{ij}] \in M_{r \times r}(F)$ is a non–singular matrix.