2. VECTOR SPACES

This chapter introduces the underlying algebraic structure of linear algebra – that of a finite dimensional vector space. The definition of a vector space involves an arbitrary field whose elements are called *scalars*. The following notation will be used (unless otherwise stated or implied):

$\mathbb{K}$	the field of scalars
$a,b,c,\ldots$	the elements of $\mathbb{K}$
V	the given vector space
$\mathbf{u},\mathbf{v},\mathbf{w}$	the elements of V

Nothing essential is lost if the reader assumes that $\mathbb{K}$ is the real field $\mathbb{R}$ or the complex field $\mathbb{C}$.

2.1 Vector spaces and subspaces

The following defines the notion of *vector space* or *linear space*.

Definition 2.1

Let $\mathbb{K}$ be a given field ($\mathbb{R}$ or $\mathbb{C}$) and let V be a non empty set with rules of addition, and scalar multiplication which assigns to any $\mathbf{u},\mathbf{v} \in V$ a sum $\mathbf{u}+\mathbf{v} \in V$ and to any $\mathbf{u} \in V$ and scalar $k \in \mathbb{K}$ a *product* $k\mathbf{u} \in V$. Then V is called a *vector space over* $\mathbb{K}$ (and the elements of V are called *vectors*) if the following axioms hold:

(A1) For any vectors $\mathbf{u},\mathbf{v},\mathbf{w} \in V, (\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(A2) There is a vector in V, denoted by 0 and called the *zero vector*, for which $\mathbf{u}+0=\mathbf{u}$ for any vector $\mathbf{u} \in V$.

(A3) For each vector $\mathbf{u} \in V$ there is a vector in V, denoted by $-\mathbf{u}$, for which $\mathbf{u}+(-\mathbf{u})=0$.

(A4) For any vectors $\mathbf{u},\mathbf{v} \in V, \mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

(M1) For any scalar $k \in \mathbb{K}$ and any vectors $\mathbf{u},\mathbf{v} \in V, k(\mathbf{u}+\mathbf{v}) = k\mathbf{u}+k\mathbf{v}$.

(M2) For any scalars $a,b \in \mathbb{K}$ and any vector $\mathbf{u} \in V, (a+b)\mathbf{u} = a\mathbf{u}+b\mathbf{u}$.

(M3) For any scalars $a,b \in \mathbb{K}$ and any vector $\mathbf{u} \in V, (ab)\mathbf{u} = a(b\mathbf{u})$.

(M4) For the unit scalar $1 \in \mathbb{K}, 1\mathbf{u} = \mathbf{u}$ for any vector $\mathbf{u} \in V$.

The above axioms naturally split into two sets. The first four are only concerned with the additive structure of V. From these four axioms it follows that any sum of vectors of the form

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_m$$

requires no parentheses and does not depend upon the order of the summands.

Theorem 2.1

For any vector space V,

a. The zero vector $\mathbf{0} \in V$ is unique.

b. For any vector $\mathbf{u} \in V$, the *additive inverse* $-\mathbf{u} \in V$ is unique.

c. For any vectors $\mathbf{u},\mathbf{v},\mathbf{w} \in V$, if $\mathbf{u}+\mathbf{w}=\mathbf{v}+\mathbf{w}$ then $\mathbf{u}=\mathbf{v}$.

Proof. a. Suppose that $\hat{0} \in V$ is a vector for which $\mathbf{u}+\hat{0}=\mathbf{u}$ for any vector $\mathbf{u} \in V$. Then,

$$\begin{aligned} \hat{0} &= \hat{0}+0 && \text{(by A2)} \\ &= 0 && \text{(by the definition of } \hat{0}\text{).} \end{aligned}$$

Since $\hat{0}=\mathbf{0}$, the zero vector is unique.

b. Given a vector $\mathbf{u} \in V$, let $\mathbf{w} \in V$ be a vector for which $\mathbf{u}+\mathbf{w}=\mathbf{0}$. Then,

$$\begin{aligned} \mathbf{w} &= \mathbf{w}+0 && \text{(by A2)} \\ &= \mathbf{w}+\left[\mathbf{u}+(-\mathbf{u})\right] && \text{(by A3)} \\ &= \left[\mathbf{w}+\mathbf{u}\right]+(-\mathbf{u}) && \text{(by A1)} \\ &= 0+(-\mathbf{u}) && \text{(by the definition of } w\text{)} \\ &= -\mathbf{u} && \text{(by A2).} \end{aligned}$$

Since $\mathbf{w}=-\mathbf{u}$, the negative of $\mathbf{u}$ is unique.

c. For any vectors $\mathbf{u},\mathbf{v},\mathbf{w} \in V$, such that $\mathbf{u}+\mathbf{w}=\mathbf{v}+\mathbf{w}$ then

$$\begin{aligned}
\mathbf{u} &= \mathbf{u} + \mathbf{0} && \text{(by A2)} \\
&= \mathbf{u} + [\mathbf{w} + (-\mathbf{w})] && \text{(by A3)} \\
&= [\mathbf{u} + \mathbf{w}] + (-\mathbf{w}) && \text{(by A1)} \\
&= [\mathbf{v} + \mathbf{w}] + (-\mathbf{w}) && \text{(by } \mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}\text{)} \\
&= \mathbf{v} + [\mathbf{w} + (-\mathbf{w})] && \text{(by A1)} \\
&= \mathbf{v} + \mathbf{0} && \text{(by A3)} \\
&= \mathbf{v} && \text{(by A2).}
\end{aligned}$$

Therefore, $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$ implies $\mathbf{u} = \mathbf{v}$. ■

Also, *subtraction* is defined by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

On the other hand, the remaining four axioms are concerned with the "action" of the field $\mathbb{K}$ on V. Observe that the labelling of the axioms reflects this splitting. Using these additional axioms we prove the following simple properties of a vector space.

Theorem 2.2

Let V be a vector space over a field $\mathbb{K}$.

a. For any scalar $k \in \mathbb{K}$ and $\mathbf{0} \in V, k\mathbf{0} = \mathbf{0}$.

b. For $0 \in \mathbb{K}$ and any vector $\mathbf{u} \in V$, then $0\mathbf{u} = \mathbf{0}$.

c. For any $k \in \mathbb{K}$ and any $\mathbf{u} \in V, (-k)\mathbf{u} = k(-\mathbf{u}) = -(k\mathbf{u})$.

Proof. a. We compute:

$$k\mathbf{0} + \mathbf{0} = k\mathbf{0} = k(\mathbf{0} + \mathbf{0}) = k\mathbf{0} + k\mathbf{0}$$

But $k\mathbf{0} + \mathbf{0} = k\mathbf{0} + k\mathbf{0}$ implies $k\mathbf{0} = \mathbf{0}$ (Theorem 2.1 part c).

b. Similarly, for any vector $\mathbf{u} \in V$,

$$0\mathbf{u} + \mathbf{0} = 0\mathbf{u} = (0 + 0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$$

which implies $0\mathbf{u} = \mathbf{0}$.

c. On one hand, for any $k \in \mathbb{K}$ and any $\mathbf{u} \in V$,

$$(-k)\mathbf{u} = (-k)\mathbf{u} + k\mathbf{u} - (k\mathbf{u}) = (-k + k)\mathbf{u} - (k\mathbf{u}) = 0\mathbf{u} - (k\mathbf{u}) = -(k\mathbf{u})$$

Therefore, $(-k)\mathbf{u} = -(k\mathbf{u})$. On the other hand,

$$k(-\mathbf{u}) = k(-\mathbf{u}) + k\mathbf{u} - (k\mathbf{u}) = k(\mathbf{u} - \mathbf{u}) - (k\mathbf{u}) = k0 - (k\mathbf{u}) = -(k\mathbf{u})$$

Therefore, $k(-\mathbf{u}) = -(k\mathbf{u})$. ■

Observe that when $k = 1$ in part c of the previous theorem, we have $-\mathbf{u} = (-1)\mathbf{u}$ for each $\mathbf{u} \in V$. That is, multiplying $\mathbf{u}$ by -1 produces its additive inverse.

Now we list a number of important examples of vector spaces which will be used throughout this course.

Example 2.1

Let $\mathbb{K}$ be an arbitrary field. The notation $\mathbb{K}^n$ is frequently used to denote the set of all column matrices with n rows and entries which are elements in $\mathbb{K}$. Here $\mathbb{K}^n$ is viewed as a vector space over $\mathbb{K}$ where vector addition and scalar multiplication defined by

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

and

$$k\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} k & a_1 \\ k & a_2 \\ & \vdots \\ k & a_n \end{bmatrix}.$$

The zero vector in $\mathbb{K}^n$ is

$$0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and the inverse of a vector is defined by

$$-\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_n \end{bmatrix}.$$

The proof that $\mathbb{K}^n$ is a vector space is left to the reader, so we now regard as stating that $\mathbb{R}^n$ with the usual operations is a vector space over $\mathbb{R}$.

Example 2.2

The notation $\mathbf{M}_{m,n}$, or simply $\mathbf{M}$, will be used to denote the set of all $m \times n$ matrices over an arbitrary field $\mathbb{K}$. Then $\mathbf{M}_{m,n}$ is a vector space over $\mathbb{K}$ with respect to the usual operations of matrix addition and scalar multiplication.

Example 2.3

Let $\mathbf{P}$ denote the set of all polynomials

$$a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n,\ n = 0,1,2,3,\ldots,$$

with coefficients a_i is some field $\mathbb{K}$. Then $\mathbf{P}$ is a vector space over $\mathbb{K}$ with respect to the usual operations of addition of polynomials and multiplication of polynomials by constants.

Example 2.4

Let X be any non empty set and let $\mathbb{K}$ be an arbitrary field $\mathbb{K}$. Consider the set $\mathbf{F}(X)$ of all functions from X into $\mathbb{K}$. The sum of two functions $f, g \in \mathbf{F}(X)$ is the function $f + g \in \mathbf{F}(X)$ defined by

$$(f+g)(x) = f(x) + g(x),\ \forall x \in X,$$

and the product of a scalar $k \in \mathbb{K}$ and a function $f \in \mathbf{F}(X)$ is the function $kf \in \mathbf{F}(X)$ defined by

$$(kf)(x) = kf(x),\ \forall x \in X.$$

Then $\mathbf{F}(X)$ with the above operations is a vector space over $\mathbb{K}$.

The zero vector in $\mathbf{F}(X)$ is the zero function 0 which maps each $x \in X$ into $0 \in \mathbb{K}$, that is,

$$0(x) = 0,\ \forall x \in X.$$

Also, for any function $f \in \mathbf{F}(X)$, the function $-f$ defined by

$$(-f)(x) = -f(x),\ \forall x \in X,$$

is the additive inverse of the function f.

Let W be a subset of a vector space V over a field $\mathbb{K}$. Then W is called a *subspace* of V if W is itself a vector space over $\mathbb{K}$ with respect to the operations of vector addition and scalar multiplication on V. Simple criteria for identifying subspaces follow.

Theorem 2.3

Suppose that W is a subset of a vector space V. Then W is a subspace of V if and only if the following hold:

a. $0 \in W$

b. W is *closed under vector addition*, that is:

For every $\mathbf{u},\mathbf{v} \in W$, the sum $\mathbf{u}+\mathbf{v} \in W$.

c. W is *closed under scalar multiplication*, that is:

For every $\mathbf{u} \in W, k \in \mathbb{K}$, the product $k\mathbf{u} \in W$.

Conditions b and c may be combined into one condition.

Corollary 2.1

W is a subspace if and only if

a. $0 \in W$

b. W is *closed under linear combinations*, that is:

$a\mathbf{u}+b\mathbf{v} \in W$ for every $\mathbf{u},\mathbf{v} \in W$ and $a,b \in \mathbb{K}$

Example 2.5

Let V be any vector space. Then the set $\{0\}$ consisting of the zero vector alone, and also the entire space V are subspaces of V.

Example 2.6

Let W be the set of vectors in $\mathbb{R}^3$ consisting of those vectors whose third component is 0; or, in other words

$$W = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a,b, \in \mathbb{R} \right\}$$

Notice that $0 \in W$. Further, for any vectors

$$\begin{bmatrix} a \\ b \\ 0 \end{bmatrix}, \begin{bmatrix} c \\ d \\ 0 \end{bmatrix} \in W,$$

and scalars $r,s \in \mathbb{R}$, we have

$$r\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} + s\begin{bmatrix} c \\ d \\ 0 \end{bmatrix} = \begin{bmatrix} ra+sc \\ rb+sd \\ 0 \end{bmatrix} \in W$$

Thus W is a subspace of $\mathbb{R}^3$.

Example 2.7

Let $V = \mathbf{M}_{n,n}$, the space of $n \times n$ matrices. Then the subset W_1 of (upper) triangular matrices and the subset W_2 of matrices satisfying $A^\top = A$, called the set of symmetric matrices, are subspaces of V since they are non empty and closed under linear combinations.

Example 2.8

Recall that $\mathbf{P}$ denotes the vector space of polynomials. Let $\mathbf{P}_n$ denote the subset of $\mathbf{P}$ that consists of all polynomials of degree $\leq n$, for a fixed n. Then $\mathbf{P}_n$ is a subspace of $\mathbf{P}$.

Recall that any solution of a homogeneous system $A\mathbf{x} = 0$. is an element of $\mathbb{R}^n$. Thus, the solution set of $A\mathbf{x} = 0$. is a subset of $\mathbb{R}^n$.

Theorem 2.4

The solution set W of a homogeneous system $A\mathbf{x} = 0$ in n unknowns is a subset of $\mathbb{R}^n$.

Proof. First, $0 \in W$ because 0 is always a solution of a homogeneous system: $A0 = 0$.

Next we show that W is closed under linear combinations. So let $\mathbf{u},\mathbf{v} \in W$ and $a,b \in \mathbb{R}$. Since $\mathbf{u}$ and $\mathbf{v}$ are solutions of the homogeneous system, we have $A\mathbf{u} = 0$ and $A\mathbf{v} = 0$. Moreover,

$$A\left(a\mathbf{u} + b\mathbf{v}\right) = aA\mathbf{u} + bA\mathbf{v} = a0 + b0 = 0$$

Consequently, $a\mathbf{u} + b\mathbf{v}$ is a solution of the homogeneous solution and, thus, $a\mathbf{u} + b\mathbf{v} \in W$. ■

2.2 Linear combinations and spans

Let V be a vector space over a field $\mathbb{K}$ and let $\mathbf{v}_1,\ldots,\mathbf{v}_m \in V$. Any vector in V of the form

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m$$

where $a_i \in \mathbb{K}$, is called a *linear combination* of $\mathbf{v}_1,\ldots,\mathbf{v}_m$. The set of all such linear combinations, denoted by

$$\text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$$

is called the *span* of $\mathbf{v}_1,\ldots,\mathbf{v}_m$.

Generally, for any subset S of V, Span $S = \{0\}$ when S is empty and Span S consists of all the linear combinations of vectors in S.

Theorem 2.5

If $\mathbf{v}_1,\ldots,\mathbf{v}_m \in V$ then $\text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ is a subspace of V.

Proof. Since $0 \in V$ can be written as a linear combination of $\mathbf{v}_1,\ldots,\mathbf{v}_m$:

$$0 = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n$$

then $0 \in \text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$. To show that $\text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ is closed under linear combinations, choose any two vectors $\mathbf{u},\mathbf{v} \in \text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$. Therefore,

$$\begin{aligned}\mathbf{u} &= a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m,\\ \mathbf{v} &= b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_m\mathbf{v}_m\end{aligned}$$

for some scalars $a_i, b_i \in \mathbb{K}$. Then, for any $r,s \in \mathbb{K}$,

$$\begin{aligned} r\mathbf{u} + s\mathbf{v} &= r\left(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m\right) + s\left(= b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_m\mathbf{v}_m\right)\\ &= \left(ra_1 + sb_1\right)\mathbf{v}_1 + \left(ra_2 + sb_2\right)\mathbf{v}_2 + \cdots + \left(ra_m + sb_m\right)\mathbf{v}_m\ .\end{aligned}$$

Then $r\mathbf{u} + s\mathbf{v}$ is a linear combination of the vectors $\mathbf{v}_1,\ldots,\mathbf{v}_m$, and, consequently, $r\mathbf{u} + s\mathbf{v} \in \text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$. ■

On the other hand, given a vector space V, the vectors $\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_r$ are said to *span* or *generate* or to form a *generating set* of V if

$$V = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}.$$

In other words, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ span V if, for every $\mathbf{v} \in V$, there are scalars $a_1, a_2, \ldots, a_r$ such that

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_r\mathbf{u}_r$$

that is, if $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$.

Example 2.9

Let

$$H = \left\{ \begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

Show that H is a subspace of $\mathbb{R}^4$.

In light of Theorem 2.5, if we can write H as the span of a set of vectors, then we can conclude that H is a subspace of $\mathbb{R}^4$. Hence, write a generic element of H as a linear combination of vectors:

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \ a, b \in \mathbb{R}.$$

This means that a generic element of H can always be expressed as a linear combination of the two vectors on the right. Therefore,

$$H = \text{Span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

So H is a subspace by Theorem 2.5.

Example 2.10

For what value(s) of h will the polynomial $q(t)$ be in the subspace of $\mathbf{P}_2$ spanned by $p_1(t), p_2(t), p_3(t)$, if

$$p_1(t) = 1 - t - 2t^2,\ p_2(t) = 5 - 4t - 7t^2,\ p_3(t) = -3 + t,\ q(t) = -4 + 3t + ht^2.$$

The polynomial $q(t)$ is in the subspace spanned by $p_1(t), p_2(t), p_3(t)$ if and only if there are scalars $a_1, a_2, a_3 \in \mathbb{R}$ such that

$$q(t) = a_1 p_1(t) + a_2 p_2(t) + a_2 p_3(t).$$

The above equality holds if and only if the coefficients multiplying $1, t$, and t^2, respectively, on the left side of the equation are equal to the correspon-ding coefficients on the right side. Therefore, comparing the corresponding coefficients we get,

$$\begin{aligned} 1: \quad -4 &= a_1 + 5a_2 - 3a_3, \\ t: \quad 3 &= -a_1 - 4a_2 + a_3, \\ t^2: \quad h &= -2a_1 - 7a_2 + 0a_3, \end{aligned}$$

or, in vector form,

$$\begin{bmatrix} 1 & 5 & -3 \\ -1 & -4 & 1 \\ -2 & -7 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}.$$

We study the solution set of this equation by reducing the augmented matrix to its row reduced echelon form:

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 7 & 1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}.$$

Observe that the third row in the last matrix corresponds to the degenerate equation $0 = h - 5$, which has a solution if and only if $h = 5$. Therefore, $q(t) \in \operatorname{Span}\{p_1(t), p_2(t), p_3(t)\}$ if and only if $h = 5$.

2.3 Null space and column space of a matrix

There are two important subspaces associated with any $m \times n$ matrix A. In this section, we introduce both of these subspaces and study some of their properties.

Definition 2.2

The null space of an $m \times n$ matrix A, written as $\mathrm{Nul}A$, is the set of all solutions of the homogeneous equation $A\mathbf{x} = 0$. In set notation,

$$\mathrm{Nul}A = \left\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = 0\right\}$$

A more dynamic description of $\mathrm{Nul}A$ is the set of all $\mathbf{x} \in \mathbb{R}^n$ that are *annihilated* by A in the sense that each $\mathrm{x} \in \mathrm{Nul}A$ is "transformed" into the zero vector.

The term *space* in *null space* is appropriate because the null space of an $m \times n$ matrix is a vector subspace of $\mathbb{R}^n$. This is a direct consequence of Theorem 2.4.

Example 2.11

Let H be the set of all vectors in $\mathbb{R}^4$ whose coordinates a,b,c,d satisfy the equations

$$\begin{aligned} a - 2b + 5c - d &= 0, \\ -a - b + c &= 0. \end{aligned}$$

In other words, each element $\mathbf{x}$ of H satisfies the homogeneous equation

$$\begin{bmatrix} 1 & -2 & 5 & -1 \\ -1 & -b & 1 & 0 \end{bmatrix}\mathbf{x} = 0.$$

So by the definition of the null space of a matrix, we have the following expression of H:

$$H = \mathrm{Nul}\begin{bmatrix} 1 & -2 & 5 & -1 \\ -1 & -b & 1 & 0 \end{bmatrix}.$$

Example 2.12

Find a generating set for the null space of the matrix

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}.$$

The first step is to find the general solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. So we reduce the augmented matrix to its row reduced echelon form:

$$\begin{bmatrix} A & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then, the solution set of $A\mathbf{x} = 0$ is the same as the solution set of the simpler system

$$\begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0, \\ x_3 + 2x_4 - 2x_5 &= 0, \\ 0 &= 0, \end{aligned}$$

with free variables x_2, x_4, x_5. If we write the leading variables in terms of the free variables, then we can express the general solution of $A\mathbf{x} = 0$ as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2a + b - 3c \\ a \\ -2b + 2c \\ b \\ c \end{bmatrix} = a\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

This means that every solution of the homogeneous equation can be expressed as a linear combination of the three vectors on the right. Then, the generating set of $\mathrm{Nul}A$ is

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Using the notation of this and the previous section, we may write

$$\mathrm{Nul}A = \mathrm{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Observation 2.1

The thoughtful reader would observe that the number of non zero vectors in the generating set of Nul A equals the number of free variables in the system $A\mathbf{x} = 0$.

Another important subspace associated with a matrix is its column space. Unlike the null space, the column space is defined explicitly via linear combinations.

Definition 2.3

The *column space* of an $m \times n$ matrix A, written as $\text{Col}A$, is the set of all linear combinations of the columns of A. If $A = \left[\mathbf{a}_1 \ldots \mathbf{a}_n\right]$, then

$$\text{Col}A = \text{Span}\left\{\mathbf{a}_1, \ldots, \mathbf{a}_n\right\}.$$

The fact that $\text{Col}A$ is a subspace of $\mathbb{R}^m$ follows from Theorem 2.5. Note that by the definition of matrix multiplication, a typical vector in $\text{Col}A$ can be written as $A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$. That is,

$$\text{Col}A = \left\{\mathbf{b} \in \mathbb{R}^m : \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\right\}.$$

Example 2.13

Find a matrix A such that $W = \text{Col}A$, where

$$W = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

We first find a generating set of W and use the vectors in the generating set as columns of the matrix A.

$$W = \left\{ a\begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \in \mathbb{R} \right\} = \text{Span}\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

The matrix A whose columns are the vectors in the generating set of W is

$$A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ 0 & -7 \end{bmatrix}.$$

Then $W = \text{Col}A$.

The following example is meant to emphasize the differences between the null space and the column space of a matrix.

Example 2.14

Let $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$.

a. If the column space of A is a subspace of $\mathbb{R}^k$, what is k?

b. If the null space of A is a subspace of $\mathbb{R}^k$, what is k?

c. Find a non zero vector of $\mathrm{Col}A$ and a non zero vector of $\mathrm{Nul}A$?

d. Is there a vector in $\mathrm{Col}A$ that also belongs to $\mathrm{Nul}A$?

The answers to these questions are:

a. Since the columns of A belong to $\mathbb{R}^3$, then they must generate a subspace of $\mathbb{R}^3$. So $k = 3$.

b. The matrix A has four columns, which means that the product $A\mathbf{x}$ is defined if and only if $\mathbf{x} \in \mathbb{R}^4$. Therefore, the general solution of the homogeneous equation $A\mathbf{x} = 0$ belongs to $\mathbb{R}^4$. Hence, $k = 4$.

c. Finding a non zero element of $\mathrm{Col}A$ is easy. Clearly, any column of A belongs to $\mathrm{Col}A$. Then it is sufficient to choose any column, say, the first one: $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$.

Finding a non zero element of $\mathrm{Nul}A$ is harder. For this, we need to solve the homogeneous equation $A\mathbf{x} = \mathbf{0}$. Reduce the augmented matrix to row reduced echelon form:

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

which corresponds to the system of equations

$$\begin{aligned} x_1 + 9x_3 &= 0, \\ x_2 - 5x_3 &= 0, \\ x_4 &= 0, \end{aligned}$$

with x_3 free. The general solution of the homogeneous equation is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9\lambda \\ 5\lambda \\ \lambda \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}.$$

Then, a no zero element of the null space of A is $\begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}$.

d. Each element of $\mathrm{Col}A$ is a vector in $\mathbb{R}^3$ and each element of $\mathrm{Nul}A$ is a vector in $\mathbb{R}^4$. Thus, with only three entries, the elements of $\mathrm{Col}A$ could not possible be in $\mathrm{Nul}A$.

Table 1: Contrast between $\mathrm{Nul}A$ an $\mathrm{Col}A$ of an $m \times n$ matrix A.

Nul$A$	**Col$A$**
1. Nul A is a subspace of $\mathbb{R}^n$.	1. Col$A$ is a subspace of $\mathbb{R}^m$.
2. Nul$A$ is defined implicitly via the condition ($A\mathbf{x} = 0$) that its elements must satisfy.	2. Col$A$ is defined explicitly via how to construct its elements.
3. A typical vector $\mathbf{v}$ is Nul$A$ has that property that $A\mathbf{v} = 0$.	3. A typical vector $\mathbf{v}$ is Col$A$ has the property that $A\mathbf{x} = \mathbf{v}$ is consistent.
4. Given a specific vector $\mathbf{v}$, it is to tell if $\mathbf{v}$ is in Nul$A$. Just compute $A\mathbf{v}$.	4. Given a specific vector $\mathbf{v}$, it takes time to tell if $\mathbf{v}$ is in Col$A$. The solutions of $A\mathbf{x} = \mathbf{v}$ must be studied.
5. Nul$A = \{0\}$ if and only if the equation $A\mathbf{x} = 0$ has a unique solution.	5. Col$A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{R}^m$.

2.4 Linear independence

We now define the notion of linear independence. This concept plays an essential role in the theory of linear algebra and in mathematics in general.

Definition 2.4

Let V be a vector space over a field $\mathbb{K}$. The vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are said to be *linearly independent* if the vector equation

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = \mathbf{0} \tag{2.1}$$

has a unique solution: $c_1 = c_2 = \cdots = c_m = 0$. The vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are said to be *linearly dependent* if (2.1) has a non trivial solution.

If $\mathbf{0}$ is one of the vectors in (2.1), say $\mathbf{u}_1$, then the vectors must be linearly dependent; for

$$1\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + 0\mathbf{u}_m = 1\mathbf{0} + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0},$$

which means that(2.1) has the non trivial solution $c_1 = 1, c_2 = \cdots = c_m = 0$.

Any non zero vector $\mathbf{u}$ is, by itself, linearly independent; for the equation $c\mathbf{u} = \mathbf{0}$ has the unique solution $c = 0$.

If two of the vectors in (2.1) are equal or one is a scalar multiple of the other, say $\mathbf{u}_1 = k\mathbf{u}_2$, then the vectors are linearly dependent. For, in this case,

$$\mathbf{u}_1 - k\mathbf{u}_1 + 0\mathbf{u}_3 + \cdots + 0\mathbf{u}_m = \mathbf{0},$$

which means that (2.1) has the non trivial solution $c_1 = 1, c_2 = -k, c_3 = \cdots = c_m = 0$. In particular, two vectors are linearly dependent if and only if one is a multiple of the other.

Clearly, if the vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are linearly independent, then any rearrangement of these vectors is also linearly independent.

If the vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are linearly independent, then any subset of these vectors, say $\mathbf{u}_1,\ldots,\mathbf{u}_r (r < m)$, is linearly independent. Indeed, consider the equation

$$d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_r\mathbf{u}_r = \mathbf{0}.$$

But we can also write

$$d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_r\mathbf{u}_r + 0\mathbf{u}_{r+1} + \cdots + 0\mathbf{u}_m = \mathbf{0}.$$

Since $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are linearly independent, this equation has only the trivial solution, which means that $d_1 = \cdots = d_r = 0$ and, thus, $\mathbf{u}_1,\ldots,\mathbf{u}_r$ are linearly independent. By contrapositive (see the Introduction), if some subset of $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$, say $\mathbf{u}_1,\ldots,\mathbf{u}_r (r < m)$, is linearly dependent, then $\mathbf{u}_1,\ldots,\mathbf{u}_m$ is linearly dependent.

The following theorem will often be useful. It states that a linearly dependent list of vectors, with the first vector not $\mathbf{0}$, one of the vectors is in the span of the previous ones.

Theorem 2.6

An indexed set $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ of two or more vectors, with $\mathbf{u}_1 \neq \mathbf{0}$, is linearly dependent if and only if some $\mathbf{u}_j (j > 1)$ is a linear combination of the preceding vectors $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1}$.

Proof. Suppose that, for some $j > 1, \mathbf{u}_j$ is a linear combination of $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1}$. Then there are scalars $d_1,\ldots,d_{j-1}$, such that

$$\mathbf{u}_j = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_{j-1}\mathbf{u}_{j-1},$$

or, equivalently,

$$-d_1\mathbf{u}_1 - \cdots - d_{j-1}\mathbf{u}_{j-1} + \mathbf{u}_j + 0\mathbf{u}_{j+1} + \cdots + 0\mathbf{u}_m = \mathbf{0}.$$

Therefore, (2.1) has the non trivial solution $c_1 = -d_1,\ldots,c_{j-1} = -d_{j-1}, c_j = 1, c_{j+1} = \cdots = c_m = 0$, which means that $\mathbf{u}_1,\ldots,\mathbf{u}_m$ are linearly dependent.

On the other hand, suppose that $\mathbf{u}_1,\ldots,\mathbf{u}_m$ are linearly dependent. Then there are scalars $a_1, a_2,\ldots,a_m$, not all equal to 0, such that

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_m\mathbf{u}_m = \mathbf{0}.$$

Not all $a_2, a_3,\ldots,a_m$ can be equal to 0. Otherwise, we would have $a_1\mathbf{u}_1 = \mathbf{0}$; but $\mathbf{u}_1 \neq \mathbf{0}$ which means that $a_1 = 0$. Then, in this case, $a_1 = \cdots = a_m = 0$ making the vectors linearly independent which contradicts our assumption. Let j be the largest element of $\{2,\ldots,m\}$ such that $a_j \neq 0$. Then,

$$\mathbf{u}_j = -\frac{a_1}{a_j}\mathbf{u}_1 - \cdots - \frac{a_{j-1}}{a_j}\mathbf{u}_{j-1}.$$

That is, there is some $\mathbf{u}_j$ that is a linear combination of the preceding vectors $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1}$. ■

Now we come to a key result. It says that, if in a set of vectors one can be written as a linear combination of the remaining ones, then we can throw out that vector without changing the span of the original list.

Theorem 2.7

Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a set in a vector space V, and let $H =$ Span S. If one of the vectors - say $\mathbf{u}_k$ - is a linear combination of the remaining vectors in S, then the set formed from S by removing $\mathbf{u}_k$ still spans H.

Proof. By rearranging the list of vectors in S, if necessary, we may suppose that $\mathbf{u}_m$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_{m-1}$, say

$$\mathbf{u}_m = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_{m-1}\mathbf{u}_{m-1}$$

Given any $\mathbf{w} \in H,$ we may write

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m$$

for suitable scalars $c_1, c_2, \ldots, c_m$. Substituting the expression for $\mathbf{u}_m$ we obtain

$$\mathbf{w} = (c_1 + c_m a_1)\mathbf{u}_1 + (c_2 + c_m a_2)\mathbf{u}_2 + \cdots + (c_{m-1} + c_m a_{m-1})\mathbf{u}_{m-1}.$$

Thus, $\mathbf{w}$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_{m-1}$. This means that $\{\mathbf{u}_1, \ldots, \mathbf{u}_{m-1}\}$ spans H because $\mathbf{w}$ was an arbitrary element of $H.$ ■

Consider a set of vectors $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ in a vector space V. We now know that for any vector $\mathbf{w} \in V$, the equation

$$w = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_m\mathbf{u}_m, \tag{2.2}$$

has a solution if and only if $x \in \text{Span}S$. So the notion of generating set deals with the existence of solutions of vector equations such as (2.2). So, what about uniqueness? It turns out that vector equations such as (2.2) with unique solutions can be characterized in terms of the notion of linear independence.

Theorem 2.8

Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a set in a vector space V. Then each $\mathbf{w} \in \text{Span}S$ has a unique expression

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m,$$

for suitable scalars $c_1, \ldots, c_m$ if and only if S is a linearly independent set.

Proof. Suppose that each $\mathbf{w} \in \text{Span}S$ has a unique expression as a linear combination of the vectors in S. In particular, $0 \in \text{Span}S$ so

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = 0,$$

which clearly is satisfied with $c_1 = c_2 = \ldots = c_m = 0$. But this expression is unique which means that S is a linearly independent set of vectors.

Now suppose that S is a linearly independent set. If a vector $\mathbf{w} \in \text{Span}S$ can be expressed as

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m,$$

and

$$\mathbf{w} = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_m\mathbf{u}_m,$$

then subtracting the expressions we get

$$0 = (c_1 - d_1)\mathbf{u}_1 + (c_2 - d_2)\mathbf{u}_2 + \cdots + (c_m - d_m)\mathbf{u}_m.$$

But S is a linearly independent set, so the last homogeneous equations has the unique solution $c_1 - d_1 = c_2 - d_2 = \ldots = c_m - d_m = 0$. Clearly, this means that $c_1 = d_1, c_2 = d_2, \ldots, c_m = d_m$. ■

Example 2.15

Consider the three vectors in $\mathbb{R}^3$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Clearly, the homogeneous equation

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

has only the trivial solution $x_1 = x_2 = x_3 = 0$. Therefore, $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. are independent vectors. In fact, since it is obvious that $\mathbb{R}^3 = \text{Span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, we conclude that each vector in $\mathbb{R}^3$ has a unique expression as linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

Example 2.16

Suppose that the scalars $c_0, c_1, \ldots, c_n$ satisfy

$$c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n = 0$$

This equality means that the polynomial on the left has the same values as the zero polynomial on the right. It is well known that a polynomial of degree n with more than n roots is the zero polynomial. Therefore, $c_0 = c_1 = \cdots = c_n = 0$. This proves that the set $\{1, t, t^2, \ldots, t^n\}$ is a linearly independent set in $\mathbf{P}_n$. It is obvious that $\mathbf{P}_n = \text{Span}\{1, t, t^2, \ldots, t^n\}$. Then each polynomial of degree at most n has a unique expression as a linear combination of $1, t, t^2, \ldots, t^n$.

Recall that any linear dependence relationship among the columns of a matrix A can be expressed in the form $A\mathbf{x} = 0$, where $\mathbf{x}$ is a column of weights. When A is row reduced to a matrix B, the columns of B are often totally different from the columns of A. However, the equations $A\mathbf{x} = 0$ and $B\mathbf{x} = 0$ have exactly the same set of solutions. That is, the columns of A have *exactly the same linear dependence relationships as the columns of B*. So we have the following important observation:

Observation 2.2

Elementary row operations on a matrix preserve the linear dependence relations among the columns of the matrix.

The previous observation provides a useful shortcut for reducing a matrix to its row reduced echelon form when the linear dependence relations among the columns of a matrix are easy to detect.

Example 2.17

Consider the matrix

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}.$$

a. Find the smallest generating set for $\text{Col}A$.

b. Use the linear dependence relations among the columns of A for compute its row reduced echelon form.

These are the solutions:

a. Recall that $\mathrm{Col}A = \mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5\}$. By Theorem 2.7, the set of columns formed by removing those columns that are linear combinations of the remaining ones still spans $\mathrm{Col}A$. Then, since

$$\mathbf{a}_2 = 4\mathbf{a}_1 \text{ and } \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3,$$

we can remove $\mathbf{a}_2$ and $\mathbf{a}_4$ from the spanning set. Therefore, $\mathrm{Col}A = \mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$. Since $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ are linearly independent (you should show this), no more columns should be removed. This means that $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ is the smallest generating set of $\mathrm{Col}A$.

b. The columns of the row reduced echelon form $B = [\mathbf{b}_1 \mathbf{b}_2 \mathbf{b}_3 \mathbf{b}_4 \mathbf{b}_5]$ should have the same linear dependence relations as the columns of A. Therefore,

$$\mathbf{b}_2 = 4\mathbf{b}_1,\ \mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3,$$

and $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$ are linearly independent. The pivot columns of a matrix in row reduced echelon form are linearly independent columns since they are the columns of the identity matrix. Moreover, all the rows of zeros should be at the bottom of the row reduced echelon matrix. From these two facts we deduce that $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$ are pivots and correspond to the first three columns of the identity matrix, respectively. Putting everything together, we obtain

$$A \sim \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

2.5 Bases and dimension

In a vector space V, sets that guarantee that each vector in V con be expressed uniquely in terms of these sets are one of the most important objects in linear algebra.

Definition 2.5

An ordered set $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of vectors is a *basis* of V if the following two conditions hold:

1. $\mathcal{B}$ is a linearly independent set.
2. $V = \mathrm{Span}\mathcal{B}$.

The following characterization of a basis is a direct consequence of Theorem 2.8.

Theorem 2.9

An ordered set $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of vectors is a *basis* of V if and only if every vector $\mathbf{v} \in V$ can be written uniquely as a linear combination of the basis vectors.

We have already seen an example of a basis of $\mathbb{R}^3$ in Example 2.15. Such basis consisting of the columns of the identity matrix is called the *canonical basis* of $\mathbb{R}^3$.

Example 2.16 presents a basis of $\mathbf{P}_n$. Such basis consisting of consecutive powers of the variable t is called the *standard basis* of $\mathbf{P}_n$.

Let us consider less obvious examples of bases.

Example 2.18

Let

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}.$$

Determine if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$.

We must show that these vector are linearly independent and generate $\mathbb{R}^3$. Both of these conditions hold if and only if the equation $\begin{bmatrix} \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3 \end{bmatrix} \mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b} \in \mathbb{R}^3$. This equation has a solution for each $\mathbf{b}$ if each row of the matrix $A = \begin{bmatrix} \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3 \end{bmatrix}$ has a pivot and the solution is unique if each column of A is a pivot column. In other words, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ if $A \sim I$. So we row reduce A to its row reduced echelon form:

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

Therefore $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$.

Example 2.19

Let

$$p_0(t) = (1-t)^3, \ p_1(t) = 3t(1-t)^2, \ p_2(t) = 3t^2(1-t), \ p_3(t) = t^3.$$

Determine if $\{p_0(t), p_1(t), p_2(t), p_3(t)\}$ is a basis for $\mathbf{P}_3$.

We need to determine if each polynomial in $\mathbf{P}_3$ can be expressed uniquely as a linear combination of $p_0(t), p_1(t), p_2(t), p_3(t)$. Let $q(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$ and suppose that the scalars c_0, c_1, c_2, c_3 satisfy

$$q(t) = c_0 p_0(t) + c_1 p_1(t) + c_2 p_2(t) + c_3 p_3(t).$$

Comparing the coefficients multiplying each power of t on both sides of the equality, we obtain

$$\begin{aligned} c_0 &= a_0, \\ -3c_0 + 3c_1 &= a_1, \\ 3c_0 - 6c_1 + 3c_2 &= a_2, \\ -c_0 + 3c_1 - 3c_2 + c_3 &= a_3, \end{aligned}$$

or in matrix form,

$$A\mathbf{c} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

It is not difficult to see that the matrix A has a pivot in every row and every column, which means that the equation above has a unique solution for each choice of the polynomial $q(t)$. Therefore, $\{p_0(t), p_1(t), p_2(t), p_3(t)\}$ is a basis for $\mathbf{P}_3$.

The following theorem characterizes an important basis of the column space of a matrix.

Theorem 2.10

The pivot columns of a matrix A form a basis for $\mathrm{Col}A$.

Proof. Let B be the reduced echelon form of A. The set of pivot columns of B is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since A is row equivalent to B, the pivot columns of A are linearly independent as well, because any linear dependence relation among the columns of A corresponds to a linear dependence relation among the columns of B. For this same reason, every non pivot column of A is a linear combination of the pivot columns of A. Thus the non pivot columns of A may be discarded from the spanning set for $\mathrm{Col}A$. This leaves the pivot columns of A as a basis for $\mathrm{Col}A$. ■

Warning: Be careful to use *pivot columns of A itself* for the basis of $\mathrm{Col}A$. Row operations can change the column space of a matrix. We illustrate this in the following example.

Example 2.20

Consider again the matrix given in Example 2.17:

$$A=\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5\end{bmatrix}=\begin{bmatrix}1 & 4 & 0 & 2 & -1\\ 3 & 12 & 1 & 5 & 5\\ 2 & 8 & 1 & 3 & 2\\ 5 & 20 & 2 & 8 & 8\end{bmatrix}.$$

We showed that

$$A\sim\begin{bmatrix}1 & 4 & 0 & 2 & 0\\ 0 & 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}.$$

It can be seen from the row reduced echelon form that the first, third, and last columns are pivot columns, which means that $\{\mathbf{a}_1,\mathbf{a}_3,\mathbf{a}_5\}$ is a basis for $\mathrm{Col}A$. However, the columns of the row reduced echelon form all have zeros in their last entries, so they cannot span the column space of A.

Every generating set in a vector space can be reduced to a basis of the vector space. Indeed, removing vectors that are linear combinations of the remaining vectors will not change the span of the original set. The deletion of vectors from the generating set must stop when the set becomes linearly independent. If additional vectors are deleted, it will not be a linear combination of the remaining vectors, and hence the smaller set will no longer span V. Thus a basis is a generating set that is as small as possible.

Also, every linearly independent list of vectors can be extended to a basis of the vector space. This is done by adjoining a vector that is not in the span of the original list. The extended list of vectors is still linearly independent because the new vector cannot be written as a linear combination of the preceding ones. This step is repeated until a generating set of V is obtained. The addition of vectors must stop as soon as a generating set for V is obtained. For if one more vector is added, then the new set cannot be linearly independent because the old set spans V, and the new vector is therefore a linear combination of the vectors preceding it. Thus, a basis is a linearly independent set that is as large as possible.

Remark 2.1

Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis of a vector space V. Then every generating set of V has at least n elements, and every linearly independent set in V has at most n elements.

We say that a vector space V is *finite-dimensional* if it has a basis with a finite number of elements.

The notion of basis would be useless if different basis of the same vector space V had different numbers of elements. Fortunately, that turns out not to be the case.

Theorem 2.11

Any two bases of a finite-dimensional vector space have the same number of elements.

Proof. Suppose that V is a finite-dimensional vector space. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases of V. Then $\mathcal{B}_1$ and $\mathcal{B}_2$ are linearly independent in V, so the number of vectors in $\mathcal{B}_1$ is at most the number of elements in $\mathcal{B}_2$. Moreover, $\mathcal{B}_1$ and $\mathcal{B}_2$ span V, then the number of vectors in $\mathcal{B}_1$ is at least the number of elements in $\mathcal{B}_2$. Thus the number of vectors in $\mathcal{B}_1$ and $\mathcal{B}_2$ must be equal. ■

Now we know that the number of vectors in a basis of a finite-dimensional vector space is an inherent property of V (that is, it does not depend on the choice of basis). Therefore we give the following definition.

Definition 2.6

The dimension of a finite-dimensional vector space, denoted by $\dim V$, is the number of vectors in any basis of V. In other words, if $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a basis of V, then $\dim V = n$.

If V is a finite-dimensional vector space, then every generating set of vectors of V with $n = \dim V$ elements is a basis of V. For suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V. However, every basis of V has n elements, so in this case the reduction must be the trivial one, meaning that no elements are deleted from $\mathbf{v}_1, \ldots, \mathbf{v}_n$. In other words, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of V, as desired.

Moreover, every linearly independent set of vectors in V with n elements is a basis of V. For suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly independent set. Then this set can be extended to a basis of V. However, every basis of V has n elements, so in this case the extension must be the trivial one, meaning that no vectors are adjoined to $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. In other words, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V.

Example 2.21

The canonical basis for $\mathbb{R}^n$ contains n vectors, so $\dim\mathbb{R}^n = n$. The standard polynomial basis $\{1, t, t^2, \ldots, t^n\}$ shows that $\dim\mathbf{P}_n = n+1$.

Example 2.22

Find the dimension of the subspace

$$H = \left\{ \begin{bmatrix} a-3b+6c \\ 5a+4d \\ b-2c-d \\ 5d \end{bmatrix} : a,b,c,d \in \mathbb{R} \right\}.$$

The strategy is to produce a basis for H by finding a generating set and then reduce it to a linearly independent generating set. The number of elements in the final set is equal to $\dim H$.

First, we find a generating set for H. A generic vector in H can be written as

$$\begin{bmatrix} a-3b+6c \\ 5a+4d \\ b-2c-d \\ 5d \end{bmatrix} = a\begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix} + d\begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix},\ a,b,c,d \in \mathbb{R}.$$

Then

$$H = \operatorname{Span}\left\{ \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix} \right\}.$$

This is not a basis, however. Notice that the third vector is a multiple of the second vector, so it can be removed from the set. The remaining vectors are linearly independent. Thus, we have found a basis for H; namely,

$$\left\{ \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix} \right\}.$$

Since it has three vectors, we conclude that $\dim H = 3$.

Example 2.23

Find the dimensions of the null space and the column space of the matrix

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

The pivot columns of A form a basis of $\operatorname{Col} A$, so the dimension of the column space is the number of pivot columns. Row reduce the matrix to row reduced echelon form:

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

Therefore, $\dim \operatorname{Col} A = 2$.

Now we find a basis for $\operatorname{Nul} A$. If we row reduce the augmented matrix $\begin{bmatrix} A 0 \end{bmatrix}$ we obtain $\begin{bmatrix} B & 0 \end{bmatrix}$. The corresponding system of linear equations is:

$$\begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0, \\ x_3 + 2x_4 - 2x_5 &= 0, \\ 0 &= 0, \end{aligned}$$

with x_2, x_4, x_5 free. Its general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2a + b - 3c \\ a \\ -2b + 2c \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \ a, b, c \in \mathbb{R}.$$

We already know that the number of generating vectors produced with this method is equal to the number of free variables. Furthermore, this method produces automatically a set of linearly independent vectors because the free variables are the weights on the spanning vectors. For instance, look at the second, fourth, and fifth entries of the three vectors above, and note that $\mathbf{0}$ is a linear combination of these vectors if and only if $a = b = c = 0$. Therefore, these vectors form a basis for $\operatorname{Nul} A$ and $\dim \operatorname{Nul} A = 3$.

2.6 Coordinates

The purpose of bases in vector spaces is to provide a method of computation, and we are going to learn to use them in this section. We will consider two topics: how to express a vector in terms of a given basis, and how to relate two different bases of the same vector space.

Let V be a vector space with $\dim V = n$ over a field $\mathbb{K}$, and suppose that

$$\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

is a basis of V. Then any vector $\mathbf{v} \in V$ can be expressed uniquely as a linear combination of the basis vectors in $\mathcal{B}$, say

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n.$$

The n scalars $a_1, a_2, \ldots, a_n$ are called the *coordinates* of $\mathbf{v}$ relative to $\mathcal{B}$: and they form the vector

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

called the *coordinate vector* of $\mathbf{v}$ relative to $\mathcal{B}$.

Example 2.24

Let $\mathcal{E}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the canonical basis for $\mathbb{R}^3$. That is,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For any vector $\mathbf{x} \in \mathbb{R}^3$, it is obvious that $[\mathbf{x}]_{\mathcal{E}_3} = \mathbf{x}$.

Example 2.25

Let $\mathcal{S}_n = \{1, t, t^2, \ldots, t^n\}$ be the standard basis of $\mathbf{P}_n$. For any polynomial $q(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$, we have

$$[q(t)]_{\mathcal{S}_n} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Example 2.26

Let

$$\mathbf{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix},\ \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mathbf{u}_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

Find the coordinates of $\mathbf{v}$ relative to the basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$.

We must find scalars a_1, a_2 such that $\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$, or, in matrix form

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Row reduce the augmented matrix:

$$\begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -2 \end{bmatrix}.$$

Therefore,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}.$$

We want to emphasize the $\mathbf{v} \neq [\mathbf{v}]_{\mathcal{B}}$. The correct relation between $\mathbf{v}$ and its coordinates $[\mathbf{v}]_{\mathcal{B}}$ is $\mathbf{v} = 7\mathbf{u}_1 - 2\mathbf{u}_2$.

The following example is to remind you that the material in this course is applicable to every vector space, not just to the well-known ones like $\mathbb{R}^n$ and $\mathbf{P}_n$.

Example 2.27

Let us consider the vector space V generated by the set of linearly independent functions $\mathcal{B} = \left\{ e^{-i\pi t}, e^{-i\frac{\pi}{2}t}, 1, e^{i\frac{\pi}{2}t}, e^{i\pi t} \right\}$. Observe that $\mathcal{B}$ is a basis for V and $\dim V = 5$. Some examples of vector in V are

$$\cos\left(\frac{\pi n}{2}t\right),\ \sin\left(\frac{\pi n}{2}t\right),\ n = -2, -1, 0, 1, 2.$$

We will use Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta,\ \theta \in \mathbb{R},\ i = \sqrt{-1},$$

to find the coordinates of some of these vectors relative to $\mathcal{B}$.

By Euler's identity

$$\cos(\pi t) = \frac{e^{i\pi t} + e^{-i\pi t}}{2} \quad \text{and} \quad \sin\left(\frac{\pi}{2}t\right) = \frac{e^{i\frac{\pi}{2}t} - e^{-i\frac{\pi}{2}t}}{2i}.$$

Therefore,

$$[\cos(\pi t)]_{\mathcal{B}} = \begin{bmatrix} 1/2 \\ 0 \\ 0 \\ 0 \\ 1/2 \end{bmatrix} \quad \text{and} \quad \left[\sin\left(\frac{\pi}{2}t\right)\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ -\frac{1}{2i} \\ 0 \\ \frac{1}{2i} \\ 0 \end{bmatrix}.$$

Another example is

$$\left[1 + 2\cos(\pi t) - 2i\sin\left(\frac{\pi}{2}t\right)\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

The coordinates relative to any basis have the property that they preserve the main structure of vector spaces: linear combinations.

Theorem 2.12

Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis for a vector space V. For any $\mathbf{v}, \mathbf{w} \in V$ and scalar r:

1. $[\mathbf{v} + \mathbf{w}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}$
2. $[r\mathbf{v}]_{\mathcal{B}} = r[\mathbf{v}]_{\mathcal{B}}$

Proof. Take two typical vectors in V, say

$$\begin{aligned} \mathbf{v} &= c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n \\ \mathbf{w} &= d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_n\mathbf{u}_n \end{aligned}$$

Then, using vector operations,

$$\mathbf{v} + \mathbf{w} = (c_1 + d_1)\mathbf{u}_1 + (c_2 + d_2)\mathbf{u}_2 + \cdots + (c_n + d_n)\mathbf{u}_n.$$

It follows that

$$[\mathbf{v}+\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} c_1+d_1 \\ c_2+d_2 \\ \vdots \\ c_n+d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}.$$

If r is any scalar, then

$$r\mathbf{v} = (rc_1)\mathbf{u}_1 + (rc_2)\mathbf{u}_2 + \cdots + (rc_n)\mathbf{u}_n.$$

So

$$[r\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ rc_2 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = r[\mathbf{v}]_{\mathcal{B}}.$$

■

We now come to a very important computational method: *change of basis*. Identifying vectors in V with column vectors in $\mathbb{K}^n$ is useful when a natural basis is presented to us, but not when the given basis is poorly suited to the problem at hand. In that case, we will want to change basis. So let us suppose that we are given two bases for the same vector space V, say $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$. We will think of $\mathcal{B}$ as the *old* basis, and $\mathcal{C}$ as a *new* basis. There are two computations which we wish to clarify. We ask first: How are the two bases related? Secondly, a vector $\mathbf{v} \in V$ will have coordinates relative to each of these bases, but of course they will be different. So we ask: How are the two coordinate vectors related? These are the computations called change of basis.

We begin by noting that since the new basis spans V, every vector in the old basis $\mathcal{B}$ is a linear combination of the new basis $\mathcal{C}$. So we can write the coordinates of each vector in $\mathcal{B}$ relative to $\mathcal{C}$:

$$[\mathbf{b}_1]_{\mathcal{C}}, \quad [\mathbf{b}_2]_{\mathcal{C}}, \quad \cdots \quad [\mathbf{b}_n]_{\mathcal{C}}.$$

Now let

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

be the coordinate vector of $\mathbf{v}$ relative to the old basis $\mathcal{B}$. That is,

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n.$$

Using Theorem 2.12, we compute the coordinate vector of $\mathbf{v}$ relative to the new basis $\mathcal{C}$:

$$[\mathbf{v}]_{\mathcal{C}} = \left[v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n\right]_{\mathcal{C}} = v_1\left[\mathbf{b}_1\right]_{\mathcal{C}} + v_2\left[\mathbf{b}_2\right]_{\mathcal{C}} + \cdots + v_n\left[\mathbf{b}_n\right]_{\mathcal{C}}.$$

If we define the matrix

$$P_{\mathcal{C}}^{\mathcal{B}} = \left[\left[\mathbf{b}_1\right]_{\mathcal{C}} \quad \left[\mathbf{b}_2\right]_{\mathcal{C}} \quad \cdots \quad \left[\mathbf{b}_n\right]_{\mathcal{C}}\right],$$

then we can write

$$[\mathbf{v}]_{\mathcal{C}} = P_{\mathcal{C}}^{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}.$$

Following the same argument as above but interchanging the roles of the bases $\mathcal{B}$ and $\mathcal{C}$, we obtain

$$[\mathbf{v}]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}},$$

where

$$P_{\mathcal{B}}^{\mathcal{C}} = \left[\left[\mathbf{c}_1\right]_{\mathcal{B}} \quad \left[\mathbf{c}_2\right]_{\mathcal{B}} \quad \cdots \quad \left[\mathbf{c}_n\right]_{\mathcal{B}}\right].$$

Then we make the following observation: $P_{\mathcal{C}}^{\mathcal{B}}$ is an invertible matrix and $\left(P_{\mathcal{C}}^{\mathcal{B}}\right)^{-1} = P_{\mathcal{B}}^{\mathcal{C}}$.

Recapitulating, we have an invertible matrix $P_{\mathcal{C}}^{\mathcal{B}}$, called the ***matrix of change of basis*** from $\mathcal{B}$ to $\mathcal{C}$, whose columns are the coordinate vectors of each element of the old basis $\mathcal{B}$ relative to the new basis $\mathcal{C}$, which transforms $[\mathbf{v}]_{\mathcal{B}}$ into $[\mathbf{v}]_{\mathcal{C}}$. Moreover, the columns of $\left(P_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}$ are the coordinate vectors of each element of the new basis $\mathcal{C}$ relative to the old basis $\mathcal{B}$.

An important consequence of the computations described above is that, given any basis $\mathcal{B}$ of V, it is possible to obtain a new basis $\mathcal{C} = \left\{\mathbf{c}_1, \ldots, \mathbf{c}_n\right\}$ by choosing any invertible $n \times n$ matrix $P = [\mathbf{a}_1\ \mathbf{a}_2 \ldots \mathbf{a}_n]$ and letting its columns be the coordinate vectors of the vectors in the new basis relative to the old one. That is,

$$\left[\mathbf{c}_1\right]_{\mathcal{B}} = \mathbf{a}_1, \left[\mathbf{c}_2\right]_{\mathcal{B}} = \mathbf{a}_2, \cdots, \left[\mathbf{c}_n\right]_{\mathcal{B}} = \mathbf{a}_n.$$

In this case, $P = P_{\mathcal{B}}^{\mathcal{C}}$.

Example 2.28

Let $\mathcal{E}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the canonical basis for $\mathbb{R}^3$. That is,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Recall that for any vector $\mathbf{x} \in \mathbb{R}^3$, it is obvious that $[\mathbf{x}]_{\mathcal{E}_3} = \mathbf{x}$. Furthermore, consider another basis for $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}.$$

The change of basis matrix from $\mathcal{B}$ to $\mathcal{E}_3$ is

$$P_{\mathcal{E}_3}^{\mathcal{B}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{E}_3} & [\mathbf{b}_2]_{\mathcal{E}_3} & [\mathbf{b}_3]_{\mathcal{E}_3} \end{bmatrix} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}.$$

We can also compute $P_{\mathcal{B}}^{\mathcal{E}_3} = \left(P_{\mathcal{E}_3}^{\mathcal{B}}\right)^{-1}$:

$$P_{\mathcal{B}}^{\mathcal{E}_3} = \begin{bmatrix} [\mathbf{e}_1]_{\mathcal{B}} & [\mathbf{e}_2]_{\mathcal{B}} & [\mathbf{e}_3]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \frac{5}{2} \\ 0 & 0 & -\frac{1}{2} \\ 2 & -1 & 3 \end{bmatrix}.$$

It is interesting to verify that the columns of $P_{\mathcal{B}}^{\mathcal{E}_3}$ are, indeed, the coordinate vectors of the elements of the canonical basis relative to $\mathcal{B}$:

$$\mathbf{e}_1 = \mathbf{b}_1 + 2\mathbf{b}_3,\ \mathbf{e}_2 = -\mathbf{b}_3,\ \mathbf{e}_3 = \frac{5}{2}\mathbf{b}_1 - \frac{1}{2}\mathbf{b}_2 + 3\mathbf{b}_3.$$

Example 2.29

Let

$$\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix},$$

and consider the bases for $\mathbb{R}^2$ given by $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Find the change of basis matrices $P_{\mathcal{C}}^{\mathcal{C}}$ and $P_{\mathcal{B}}^{\mathcal{C}}$.

First, the matrix $P_{\mathcal{C}}^{\mathcal{B}}$ involves the coordinate vectors of $\mathbf{b}_1$ and $\mathbf{b}_2$ relative to $\mathcal{C}$. Let $\left[\mathbf{b}_1\right]_{\mathcal{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\left[\mathbf{b}_2\right]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then, by defintion

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \text{ and } \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2.$$

To solve both systems simultaneously, augment the coefficient matrix with $\mathbf{b}_1$ *and* $\mathbf{b}_2$, and row reduce:

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{bmatrix}.$$

Thus,

$$\left[\mathbf{b}_1\right]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \text{ and } \left[\mathbf{b}_2\right]_{\mathcal{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}.$$

The matrix $P_{\mathcal{C}}^{\mathcal{B}}$ is therefore

$$P_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} \left[\mathbf{b}_1\right]_{\mathcal{C}} & \left[\mathbf{b}_2\right]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}.$$

We can also compute $P_{\mathcal{B}}^{\mathcal{C}} = \left(P_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}$:

$$P_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} \left[\mathbf{c}_1\right]_{\mathcal{B}} & \left[\mathbf{c}_2\right]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & -2 \\ \frac{5}{2} & 3 \end{bmatrix}.$$

Example 2.30

Let

$$p_0(t) = (1-t)^3,\ p_1(t) = 3t(1-t)^2,\ p_2(t) = 3t^2(1-t),\ p_3(t) = t^3.$$

and consider the bases for $\mathbf{P}_3$ given by $\mathcal{B} = \{p_0(t), p_1(t), p_2(t), p_3(t)\}$ and $\mathcal{S}_3 = \{1, t, t^2, t^3\}$. Use a change of basis matrix to write the polynomial $q(t) = t^2 + 1$ as a linear combination of the basis $\mathcal{B}$.

Observe that the polynomial $q(t)$ is written as a linear combination of the basis $\mathcal{S}_3$. So we need to compute the change of basis matrix $P_{\mathcal{B}}^{\mathcal{S}_3}$. We will use this matrix to compute $\left[q(t)\right]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{S}_3}\left[q(t)\right]_{\mathcal{S}_3}$ since we know that $\left[q(t)\right]_{\mathcal{S}_3} = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$.

We will compute $P_{\mathcal{B}}^{\mathcal{S}_3}$ in two steps. We first compute $P_{\mathcal{S}_3}^{\mathcal{B}}$:

Now, we compute $P_{\mathcal{B}}^{\mathcal{S}_3} = \left(P_{\mathcal{S}_3}^{\mathcal{B}}\right)^{-1}$:

$$P_{\mathcal{B}}^{\mathcal{S}_3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1/3 & 0 & 0 \\ 1 & 2/3 & 1/3 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Since $\left[q(t)\right]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{S}_3}\left[q(t)\right]_{\mathcal{S}_3} = \begin{bmatrix} 1 & 1 & \frac{4}{3} & 2 \end{bmatrix}$, we have $q(t) = p_0(t) + p_1(t) + \frac{4}{3}p_2(t) + 2p_3(t).$