EXAMPLE 1. In V_4, find an orthonormal basis for the subspace spanned by the three vectors $x_1 = (1, -1, 1, -1)$, $x_2 = (5, 1, 1, 1)$, and $x_3 = (-3, -3, 1, -3)$.

Solution. Applying the Gram-Schmidt process, we find

$$y_1 = x_1 = (1, -1, 1, -1),$$

$$y_2 = x_2 - \frac{(x_2, y_1)}{(y_1, y_1)} y_1 = x_2 - y_1 = (4, 2, 0, 2),$$

$$y_3 = x_3 - \frac{(x_3, y_1)}{(y_1, y_1)} y_1 - \frac{(x_3, y_2)}{(y_2, y_2)} y_2 = x_3 - y_1 + y_2 = (0, 0, 0, 0).$$

Since $y_3 = O$, the three vectors x_1, x_2, x_3 must be dependent. But since y_1 and y_2 are nonzero, the vectors x_1 and x_2 are independent. Therefore $L(x_1, x_2, x_3)$ is a subspace of dimension 2. The set $\{y_1, y_2\}$ is an orthogonal basis for this subspace. Dividing each of y_1 and y_2 by its norm we get an orthonormal basis consisting of the two vectors

$$\frac{y_1}{\|y_1\|} = \tfrac{1}{2}(1, -1, 1, -1) \quad \text{and} \quad \frac{y_2}{\|y_2\|} = \frac{1}{\sqrt{6}}(2, 1, 0, 1).$$

EXAMPLE 2. *The Legendre polynomials.* In the linear space of all polynomials, with the inner product $(x, y) = \int_{-1}^{1} x(t)y(t)\,dt$, consider the infinite sequence $x_0, x_1, x_2, \ldots$, where $x_n(t) = t^n$. When the orthogonalization theorem is applied to this sequence it yields another sequence of polynomials $y_0, y_1, y_2, \ldots$, first encountered by the French mathematician A. M. Legendre (1752-1833) in his work on potential theory. The first few polynomials are easily calculated by the Gram-Schmidt process. First of all, we have $y_0(t) = x_0(t) = 1$. Since

$$(y_0, y_0) = \int_{-1}^{1} dt = 2 \quad \text{and} \quad (x_1, y_0) = \int_{-1}^{1} t\,dt = 0,$$

we find that

$$y_1(t) = x_1(t) - \frac{(x_1, y_0)}{(y_0, y_0)} y_0(t) = x_1(t) = t.$$

Next, we use the relations

$$(x_2, y_0) = \int_{-1}^{1} t^2\,dt = \tfrac{2}{3}, \quad (x_2, y_1) = \int_{-1}^{1} t^3\,dt = 0, \quad (y_1, y_1) = \int_{-1}^{1} t^2\,dt = \tfrac{2}{3},$$

to obtain

$$y_2(t) = x_2(t) - \frac{(x_2, y_0)}{(y_0, y_0)} y_0(t) - \frac{(x_2, y_1)}{(y_1, y_1)} y_1(t) = t^2 - \tfrac{1}{3}.$$

Similarly, we find that

$$y_3(t) = t^3 - \tfrac{3}{5}t, \qquad y_4(t) = t^4 - \tfrac{6}{7}t^2 + \tfrac{3}{35}, \qquad y_5(t) = t^5 - \tfrac{10}{9}t^3 + \tfrac{5}{21}t.$$

We shall encounter these polynomials again in Chapter 6 in our further study of differential equations, and we shall prove that

$$y_n(t) = \frac{n!}{(2n)!} \frac{d^n}{dt^n} (t^2 - 1)^n.$$

The polynomials P_n given by

$$P_n(t) = \frac{(2n)!}{2^n(n!)^2} y_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} (t^2 - 1)^n$$

are known as the ***Legendre polynomials.*** The polynomials in the corresponding orthonormal sequence $\varphi_0, \varphi_1, \varphi_2, \ldots,$ given by $\varphi_n = y_n/\|y_n\|$ are called the ***normalized Legendre polynomials.*** From the formulas for $y_0, \ldots, y_5$ given above, we find that

$$\varphi_0(t) = \sqrt{\tfrac{1}{2}}\,, \qquad \varphi_1(t) = \sqrt{\tfrac{3}{2}}\, t\,, \qquad \varphi_2(t) = \tfrac{1}{2}\sqrt{\tfrac{5}{2}}\,(3t^2 - 1), \qquad \varphi_3(t) = \tfrac{1}{2}\sqrt{\tfrac{7}{2}}\,(5t^3 - 3t),$$

$$\varphi_4(t) = \tfrac{1}{8}\sqrt{\tfrac{9}{2}}\,(35t^4 - 30t^2 + 3), \qquad \varphi_5(t) = \tfrac{1}{8}\sqrt{\tfrac{11}{2}}\,(63t^5 - 70t^3 + 15t).$$

1.15. Orthogonal complements. Projections

Let V be a Euclidean space and let S be a finite-dimensional subspace. We wish to consider the following type of approximation problem: ***Given an element x in V, to determine an element in S whose distance from x is as small as possible.*** The distance between two elements x and y is defined to be the norm $\|x - y\|$.

Before discussing this problem in its general form, we consider a special case, illustrated in Figure 1.2. Here V is the vector space V_3 and S is a two-dimensional subspace, a plane through the origin. Given x in V, the problem is to find, in the plane S, that point s nearest to x.

If $x \in S$, then clearly $s = x$ is the solution. If x is not in S, then the nearest point s is obtained by dropping a perpendicular from x to the plane. This simple example suggests an approach to the general approximation problem and motivates the discussion that follows.

DEFINITION. *Let S be a subset of a Euclidean space V. An element in V is said to be orthogonal to S if it is orthogonal to every element of S. The set of all elements orthogonal to S is denoted by $S^\perp$ and is called "S perpendicular."*

It is a simple exercise to verify that $S^\perp$ is a subspace of V, whether or not S itself is one. In case S is a subspace, then $S^\perp$ is called the ***orthogonal complement*** of S.

EXAMPLE. If S is a plane through the origin, as shown in Figure 1.2, then $S^\perp$ is a line through the origin perpendicular to this plane. This example also gives a geometric interpretation for the next theorem.

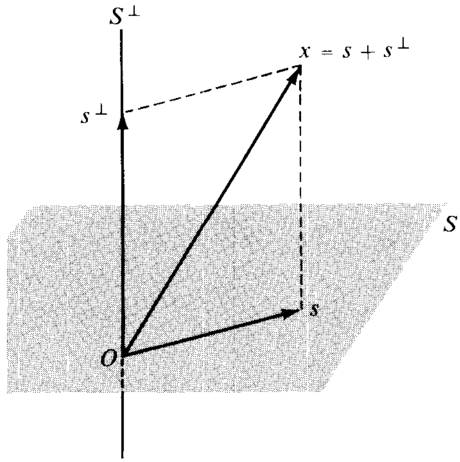

FIGURE 1.2 Geometric interpretation of the orthogonal decomposition theorem in V_3.

THEOREM 1.15. ORTHOGONAL DECOMPOSITION THEOREM. *Let V be a Euclidean space and let S be a finite-dimensional subspace of V. Then every element x in V can be represented uniquely as a sum of two elements, one in S and one in $S^\perp$. That is, we have*

$$x = s + s^\perp, \quad \text{where } s \in S \quad \text{and} \quad s^\perp \in S^\perp. \tag{1.17}$$

Moreover, the norm of x is given by the Pythagorean formula

$$\|x\|^2 = \|s\|^2 + \|s^\perp\|^2. \tag{1.18}$$

Proof. First we prove that an orthogonal decomposition (1.17) actually exists. Since S is finite-dimensional, it has a finite orthonormal basis, say {e, , . . . , e,}. Given x, define the elements s and $s^\perp$ as follows:

$$s = \sum_{i=1}^{n} (x,\ e_i)e_i, \qquad s^\perp = x - s. \tag{1.19}$$

Note that each term (x, $e_i)e_i$ is the projection of x along e_i . The element s is the sum of the projections of x along each basis element. Since s is a linear combination of the basis elements, s lies in S. The definition of $s^\perp$ shows that Equation (1.17) holds. To prove that $s^\perp$ lies in $S^\perp$, we consider the inner product of $s^\perp$ and any basis element e_j . We have

$$(s^\perp,\ e_j) = (x - s, e_j) = (x, e,) - (s, e_j).$$

But from (1.19), we find that (s, e_j) = (x, e,), so $s^\perp$ is orthogonal to e_j. Therefore $s^\perp$ is orthogonal to every element in S, which means that $s^\perp \in S^\perp$.

Next we prove that the orthogonal decomposition (1.17) is unique. Suppose that x has two such representations, say

$$x = s + s^\perp \qquad \text{and} \qquad x = t + t^\perp, \tag{1.20}$$

where s and t are in S, and $s^\perp$ and $t^\perp$ are in $S^\perp$. We wish to prove that $s = t$ and $s^\perp = t^\perp$. From (1.20), we have $s - t = t^\perp - s^\perp$, so we need only prove that $s - t = 0$. But $s - t \in S$ and $t^\perp - s^\perp \in S^\perp$ so $s - t$ is both orthogonal to $t^\perp - s^\perp$ and equal to $t^\perp - s^\perp$. Since the zero element is the only element orthogonal to itself, we must have $s - t = 0$. This shows that the decomposition is unique.

Finally, we prove that the norm of x is given by the Pythagorean formula. We have

$$\|x\|^2 = (x, x) = (s + s^\perp, s + s^\perp) = (s, s) + (s^\perp, s^\perp),$$

the remaining terms being zero since s and $s^\perp$ are orthogonal. This proves (1.18).

DEFINITION. *Let S be a finite-dimensional subspace of a Euclidean space V, and let $\{e_1, \ldots, e_n\}$ be an orthonormal basis for S. If $x \in V$, the element s defined by the equation*

$$s = \sum_{i=1}^{n} (x, e_i)e_i$$

is called the projection of x on the subspace S.

We prove next that the projection of x on S is the solution to the approximation problem stated at the beginning of this section.

1.16 Best approximation of elements in a Euclidean space by elements in a finite-dimensional subspace

THEOREM 1.16. APPROXIMATION THEOREM. *Let S be a finite-dimensional subspace of a Euclidean space V, and let x be any element of V. Then the projection of x on S is nearer to x than any other element of S. That is, if s is the projection of x on S, we have*

$$\|x - s\| \le \|x - t\|$$

for all t in S; the equality sign holds if and only if $t = s$.

Proof. By Theorem 1.15 we can write $x = s + s^\perp$, where $s \in S$ and $s^\perp \in S^\perp$. Then, for any t in S, we have

$$x - t = (x - s) + (s - t).$$

Since $s - t \in S$ and $x - s = s^\perp \in S^\perp$, this is an orthogonal decomposition of $x - t$, so its norm is given by the Pythagorean formula

$$\|x - t\|^2 = \|x - s\|^2 + \|s - t\|^2.$$

But $\|s - t\|^2 \ge 0$, so we have $\|x - t\|^2 \ge \|x - s\|^2$, with equality holding if and only if $s = t$. This completes the proof.

EXAMPLE 1. *Approximation of continuous functions on* $[0, 2\pi]$ *by trigonometric polynomials.* Let $V = C(0, 2\pi)$, the linear space of all real functions continuous on the interval $[0, 2\pi]$, and define an inner product by the equation $(f, g) = \int_0^{2\pi} f(x)g(x)\,dx$. In Section 1.12 we exhibited an orthonormal set of trigonometric functions $\varphi_0, \varphi_1, \varphi_2, \ldots,$ where

$$(1.21)\quad \varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \qquad \varphi_{2k-1}(x) = \frac{\cos kx}{\sqrt{\pi}}, \qquad \varphi_{2k}(x) = \frac{\sin kx}{\sqrt{\pi}}, \qquad \text{for } k \geq 1.$$

The $2n + 1$ elements $\varphi_0, \varphi_1, \ldots, \varphi_{2n}$ span a subspace S of dimension $2n + 1$. The elements of S are called *trigonometric polynomials.*

If $f \in C(0, 2\pi)$, let f_n denote the projection off on the subspace S. Then we have

$$(1.22)\qquad f_n = \sum_{k=0}^{2n} (f, \varphi_k)\varphi_k, \qquad \text{where } (f, \varphi_k) = \int_0^{2\pi} f(x)\varphi_k(x)\,dx.$$

The numbers (f, φ_k) are called *Fourier coefficients* off. Using the formulas in (1.21), we can rewrite (1.22) in the form

$$(1.23)\qquad f_n(x) = \tfrac{1}{2}a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

where

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,dx, \qquad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,dx$$

for $k = 0, 1, 2, \ldots, n$. The approximation theorem tells us that the trigonometric polynomial in (1.23) approximates f better than any other trigonometric polynomial in S, in the sense that the norm $\|f - f_n\|$ is as small as possible.

EXAMPLE 2. *Approximation of continuous functions on* $[-1, 1]$ *by polynomials of degree* $\leq n$. Let $V = C(-1, 1)$, the space of real continuous functions on $[-1, 1]$, and let $(f, g) = \int_{-1}^{1} f(x)g(x)\,dx$. The $n + 1$ normalized Legendre polynomials $\varphi_0, \varphi_1, \ldots, \varphi_n$, introduced in Section 1.14, span a subspace S of dimension $n + 1$ consisting of all polynomials of degree $\leq n$. If $f \in C(-1, 1)$, let f_n denote the projection off on S. Then we have

$$f_n = \sum_{k=0}^{n} (f, \varphi_k)\varphi_k, \qquad \text{where } (f, \varphi_k) = \int_{-1}^{1} f(t)\varphi_k(t)\,dt.$$

This is the polynomial of degree $\leq n$ for which the norm $\|f - f_n\|$ is smallest. For example, when $f(x) = \sin \pi x$, the coefficients (f, φ_k) are given by

$$(f, \varphi_k) = \int_{-1}^{1} \sin \pi t\, \varphi_k(t)\,dt.$$

In particular, we have $(f, \varphi_0) = 0$ and

$$(f, \varphi_1) = \int_{-1}^{1} \sqrt{\frac{3}{2}}\, t \sin \pi t\,dt = \sqrt{\frac{3}{2}}\,\frac{2}{\pi}.$$

Therefore the linear polynomial $f_1(t)$ which is nearest to sin πt on [- 1, 1] is

$$f_1(t) = \sqrt{\frac{3}{2}}\frac{2}{\pi}\,\varphi_1(t) = \frac{3}{\pi}\,t\,.$$

Since $(f, \varphi_2) = 0$, this is also the nearest quadratic approximation.

1.17 Exercises

1. In each case, find an orthonormal basis for the subspace of V_3 spanned by the given vectors.
(a) $x_1 = (1, 1, 1)$, $x_2 = (1, 0, 1)$, $x_3 = (3, 2, 3)$.
(b) $x_1 = (1, 1, 1)$, $x_2 = (-1, 1, -1)$, $x_3 = (1, 0, 1)$.
2. In each case, find an orthonormal basis for the subspace of V_4 spanned by the given vectors.
(a) $x_1 = (1, 1, 0, 0)$, $x_2 = (0, 1, 1, 0)$, $x_3 = (0, 0, 1, 1)$, $x_4 = (1, 0, 0, 1)$.
(b) $x_1 = (1, 1, 0, 1)$, $x_2 = (1, 0, 2, 1)$, $x_3 = (1, 2, -2, 1)$.
3. In the real linear space $C(0, \pi)$, with inner product (x, $y) = \int_0^\pi x(t)y(t)\,dt$, let x,(t) = cos nt forn $= 0, 1, 2, \ldots$. Prove that the functions $y_0, y_1, y_2, \ldots$, given by

$$y_0(t) = \frac{1}{\sqrt{\pi}} \quad \text{and} \quad y_n(t) = \sqrt{\frac{2}{\pi}}\cos nt \quad \text{for } n \geq 1,$$

form an orthonormal set spanning the same subspace as $x_0, x_1, x_2, \ldots$.
4. In the linear space of all real polynomials, with inner product $(x, y) = \int_0^1 x(t)y(t)\,dt$, let x,(t) = t^n for n = 0, 1, 2, ... Prove that the functions

$$y_0(t) = 1, \quad y_1(t) = \sqrt{3}\,(2t - 1), \quad y,(t) = \sqrt{5}\,(6t^2 - 6t + 1)$$

form an orthonormal set spanning the same subspace as $\{x_0, x_1, x_2\}$.
5. Let V be the linear space of all real functions f continuous on $[0, +\infty)$ and such that the integral $\int_0^\infty e^{-t}f^2(t)\,dt$ converges. Define $(f, g) = \int_0^\infty e^{-t}f(t)g(t)\,dt$, and let $y_0, y_1, y_2, \ldots$, be the set obtained by applying the Gram-Schmidt process to $x_0, x_1, x_2, \ldots$, where x,(t) = t^n for $n \geq 0$. Prove that $y_0(t) = 1$, $y_1(t) = t - 1$, $y_2(t) = t^2 - 4t + 2$, $y_3(t) = t^3 - 9t^2 + 18t - 6$.
6. In the real linear space $C(1, 3)$ with inner product $(f, g) = \int_1^3 f(x)g(x)\,dx$, let $f(x)$ = l/x and show that the constant polynomial g nearest to f is $g = \frac{1}{2}\log 3$. Compute $\|g - f\|^2$ for this g.
7. In the real linear space $C(0, 2)$ with inner product $(f, g) = \int_0^2 f(x)g(x)\,dx$, let $f(x) = e^x$ and show that the constant polynomial g nearest to f is $g = \frac{1}{2}(e^2 - 1)$. Compute $\|g - f\|^2$ for this g.
8. In the real linear space C($-$ 1, 1) with inner product $(f, g) = \int_{-1}^1 f(x)g(x)\,dx$, let $f(x) = e^x$ and find the linear polynomial g nearest to f. Compute $\|g - f\|^2$ for this g.
9. In the real linear space $C(0, 2\pi)$ with inner product $(f, g) = \int_0^{2\pi} f(x)g(x)\,dx$, let $f(x) = x$. In the subspace spanned by u,,(x) = 1, $u_1(x)$ = cos x, $u_2(x)$ = sin x, find the trigonometric polynomial nearest to f.
10. In the linear space V of Exercise 5, let $f(x) = e^{-x}$ and find the linear polynomial that is nearest to f.

2

LINEAR TRANSFORMATIONS AND MATRICES

2.1 Linear transformations

One of the ultimate goals of analysis is a comprehensive study of functions whose domains and ranges are subsets of linear spaces. Such functions are called ***transformations, mappings,*** or ***operators.*** This chapter treats the simplest examples, called ***linear*** transformations, which occur in all branches of mathematics. Properties of more general transformations are often obtained by approximating them by linear transformations.

First we introduce some notation and terminology concerning arbitrary functions. Let V and W be two sets. The symbol

$$T: V \to W$$

will be used to indicate that T is a function whose domain is V and whose values are in W. For each x in V, the element $T(x)$ in W is called the ***image of x under T***, and we say that T ***maps*** x ***onto*** $T(x)$. If A is any subset of V, the set of all images $T(x)$ for x in A is called the ***image of A under T*** and is denoted by $T(A)$. The image of the domain V, $T(V)$, is the range of T.

Now we assume that V and W are linear spaces having the same set of scalars, and we define a linear transformation as follows.

DEFINITION. *If V and W are linear spaces, a function $T: V \to W$ is called a linear transformation of V into W if it has the following two properties:*

(a) $T(x + y) = T(x) + T(y)$ *for all x and y in V,*

(b) $T(cx) = cT(x)$ *for all x in V and all scalars c.*

These properties are verbalized by saying that T preserves addition and multiplication by scalars. The two properties can be combined into one formula which states that

$$T(ax + by) = aT(x) + bT(y)$$

for all x, y in V and all scalars a and b. By induction, we also have the more general relation

$$T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i)$$

for any n elements $x_1, \ldots, x_n$ in V and any n scalars $a_1, \ldots, a_n$.

The reader can easily verify that the following examples are linear transformations.

EXAMPLE 1. ***The identity transformation.*** The transformation $T: V \to V$, where $T(x) = x$ for each x in V, is called the identity transformation and is denoted by I or by I_V.

EXAMPLE 2. ***The zero transformation.*** The transformation $T: V \to V$ which maps each element of V onto 0 is called the zero transformation and is denoted by 0.

EXAMPLE 3. ***Multiplication by a fixed scalar c.*** Here we have $T: V \to V$, where $T(x) = cx$ for all x in V. When $c = 1$, this is the identity transformation. When $c = 0$, it is the zero transformation.

EXAMPLE 4. ***Linear equations.*** Let $V = V_n$ and $W = V_m$. Given mn real numbers a_{ik}, where $i = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, n$, define $T: V_n \to V_m$ as follows: T maps each vector $x = (x_1, \ldots, x_n)$ in V_n onto the vector $y = (y_1, \ldots, y_m)$ in V_m according to the equations

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \qquad \text{for} \quad i = 1, 2, \ldots, m.$$

EXAMPLE 5. ***Inner product with a fixed element.*** Let V be a real Euclidean space. For a fixed element z in V, define $T: V \to R$ as follows: If $x \in V$, then $T(x) = (x, z)$, the inner product of x with z.

EXAMPLE 6. ***Projection on a subspace.*** Let V be a Euclidean space and let S be a finite-dimensional subspace of V. Define $T: V \to S$ as follows: If $x \in V$, then $T(x)$ is the projection of x on S.

EXAMPLE 7. ***The differentiation operator.*** Let V be the linear space of all real functions f differentiable on an open interval (a, b). The linear transformation which maps each function f in V onto its derivative f' is called the differentiation operator and is denoted by D. Thus, we have $D: V \to W$, where $D(f) = f'$ for each f in V. The space W consists of all derivatives f'.

EXAMPLE 8. ***The integration operator.*** Let V be the linear space of all real functions continuous on an interval $[a, b]$. If $f \in V$, define $g = T(f)$ to be that function in V given by

$$g(x) = \int_a^x f(t)\, dt \qquad \text{if} \quad a \le x \le b.$$

This transformation T is called the integration operator.

2.2 Null space and range

In this section, T denotes a linear transformation of a linear space V into a linear space W.

THEOREM 2.1. ***The set $T(V)$ (the range of T) is a subspace of W. Moreover, T maps the zero element of V onto the zero element of W.***

Proof. To prove that $T(V)$ is a subspace of W, we need only verify the closure axioms. Take any two elements of T(V), say $T(x)$ and $T(y)$. Then $T(x) + T(y) = T(x + y)$, so $T(x) + T(y)$ is in $T(V)$. Also, for any scalar c we have $cT(x) = T(cx)$, so $cT(x)$ is in T(V). Therefore, $T(V)$ is a subspace of W. Taking $c = 0$ in the relation $T(cx) = cT(x)$, we find that $T(0) = 0$.

DEFINITION. ***The set of all elements in V that T maps onto 0 is called the null space of T and is denoted by N(T). Thus, we have***

$$N(T) = \{x \mid x \in V \text{ and } T(x) = 0\}.$$

The null space is sometimes called the kernel of T.

THEOREM **2.2.** ***The null space of T is a subspace of V.***

Proof. If x and y are in N(T), then so are x + y and cx for all scalars c, since

$$T(x + y) = T(x) + T(y) = 0 \quad \text{and} \quad T(cx) = cT(x) = 0.$$

The following examples describe the null spaces of the linear transformations given in Section 2.1.

EXAMPLE **1. *Identity transformation.*** The null space is {0}, the subspace consisting of the zero element alone.

EXAMPLE **2. *Zero transformation,*** Since every element of V is mapped onto zero, the null space is V itself.

EXAMPLE **3. *Multiplication by a fixed scalar*** c. If $c \neq 0$, the null space contains only 0. If $c = 0$, the null space is V.

EXAMPLE **4. *Linear equations.*** The null space consists of all vectors $(x_1, \ldots, x_n)$ in V_n for which

$$\sum_{k=1}^{n} a_{ik}x_k = 0 \qquad \text{for} \quad i = 1, 2, \ldots, m.$$

EXAMPLE **5. *Inner product with a fixed element*** z. The null space consists of all elements in V orthogonal to z.

EXAMPLE **6. *Projection on a subspace*** S. If $x \in V$, we have the unique orthogonal decomposition $x = s + s^{\perp}$ (by Theorem I .15). Since $T(x) = s$, we have $T(x) = 0$ if and only if $x = s^{\perp}$. Therefore, the null space is $S^{\perp}$, the orthogonal complement of S.

EXAMPLE **7. *Differentiation operator.*** The null space consists of all functions that are constant on the given interval.

EXAMPLE **8. *Integration operator.*** The null space contains only the zero function.

2.3 Nullity and rank

Again in this section T denotes a linear transformation of a linear space V into a linear space W. We are interested in the relation between the dimensionality of V, of the null space $N(T)$, and of the range $T(V)$. If V is finite-dimensional, then the null space is also finite-dimensional since it is a subspace of V. The dimension of $N(T)$ is called the ***nullity*** of T. In the next theorem, we prove that the range is also finite-dimensional; its dimension is called the ***rank*** of T.

THEOREM 2.3. NULLITY PLUS RANK THEOREM. *If V is finite-dimensional, then $T(V)$ is also finite-dimensional, and we have*

$$\dim N(T) + \dim T(V) = \dim V. \tag{2.1}$$

In other words, the nullity plus the rank of a linear transformation is equal to the dimension of its domain.

Proof. Let $n = \dim V$ and let $e_1, \ldots, e_k$ be a basis for $N(T)$, where $k = \dim N(T) \leq n$. By Theorem 1.7, these elements are part of some basis for V, say the basis

$$e_1, \ldots, e_k, e_{k+1}, \ldots, e_{k+r}, \tag{2.2}$$

where $k + r = n$. We shall prove that the r elements

$$T(e_{k+1}), \ldots, T(e_{k+r}) \tag{2.3}$$

form a basis for $T(V)$, thus proving that $\dim T(V) = r$. Since $k + r = n$, this also proves (2.1).

First we show that the r elements in (2.3) span $T(V)$. If $y \in T(V)$, we have $y = T(x)$ for some x in V, and we can write $x = c_1e_1 + \cdots + c_{k+r}e_{k+r}$. Hence, we have

$$y = T(x) = \sum_{i=1}^{k+r} c_iT(e_i) = \sum_{i=1}^{k} c_iT(e_i) + \sum_{i=k+1}^{k+r} c_iT(e_i) = \sum_{i=k+1}^{k+r} c_iT(e_i)$$

since $T(e_1) = \cdots = T(e_k) = 0$. This shows that the elements in (2.3) span $T(V)$.

Now we show that these elements are independent. Suppose that there are scalars $c_{k+1}, \ldots, c_{k+r}$ such that

$$\sum_{i=k+1}^{k+r} c_iT(e_i) = 0.$$

This implies that

$$T\left(\sum_{i=k+1}^{k+r} c_ie_i\right) = O$$

so the element $x = c_{k+1}e_{k+1} + \cdots + c_{k+r}e_{k+r}$ is in the null space $N(T)$. This means there are scalars $c_1, \ldots, c_k$ such that $x = c_1e_1 + \cdots + c_ke_k$, so we have

$$x - x = \sum_{i=1}^{k} c_ie_i - \sum_{i=k+1}^{k+r} c_ie_i = 0 .$$

But since the elements in (2.2) are independent, this implies that all the scalars c_i are zero. Therefore, the elements in (2.3) are independent.

Note: If V is infinite-dimensional, then at least one of N(T) or T(V) is infinite-dimensional. A proof of of this fact is outlined in Exercise 30 of Section 2.4.

2.4 Exercises

In each of Exercises 1 through 10, a transformation $T: V_2 \to V_2$ is defined by the formula given for $T(x, y)$, where (x, y) is an arbitrary point in V_2. In each case determine whether T is linear. If T is linear, describe its null space and range, and compute its nullity and rank.

1. $T(x, y) = (y, x)$.
2. $T(x, y) = (x, -y)$.
3. $T(x, y) = (x, 0)$.
4. $T(x, y) = (x, x)$.
5. $T(x, y) = (x^2, y^2)$.
6. $T(x, y) = (e^x, e^y)$.
7. $T(x, y) = (x, 1)$.
8. $T(x, y) = (x + 1, y + 1)$.
9. $T(x, y) = (x - y, x + y)$.
10. $T(x, y) = (2x - y, x + y)$.

Do the same as above for each of Exercises 11 through 15 if the transformation $T: V_2 \to V_2$ is described as indicated.

11. T rotates every point through the same angle φ about the origin. That is, T maps a point with polar coordinates (r, θ) onto the point with polar coordinates $(r, \theta + \varphi)$, where φ is fixed. Also, T maps O onto itself.
12. T maps each point onto its reflection with respect to a fixed line through the origin.
13. T maps every point onto the point (1, 1).
14. T maps each point with polar coordinates (r, θ) onto the point with polar coordinates $(2r, \theta)$. Also, T maps O onto itself.
15. T maps each point with polar coordinates (r, θ) onto the point with polar coordinates $(r, 2\theta)$. Also, T maps O onto itself.

Do the same as above in each of Exercises 16 through 23 if a transformation $T: V_3 \to V_3$ is defined by the formula given for $T(x, y, z)$, where (x, y, z) is an arbitrary point of V_3.

16. $T(x, y, z) = (z, y, x)$.
17. $T(x, y, z) = (x, y, 0)$.
18. $T(x, y, z) = (x, 2y, 3z)$.
19. $T(x, y, z) = (x, y, 1)$.
20. $T(x, y, z) = (x + 1, y + 1, z - 1)$.
21. $T(x, y, z) = (x + 1, y + 2, z + 3)$.
22. $T(x, y, z) = (x, y^2, z^3)$.
23. $T(x, y, z) = (x + z, 0, x + y)$.

In each of Exercises 24 through 27, a transformation $T: V \to V$ is described as indicated. In each case, determine whether T is linear. If T is linear, describe its null space and range, and compute the nullity and rank when they are finite.

24. Let V be the linear space of all real polynomials $p(x)$ of degree $\leq n$. If $p \in V$, $q = T(p)$ means that $q(x) = p(x + 1)$ for all real x.
25. Let V be the linear space of all real functions differentiable on the open interval $(-1, 1)$. If $f \in V$, $g = T(f)$ means that $g(x) = xf'(x)$ for all x in $(-1, 1)$.