Chapter 3

VECTOR SPACES

3.1 EUCLIDEAN SPACE E^n

Real euclidean space is a generalization of two-space and three-space. Generalization is needed since in some applications more than three variables may be needed to describe a situation. Recall the example of an absorption column at the beginning of Chap. 2: The number of plates in the column dictates the number of components in the concentration vector. Or consider the problem of estimating the thermal conductivity (a real number associated with a given material) of a heat-conducting body. Solution may require that temperatures be measured at several points at several times. Arrangement of these data in a vector requires much more than three components.

Definition 3.1.1

REAL EUCLIDEAN n-SPACE, DENOTED E^n, CONSISTS OF	PARALLEL SITUATION FOR THREE-SPACE
1. The set of all ordered n-tuples $(x_1, x_2, \dots, x_n)$, where $x_1, x_2, \dots, x_n$ are all real numbers. An n-tuple is called a **vector**, and the numbers $x_1, \dots, x_n$ are called **components** of the vector.	The set of ordered triples (x_1, x_2, x_3).
2. An operation of **vector addition** + defined by $(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$	
3. An operation of **scalar multiplication** defined by $r(x_1, \dots, x_n) = (rx_1, \dots, rx_n)$	

Addition here was given by

$$(x_1, x_2, x_3) + (y_1, y_2, y_3)$$
$$= (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

Multiplication here was given by

$$r(x_1, x_2, x_3) = (rx_1, rx_2, rx_3)$$

Definition 3.1.2. Two vectors $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ in E^n are called **equal** if $x_1 = y_1$, $x_2 = y_2, \dots, x_n = y_n$.

Example 1. In E^5 let $\mathbf{v} = (-1, 0, 2, 3, -6), \mathbf{w} = (1, 0, -2, -3, 6)$, $\mathbf{u} = (3, 7, -6, 2, -1)$, and $\mathbf{z} = (0, 0, 0, 0, 0)$. Calculate (**a**) $\mathbf{u} + \mathbf{v}$, (**b**) $\mathbf{v} + \mathbf{u}$, (**c**) $\mathbf{u} + \mathbf{z}$,(**d**) $3\mathbf{u}$. (**e**) $-7\mathbf{v}$, (**f**) $3\mathbf{u} + 7\mathbf{v}$, (**g**) $\mathbf{v} + \mathbf{w}$, (**h**) $0\mathbf{u}$, and (**i**) $2\mathbf{z}$.

Solution

(**a**)

$$\begin{aligned}&(3, 7, -6, 2, -1) + (-1, 0, 2, 3, -6)\\ &\quad = (3 + (-1), 7 + 0, -6 + 2, 2 + 3, -1 + (-6))\\ &\quad = (2, 7, -4, 5, -7)\end{aligned}$$

(**b**)

$$\begin{aligned}&(-1, 0, 2, 3, -6) + (3, 7 - 6, 2, -1)\\ &\quad = (-1 + 3, 0 + 7, 2 + (-6), 3 + 2, -6 + (-1))\\ &\quad = (2, 7, -4, 5, -7)\end{aligned}$$

Note that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

(**c**)

$$\begin{aligned}\mathbf{u} + \mathbf{z} &= (3, 7, -6, 2, -1) + (0, 0, 0, 0, 0)\\ &= (3 + 0, 7 + 0, -6 + 0, 2 + 0, -1 + 0)\\ &= (3, 7, -6, 2, -1)\end{aligned}$$

(**d**)

$$\begin{aligned}3\mathbf{u} = 3(3, 7, -6, 2, -1) &= (3 \cdot 3, 3 \cdot 7, 3 \cdot (-6), 3 \cdot 2, 3 \cdot (-1))\\ &= (9, 21, -18, 6, -3)\end{aligned}$$

(**e**)

$$\begin{aligned}-7\mathbf{v} = -7(-1,0,2,3,-6) &= (-7\cdot(-1), -7\cdot 0, -7\cdot 2, -7\cdot 3, -7\cdot(-6))\\ &= (7,0,-14,21,42)\end{aligned}$$

(**f**)

$$\begin{aligned}3\mathbf{u} + 7\mathbf{v} &= (9,21,-18,6,-3) + (-7,0,14,21,-42)\\ &= (2,21,-4,27,-45)\end{aligned}$$

(**g**)

$$\begin{aligned}\mathbf{v} + \mathbf{w} &= (-1,0,2,3,-6) + (1,0,-2,-3,6)\\ &= (0,0,0,0,0)\end{aligned}$$

(**h**)

$$\begin{aligned}0\mathbf{u} &= 0(3,7,-6,2,-1) = (0\cdot 3, 0\cdot 7, 0\cdot(-6), 0\cdot 2, 0\cdot(-1))\\ &= (0,0,0,0,0)\end{aligned}$$

Notice that $0(x_1, x_2, x_3, x_4, x_5) = (0,0,0,0,0)$ regardless of the values for x_1, x_2, x_3, x_4, and x_5.

(**i**)

$$\begin{aligned}2\mathbf{z} = 2(0,0,0,0,0) &= (2\cdot 0, 2\cdot 0, 2\cdot 0, 2\cdot 0, 2\cdot 0)\\ &= (0,0,0,0,0) = \mathbf{z}\end{aligned}$$

Note that $r\mathbf{z} = \mathbf{z}$ for any real number r.

Example 1**b** leads us to believe that the commutative property for addition of vectors in three-space carries over to E^n. Also the vector with all zero components may have a special role (Example 1**c**) as an additive identity. In fact, this intuition is correct. Let us define some special vectors first.

Definition 3.1.3. The **zero vector** in E^n is denoted by θ and is defined by $\theta = (0,0,\ldots,0)$.

Definition 3.1.4. Given a vector $\mathbf{x} = (x_1, \dots, x_n)$ in E^n, the **negative** of $\mathbf{x}$ in E^n is denoted $-\mathbf{x}$ and is defined by

$$-\mathbf{x} = (-x_1, -x_2, \dots, -x_n)$$

Example 2. Let $\mathbf{x} = (1, 2, -3, 4)$. Find $-\mathbf{x}$. Calculate $\mathbf{x} + (-\mathbf{x})$.

Solution By definition $-\mathbf{x} = (-1, -2, -(-3), -4) = (-1, -2, 3, -4)$. Now $\mathbf{x} + (-\mathbf{x}) = (1, 2, -3, 4) + (-1, -2, 3, -4) = (0, 0, 0, 0) = \theta$.

Example 2 illustrates the fact that for any $\mathbf{x}$ in E^n, $\mathbf{x} + (-\mathbf{x}) = \theta$. This is part (**d**) of the next theorem.

Theorem 3.1.1. *Let* $\mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{y} = (y_1, y_2, \dots, y_n)$, *and* $\mathbf{z} = (z_1, z_2, \dots, z_n)$ *be vectors in* E^n, *and let* r *and* s *be real numbers. Then*

(**a**) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$	Commutative law
(**b**) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$	Associative law
(**c**) $\mathbf{x} + \boldsymbol{\theta} = \mathbf{x}$	Additive identity
(**d**) $\mathbf{x} + (-\mathbf{x}) = \boldsymbol{\theta}$	Additive inverse
(**e**) $(rs)\mathbf{x} = r(s\mathbf{x})$	Associative law
(**f**) $(r + s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$	Distributive laws
(**g**) $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$	Distributive laws
(**h**) $1\mathbf{x} = \mathbf{x}$	Multiplicative identity

Proof. (**a**) $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$. The additions in each component are additions in the set of real numbers. Since addition is commutative for real numbers, we have

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n) = (y_1 + x_1, \dots, y_n + x_n) = \mathbf{y} + \mathbf{x}$$

(**d**) Since $-\mathbf{x} = (-x_1, -x_2, \dots, -x_n)$, we know that

$$\begin{aligned}\mathbf{x} + (-\mathbf{x}) &= (x_1, x_2, \dots, x_n) + (-x_1, -x_2, \dots, -x_n) \\ &= (x_1 + (-x_1), x_2 + (-x_2), \dots, x_n + (-x_n)) \\ &= (0, 0, \dots, 0) = \theta\end{aligned}$$

(**g**)

$$\begin{aligned}r(\mathbf{x} + \mathbf{y}) &= r(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ &= (r(x_1 + y_1), r(x_2 + y_2), \dots, r(x_n + y_n))\end{aligned}$$

Now since in each component we can apply the distributive law for real numbers, the last vector is

$$r(\mathbf{x}+\mathbf{y}) = (rx_1 + ry_1, rx_2 + ry_2, \dots, rx_n + ry_n)$$

However,

$$\begin{aligned} r\mathbf{x} + r\mathbf{y} &= r(x_1, x_2, \dots, x_n) + r(y_1, y_2, \dots, y_n) \\ &= (rx_1, rx_2, \dots, rx_n) + (ry_1, ry_2, \dots, ry_n) \\ &= (rx_1 + ry_1, rx_2 + ry_2, \dots, rx_n + ry_n) \end{aligned}$$

Therefore $r(\mathbf{x}+\mathbf{y}) = r\mathbf{x} + r\mathbf{y}$. □

Proofs of the other parts are left to the problems.

Example 3. Illustrate parts (**b**), (**e**) and (**f**) of Theorem 3.1.1 by means of examples from E^3.

Solution For (**b**), let $\mathbf{x} = (1, -1, 2), \mathbf{y} = (0, 3, 6)$, and $\mathbf{z} = (-5, 0, 2)$. Then

$$\begin{aligned} (\mathbf{x}+\mathbf{y}) + \mathbf{z} &= ((1, -1, 2) + (0, 3, 6) + (-5, 0, 2) \\ &= (1, 2, 8) + (-5, 0, 2) \\ &= (-4, 2, 10) \end{aligned}$$

and

$$\begin{aligned} \mathbf{x} + (\mathbf{y}+\mathbf{x}) &= (1, -1, 2) + ((0, 3, 6) + (-5, 0, 2)) \\ &= (1, -1, 2) + (-5, 3, 8) \\ &= (-4, 2, 10) \end{aligned}$$

For (**e**) let $r = 3, s = -5$, and $\mathbf{x} = (4, -2, 6)$. Then

$$\begin{aligned} (rs)\mathbf{x} &= [3(-5)](4, -2, 6) = -15(4, -2, 6) = (-60, 30, -90) \\ r(s\mathbf{x}) &= 3[-5(4, -2, 6)] = 3(-20, 10, -30) = (-60, 30, -90) \end{aligned}$$

For (**f**) use the same values as in (**e**). Then

$$\begin{aligned} (r+s)\mathbf{x} &= -2(4, -2, 6) = (-8, 4, -12) \\ r\mathbf{x} + s\mathbf{x} &= 3(4, -2, 6) + (-5)(4, -2, 6) = (12, -6, 18) + (-20, 10, -30) \\ &= (-8, 4, -12) \end{aligned}$$

Theorem 3.1.1 shows that E^n has the same generic properties as the set of two- and three-vectors had. In the next section we will see how the important properties of E^n outlined in Theorem 3.1.1 can be used to define the concept of a vector space. We postpone the discussion of generalizing the dot product until Sec. 3.6.

PROBLEMS 3.1

1. From the following list of vectors pick all possible pairs of vectors which are equal.

   $\mathbf{A} = (1, -1, 0)$ $\qquad$ $\mathbf{D} = ((a-b)^2, 1, 2)$
   $\mathbf{B} = (a^2 - b^2, 1, 2)$ $\qquad$ $\mathbf{E} = ((a-b)(a+b), (-1)^2, 2)$
   $\mathbf{C} = (\sin \pi/2, \cos \pi, \sin \pi)$

2. Let $\mathbf{x}$ and $\mathbf{y}$ be vectors in E^n. Define $\mathbf{x} - \mathbf{y}$ as $\mathbf{x} + (-\mathbf{y})$. Setting $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, write out the vector $\mathbf{x} - \mathbf{y}$.

3. Show by example that $(\mathbf{x} - \mathbf{y}) - \mathbf{z} \neq \mathbf{x} - (\mathbf{y} - \mathbf{z})$ in E^n. That is, the associative law does not hold for vector subtraction. (**Note**: To show that a statement does not hold in E^n, a particular value of n can be chosen. So in this problem the example can come from E^1, E^2, or E^3.)

4. Now E^1 is just the set of real numbers. In E^1 multiplication by a scalar is quite simple. What is it?

5. In E^3 consider the vectors $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, and $\mathbf{e}_3 = (0, 0, 1)$. A sum of the form $a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3$, where a, b, and c are real numbers, is called a **linear combination** of $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$. Write the vector $(3, -2, 1)$ as a linear combination of $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$. (These vectors are also called $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$.)

6. In E^n consider the vectors

$$\begin{aligned} &\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \\ &\mathbf{e}_k = (0, 0, \ldots, 0, \underset{\substack{\uparrow \\ k\text{th component}}}{1}, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 1) \end{aligned}$$

A sum of the form $a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n$ is called a **linear combination** of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Show that any vector $\mathbf{x} = (x_1, x_2, x_3, \ldots, x_n)$ can be written as a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$.

7. Let $\mathbf{x} = (1, -2, 3)$ and $\mathbf{y} = (3, 1, 0)$. Calculate $\mathbf{x} + \mathbf{y}$. Place the three vectors $\mathbf{x}, \mathbf{y}$, and $\mathbf{x} + \mathbf{y}$ in a matrix as rows of the matrix (remove the commas). Calculate the determinant of the matrix.

8. Work Prob. 7 with $\mathbf{x}$ and $\mathbf{y}$ as given, but replace $\mathbf{x} + \mathbf{y}$ with $3\mathbf{x} - 2\mathbf{y}$.

9. Work Prob. 7 with $\mathbf{x}$ and $\mathbf{y}$ as given, but replace $\mathbf{x} + \mathbf{y}$ with $a\mathbf{x} + b\mathbf{y}$, where a and b are nonzero real numbers.

10. Find constants c_1 and c_2 which satisfy

$$(3, 5) = c_1(1, -2) + c_2(2, -3)$$

11. Show that there do not exist constants c_1, c_2, c_3 which satisfy

$$(1, -1, 4) = c_1(2, 3, 5) + c_2(-1, 0, 6) + c_3(1, 3, 11)$$

12. Prove part (**b**) of Theorem 3.1.1. Also illustrate with an example.

13. Prove part (**c**) of Theorem 3.1.1, and illustrate with an example.

14. Prove part (**e**) of Theorem 3.1.1. Illustrate with an example.

15. Prove part (**f**) of Theorem 3.1.1. Also illustrate with an example.

16. Prove part (**h**) of Theorem 3.1.1, and illustrate with an example.

17. State the commutative, associative, and distributive laws in your own words.

3.2 VECTOR SPACES

The structure of euclidean space (as outlined in Theorem 3.1.1) is important for other "spaces" used in applied mathematics. For example, consider the mass-spring arrangement in Fig. 3.2.1**a**. If the body of mass m is displaced from rest position and released or pushed, oscillatory motion results. If this is done in the absence of external forces and friction is ignored, the displacement $y(t)$ of the body from rest is given by

$$y(t) = C_1 \cos \omega t + C_2 \sin \omega t$$

where $\omega = \sqrt{k/m}$ with k being the spring constant. Constants C_1 and C_2 are determined by how far the body is started from rest and how fast and in what direction it is pushed initially. The set of functions

$$V = \{f | f(t) = C_1 \cos \omega t + C_2 \sin \omega t, C_1 \in \mathbb{R}, C_2 \in \mathbb{R}\}$$

along with the usual definition (from calculus) of addition of functions and multiplication of a function by a constant turns out to have the properties of E^n outlined in Sec. 3.1.

Because there are many important structures with the properties of euclidean spaces, we put them all under the umbrella concept of vector space.

Definition 3.2.1. A **real vector space** consists of the following.

(**a**) A set V of objects. These objects are called **vectors** even though they may be functions or matrices in a specific case.

(**b**) An operation denoted by + which associates with each pair of vectors **v**,**w** in V a vector $\mathbf{v} + \mathbf{w}$ in V, called the **sum** of **v** and **w**.

(**c**) An operation called **scalar multiplication** which associates with each real number r and vector **v** in V a vector $r\mathbf{v}$ in V that is called the **product** of r and **v**.

The operations must be defined in such a way that

$$\left.\begin{array}{l} \text{1. Addition is commutative: } \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}. \\ \text{2. Addition is associative: } \mathbf{v} + (\mathbf{w} + \mathbf{u}) = (\mathbf{y} + \mathbf{w}) + \mathbf{u}. \end{array}\right\} \mathbf{v}, \mathbf{w}, \mathbf{u} \text{ in } V$$

3. There exists a zero vector θ in V such that $\mathbf{u} + \theta = \mathbf{u}$ for all vectors **u** in V. Vector θ is called an **additive identity**.

4. For each vector **v** in V there exists an additive inverse $-\mathbf{v}$ in V such that $\mathbf{v} + (-\mathbf{v}) = \theta$.

$$\left.\begin{array}{l} \text{5. } (rs)\mathbf{v} = r(s\mathbf{v}) \\ \text{6. } (r+s)\mathbf{v} = r\mathbf{v} + s\mathbf{v} \\ \text{7. } r(\mathbf{v} + \mathbf{w}) = r\mathbf{v} + r\mathbf{w} \end{array}\right\} r, s \in \mathbb{R}$$

8. $1\mathbf{v} = \mathbf{v}$ for every $\mathbf{v}$ in V.

It is important that a real vector space consist of the set of vectors and the two operations with certain properties. The same set of vectors with different operations may not satisfy the required properties.

Note that the properties which we **derived** for E^n have become the **defining properties**, or **axioms**, for a vector space. The reason that this works well is that E^n, although a specific example, possesses the important qualities for generalization. This happens often in applied mathematics: A specific problem leads to a specific solution—yet the solution actually solves many more problems when it is seen in a larger context.

To show that an object is a vector space, we must show that **closure for both operations** holds [parts (**b**) and (**c**) of the definition] and that properties 1 through 8 hold. Altogether these 10 properties are called the **axioms for a real vector space**. To show that an object is not a vector space, we need only show that 1 of the 10 axioms fails to hold.

Example 1. Is the set of ordered pairs $V = \{(x_1, x_2) | x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$ a vector space?

Solution Until operations of vector addition and scalar multiplication are specified, we cannot test for vector space structure. Therefore, we do not have a vector space.

Example 2. Now E^n is a vector space, because the definitions of the operations imply closure and Theorem 3.1.1 shows that properties 1 to 8 hold.

At this point, we recall that in E^n only the vector $(0, 0, \ldots, 0)$ had the property of being an additive identity. Also, given $(x_1, \ldots, x_n)$ in E^n, only $(-x_1, \ldots, -x_n)$ is an additive inverse. In general, this uniqueness holds in any vector space.

Theorem 3.2.1. *Let V be a vector space.*

*(**a**) There exists only one additive identity in V.*

*(**b**) Given $\mathbf{x}$ in V, there is only additive inverse of $\mathbf{x}$ in V.*

Proof. (**a**) Let θ_1 and θ_2 be additive identities for V. We will show that they are equal. By the additive-identity axiom

$$\theta_1 + \theta_2 = \theta_1$$

However,

$$\theta_1 + \theta_2 \underset{\text{Commutativity}}{=} \theta_2 + \theta_1 \underset{\text{Additive-identity axiom}}{=} \theta_2$$

Therefore $\theta_1 = \theta_1 + \theta_2 = \theta_2$.

(**b**) Let $\mathbf{x}$ be in V. Let $\mathbf{w}$ and $\mathbf{u}$ be additive inverses. We will show that $\mathbf{w} = \mathbf{u}$. We have

$$\mathbf{x} + \mathbf{w} = \theta = \mathbf{x} + \mathbf{u}$$

By commutativity, $\mathbf{w} + \mathbf{x} = \mathbf{u} + \mathbf{x}$, and upon adding $\mathbf{w}$ to both sides of this last equation, we have

$$(\mathbf{w} + \mathbf{x}) + \mathbf{w} = (\mathbf{u} + \mathbf{x}) + \mathbf{w}$$

Finally, by associativity

$$\begin{aligned} \mathbf{w} + (\mathbf{x} + \mathbf{w}) &= \mathbf{u} + (\mathbf{x} + \mathbf{w}) \\ \mathbf{w} + \theta &= \mathbf{u} + \theta \\ \mathbf{w} &= \mathbf{u} \end{aligned}$$

□

Example 3. Let $V = \mathcal{M}_{mn} = \{m \times n$ matrices with real entries$\}$, let vector addition be the addition of matrices, and let scalar multiplication be the multiplication of matrices by scalars. Is V with these operations a vector space?

Solution From our work with matrices we know that closure, commutativity, associativity, and distributivity hold. The additive identity is the $m \times n$ matrix with all entries zero. The additive inverse of $A_{m\times n}$ is $-A_{m\times n}$. Clearly $1A_{m\times n} = A_{m\times n}$. Therefore this is a vector space.

Example 4. Let $\mathcal{P}_2 = \{$polynomials $f(x) = a_2x^2 + a_1x + a_0$, where $a_0, a_1, a_2 \in \mathbb{R}\}$, and define addition and scalar multiplication as follows:

(**a**) Addition. Let $f(x) = a_2x^2 + a_1x + a_0$ and $g(x) = b_2x^2 + b_1x + b_0$. Then define

$$(f + g)(x) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)$$

(**b**) Scalar multiplication. Let $r \in \mathbb{R}, f(x) = a_2x^2 + a_1x + a_0$. Define

$$(rf)(x) = (ra_2)x^2 + (ra_1)x + (ra_0)$$

Show that $\mathcal{P}_2$ is a vector space.

Solution Closure follows easily from the definitions since the right-hand sides of the equations in (**a**) and (**b**) are polynomials of degree ≤ 2. For commutativity

$$\begin{aligned}(f+g)(x) &= (a_2+b_2)x^2 + (a_1+b_1)x + (a_0+b_0) \\ &= (b_2+a_2)x^2 + (b_1+a_1)x + (b_0+a_0) \\ &= (g+f)(x)\end{aligned}$$

Associativity follows similarly. The θ is $0x^2 + 0x + 0$, which is the function $f(x) \equiv 0$. The additive inverse $-f$ is simply $-a_2x^2 - a_1x - a_0$. The other axioms are easily verified.

Example 5. Let $\mathcal{P}_n = \{$polynomials $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0\}$ with operations defined by [let $g(x) = b_nx^n + \cdots + b_1x + b_0$ and $r \in \mathbb{R}$]

$$\begin{aligned}(f+g)(x) &= (a_n+b_n)x^n + \cdots + (a_1+b_1)x + (a_0+b_0) \\ (rf)(x) &= (ra_n)x^n + \cdots + (ra_1)x + (ra_0)\end{aligned}$$

So $\mathcal{P}_n$ is a real vector space.

As we see more examples of vector spaces, we will be led to theorems about their structure. Theorems are formed by considering examples. For instance, in calculus after showing

$$\frac{d}{dx}x^2 = 2x \qquad \frac{d}{dx}x^3 = 3x^2$$

by using the definition of derivative, we can guess that

$$\frac{d}{dx}x^4 = 4x^3$$

and so on until we formulate the theorem:

$$\frac{d}{dx}x^n = nx^{n-1} \qquad n \geq 1, n \text{ an integer}$$

In proving theorems we can use axioms, previous theorems, and facts from earlier mathematics courses.

Example 6. In $E^n, \mathcal{M}_{mn}$, and $\mathcal{P}_n$, what is the result of multiplying a vector by the scalar $r = 0$? State a possible theorem.

Solution In E^n,

$$0(x_1, \ldots, x_n) = (0x_1, \ldots, 0x_n) = (0, 0, \ldots, 0) = \theta$$

In $\mathcal{M}_{mn}$,

$$0 \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}$$

And in $\mathcal{P}_n$,

$$0(a_n x^n + \cdots + a_1 x + a_0) = 0x^n + \cdots + 0x + 0 = \theta$$

A possible theorem is: If V is a vector space and $\mathbf{x} \in V$, then $0\mathbf{x} = \theta$.

The guess in the solution to Example 6 is actually correct.

Theorem 3.2.2. *If V is a real vector space and $\mathbf{x} \in V$, then $0\mathbf{x} = \theta$.*

Proof.

$$\begin{aligned} 0\mathbf{x} = (0+0)\mathbf{x} &= 0\mathbf{x} + 0\mathbf{x} \\ \theta = 0\mathbf{x} + [-(0\mathbf{x})] &= (0\mathbf{x} + 0\mathbf{x}) + [-(0\mathbf{x})] \\ &= 0\mathbf{x} + \{0\mathbf{x} + [-(0\mathbf{x})]\} \\ &= 0\mathbf{x} + \theta \\ &= 0\mathbf{x} \end{aligned}$$

Therefore, $\theta = 0\mathbf{x}$. □

Example 7. This is an important example because it shows that vector addition need not be related to ordinary addition and the zero vector θ need not involve the real number 0. Let $V = \{x | x \in \mathbb{R}, x > 0\}$. Define addition and scalar multiplication as follows:

Addition. For $x \in V, y \in V$, define

$$\underbrace{x \oplus y}_{\text{Vector sum}} = \underbrace{xy}_{\text{Ordinary multiplication}}$$

Multiplication. For $r \in \mathbb{R}, x \in V$, define

$$\underbrace{r \odot x}_{\substack{\text{Scalar} \\ \text{multiplication}}} = \underbrace{x^r}_{\substack{\text{Ordinary} \\ \text{exponentiation}}}$$

Show that V with these operations is a vector space. We have used $\oplus$ and $\odot$ to denote the vector space operations, to distinguish them from the ordinary operations in the solution.

Solution Before showing that V with these operations is a vector space, we look at some specific vector sums and scalar multiples. First, note that the vectors are just positive real numbers. So, for addition we have

$$2 \oplus 3 = 2 \cdot 3 = 6$$

Vector sum

From definition

$$\begin{aligned} 4 \oplus 6 &= 24 & \qquad 10 \oplus 10 &= 100 \\ 2 \odot 3 = 3^2 &= 9 & \qquad -2 \odot 3 &= 3^{-2} = \tfrac{1}{9} \end{aligned}$$

Scalar Vector

From definition

Now to show that we have a vector space, we must show that all the properties of the definition are satisfied.

Closure of addition. Let $x \in V, y \in V$. Then $x \oplus y = xy$. Since x and y are positive real numbers and the product of positive real numbers is a positive real number, $xy \in V$. Therefore, $x \oplus y \in V$, and we have closure for addition.

Closure for scalar multiplication. Let $r \in \mathbb{R}$ and $x \in V$. Then $r \odot x = x^r$. Since x is a positive real number, any real power of it is also a positive real number. Therefore $r \odot x \in V$, and we have closure for addition.

1. **Commutativity for addition**. Let $x \in V, y \in V$. Then

$$x \oplus y = xy = yx = y \oplus x$$

By definition

By definition

By commutativity of multiplication for real numbers

Therefore $x \oplus y = y \oplus x$.

2. **Associativity for addition**. Let $x \in V, y \in V, z \in V$. Then

$$(x \oplus y) \oplus z \underset{\text{By definition}}{=} (xy) \oplus z \underset{\text{By definition}}{=} (xy)z \underset{\substack{\text{Associativity}\\ \text{for multiplication}\\ \text{of real numbers}}}{=} x(yz) \underset{\text{By definition}}{=} x \oplus (yz) \underset{\text{By definition}}{=} x \oplus (y \oplus z)$$

3. **Additive identity**. To determine θ, we ask for the equation

$$x \oplus \theta = x$$

to hold. The left-hand side is $x\theta$ by definition of $\oplus$. Therefore, $\theta = 1$. That is, the positive real number 1 is the only additive identity or "zero" for this space. That is,

$$x \oplus 1 = x$$

4. **Additive inverse.** For any $x \in V, -x$ must satisfy $x \oplus (-x) = \theta$. Now we must be very careful: $-x$ is not "minus x"; it is the symbol for additive inverse, and θ is not the real number zero, it is 1. So the equation is

$$x \oplus (-x) = \theta$$

By definition From 3 above

$$x \cdot (-x) = 1$$

Therefore the unique solution for $-x$ is $(-x) = 1/x$. Since $x > 0$, we know that $1/x > 0$, so $-x \in V$.

5. **Associative law for scalar multiplication**. Let $r \in \mathbb{R}, s \in \mathbb{R}, x \in V$. Then

$$(rs) \odot x = x^{rs} \quad \text{and} \quad r \odot (s \odot x) = r \odot (x^s) = (x^s)^r$$

By laws of exponents, $(x^s)^r = x^{sr} = x^{rs}$. Therefore, $(rs) \odot x = r \odot (s \odot x)$.

6,7. **Distributive laws**. Let $r \in \mathbb{R}, s \in \mathbb{R}, x \in V, y \in V$. Then

$$(r+s) \odot x \underset{\substack{\text{By definition}\\ \text{of scalar}\\ \text{multiplication}}}{=} x^{r+s} \underset{\substack{\text{Properties}\\ \text{of exponents}}}{=} x^r x^s \underset{\substack{\text{Definition}\\ \text{of addition}}}{=} x^r \oplus x^s \underset{\substack{\text{By definition}\\ \text{of scalar}\\ \text{multiplication}}}{=} (r \odot x) \oplus (s \odot x)$$

$$r \odot (x \oplus y) = r \odot (xy) = (xy)^r = x^r y^r = x^r \oplus y^r = (r \odot x) \oplus (r \odot y)$$

8. Let $x \in V$. Then $1 \odot x = x^1 = x$.

Therefore, V with the operations as defined is a vector space.

Example 8. Let $V = \{x | x \in \mathbb{R}, x > 0\}$, and define addition and scalar multiplication as follows:

1. **Addition**. Let $x \in V, y \in V$. Define

$$x \underset{\substack{\text{Vector}\\ \text{addition}}}{+} y = x \underset{\substack{\text{Ordinary addition}\\ \text{of real numbers}}}{+} y$$

2. **Scalar multiplication**. Let $r \in \mathbb{R}, x \in V$. Define

$$rx = r \underset{\text{Ordinary multiplication}}{\cdot} x$$

Show that V is not a vector space .

Solution To show that the structure is not a vector space, all we have to show is that at least one of the axioms fails to hold true. In this case closure for scalar multiplication fails to hold since if $r < 0$, $rx = r \cdot x < 0$ and $rx \notin V$. Again we see that the truth of the axioms depends on the set V **and** the operations.

Example 9. Regarding the mass spring apparatus in Fig. 3.2.1 and the related discussion,

$$V = \{f | f(t) = c_1 \cos wt + c_2 \sin wt, t \in \mathbb{R}\}$$

with the usual definition of addition of functions and multiplication of functions is a vector space.

Solution Closure. Let $f(t) = c_1 \cos wt + c_2 \sin wt$ and $g(t) = c_3 \cos wt + c_4 \sin wt$. We have

$$\begin{aligned}(f+g)(t) &= (c_1+c_3)(\cos wt) + (c_2+c_4)(\sin wt)\\ (rf)(t) &= rc_1 \cos wt + rc_2 \sin wt\end{aligned}$$

which are in V. Commutative and associative properties hold. The zero vector is $0\ \cos wt + 0 \sin wt$, which is the identically zero function. The additive inverse of f is

$$-c_1 \cos wt - c_2 \sin wt$$

The other properties are left to the reader.

Complex Vector Spaces If in the definition of real vector space we replace real numbers by complex numbers, we have the notion of a complex vector space. The complex vector spaces we use most often in this text are $\mathbb{C}^n$ and $\mathcal{C}_{mn}$, which are defined below.

Definition 3.2.2. Vector space $\mathbb{C}^n$ is the complex vector space consisting of n-tuples $(z_1, z_2, \ldots, z_n)$ of complex numbers with the operations

$$\begin{aligned}(z_1, \ldots, z_n) + (w_1, \ldots, w_n) &= (z_1 + w_1, \ldots, z_n + w_n)\\ c(z_1, \ldots, z_n) &= (cz_1, \ldots, cz_n)\end{aligned}$$

where $(w_1, \ldots, w_n)$ is a complex n-tuple and c is any complex number.

Definition 3.2.3. $\mathcal{C}_{mn}$ is the complex vector space of $m \times n$ matrices with complex number entries along with the standard matrix operations of addition and scalar multiplication.

The zero vectors for $\mathbb{C}^n$ and $\mathcal{C}_{mn}$ are, respectively, the same as the zero vectors for E^n and $\mathcal{M}_{mn}$, as can be verified directly.

In the remainder of the text, the term **vector space** will mean the **real** vector space unless we are working specifically with a complex vector space. When you are working with complex vector spaces, it is important to remember that the vectors can be constructed by using complex numbers and that the scalars for scalar multiplication can be any complex number.

Example 10. Let $V = \{\text{hermitian } n \times n \text{ matrices}\}$, and give V the usual matrix operations. Is V a vector space?

Solution Since there is no restriction on the entries of the matrices, we are checking to see whether V with these operations is a complex vector space. Recall that the main-diagonal entries of a hermitian matrix are real numbers. Thus, in general, if A is hermitian, the scalar multiple iA is not hermitian. Therefore V is not closed under scalar multiplication and is not a vector space.

Problems 3.2

In Probs. 1 to 20, a set V and operations of scalar multiplication and vector addition are given. Determine whether V is a vector space. If V fails to be a vector space, state an axiom which fails to hold. In a given problem, if the objects in V can be constructed by using complex numbers, the problem is to determine whether V is a complex vector space.

1. Let $V = \{\text{ordered triples } (x_1, x_2, 0); x_1, x_2 \in \mathbb{R}\}$, the operations as in E^3.

2. Let $V = \{\text{ordered triples } (x_1, x_2, 1); x_1, x_2 \in \mathbb{R}\}$, the operations as in E^3.

3. Let $V = \{\text{ordered pairs } (x_1, x_2); x_1, x_2 \in \mathbb{R}\}$ with the operations $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, y_2 + x_2)$ and $r(x_1, x_2) = (rx_1, rx_2)$.

4. Let $V = \{\text{ordered pairs } (x_1, x_2); x_1, x_2 \in \mathbb{R}\}$ with the operations $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $r(x_1, x_2) = (rx_1, x_2)$.

5. Let $V = \mathbb{R}$, and let the operations be the standard addition and multiplication in $\mathbb{R}$.

6. Let $V = \{n \times n \text{ matrices with positive entries}\}$ with the usual matrix operations. (Entries must be real to be compared to zero.)

7. Let $V = \{n \times n \text{ real symmetric matrices}\}$ with the usual matrix operations.

8. Let $V = \{n \times n \text{ real skew-symmetric matrices}\}$ with the usual matrix operations.

9. Let $V = \{n \times n \text{ upper triangular matrices}\}$ with the usual matrix operations.

10. Let $V = \{n \times n$ diagonal matrices$\}$ with the usual matrix operations.

11. Let $V = \{$functions defined for all x in $\mathbb{R}$ with $f(0) = 0\}$ with operations $(f + g)(x) = f(x) + g(x)$ and $(rf)(x) = r[f(x)]$.

12. Let $V = \{$functions defined for all x with $f(0) = 1\}$ with operations as in Prob. 11.

13. Let $V = \{n \times n$ nonsingular matrices$\}$ with the usual matrix operations.

14. Let $V = \{n \times n$ singular matrices$\}$ with the usual matrix operations.

15. Let $V = \{n \times n$ nilpotent matrices$\}$ with the usual matrix operations.

16. Let $V = \{n \times n$ idempotent matrices$\}$ with the usual matrix operations.

17. Let $V = \{n \times n$ matrices A with $A^2 = I\}$ with the usual matrix operations.

18. Let $V = \{$real-valued functions defined on $\mathbb{R}$ with $f(x) > 0$, for all $x\}$ with the usual operations.

19. Let V be as in Prob. 18, but give V the operations $(f+g)(x) = f(g)g(x)$ and $(rf)(x) = [f(x)]^r$, for r a real number.

20. Let $V = \{n \times n$ matrices with sum of main diagonal entries equal to zero$\}$ with the usual matrix operations.

21. Let C be a fixed $n \times n$ matrix. Let $V = \{A_{n \times n}$ such that $AC = 0\}$. Given the usual matrix operations, is V a vector space?

22. Prove that in a real vector space V, $r\theta = \theta$ for all $r \in \mathbb{R}$. (**Hint:** Mimic the proof of Theorem 3.2.2.)

23. Prove that in a real vector space $V, (-1)\mathbf{x} = -\mathbf{x}$ for all $x \in V$. (**Hint:** Mimic the proof of Theorem 3.2.2.)

24. Prove that in a real vector space V, if $\mathbf{x} \in V$, $r \in \mathbb{R}$, and $r\mathbf{x} = \theta$ then $\mathbf{x} = \theta$ or $r = 0$.

25. Let $V = \{Z\}$ (a set consisting of one element) and define

$$\begin{aligned} Z + Z &= Z \\ rZ &= Z \qquad \text{For all } r \in \mathbb{R} \end{aligned}$$

Is V a real vector space?

26. Work Prob. 25 with the second operation being $cZ = Z$ for all $c \in \mathbb{C}$. Is V a complex vector space? (Check the vector space axioms.)

27. In the list of properties of vector space operations, could 3 and 4 be reversed in order? Explain.

28. Let $V = \{(x_1, x_2) | x_1 + x_2 = 1,\ 0 \le x_1 \le 1, 0 \le x_2 \le 1\}$ with the operations $(x_1, x_2) + (y_1, y_2) = \left(\frac{1}{2}(x_1 + y_1), \frac{1}{2}(x_2 + y_2)\right)$ and $r(x_1, x_2) = (x_1, x_2)$. Are the closure axioms satisfied? Is V a real vector space?

3.3 THE SUBSPACE PROBLEM

Certain applications involve the use of subsets of vector spaces which are vector spaces also. We will see an example from coding theory at the end of the section. Another example is the set $\mathcal{S}_n$ of all $n \times n$ symmetric matrices (used in least squares problems; see App III) with the usual matrix operations which is a subset of $\mathcal{M}_{nn}$. $\mathcal{S}_n$ is a vector space (Prob. 7 of Sec. 3.2). In this case we say that $\mathcal{S}_n$ is a **subspace** of $\mathcal{M}_{nn}$.

Definition 3.3.1. Let V be a real or complex vector space, and let W be a subset of V with W inheriting the operations of V. We say that W is a **subspace** of V if W with the inherited operations is a vector space.

Given a subset of a real or complex vector space, we often need to know whether the subset is a subspace. Thus we have the following.

> **Subspace problem.**
> Given a subset W of a vector space V, with W having the same operations as V, determine whether W is a subspace of V.

Example 1. Let $V = E^3$, and let $W = \{\mathbf{x} \in E^e | \mathbf{x} = (x_1, x_2, 0)\}$. Solve the subspace problem for W and V.

Solution By its definition W is a subset of V; we must determine whether W with the operations inherited from E^3 is a vector space.

Closure. Let $r \in \mathbb{R}, \mathbf{x} \in W, \mathbf{y} \in W$. We check to see whether $\mathbf{x} + \mathbf{y} \in W$ and $r\mathbf{x} \in W$. We have

$$\mathbf{x} + \mathbf{y} = (x_1, x_2, 0) + (y_1, y_2, 0) = (x_1 + y_1, x_2 + y_2, 0)$$

Since the third component is 0, $\mathbf{x} + \mathbf{y} \in W$. For the scalar multiplication

$$r\mathbf{x} = (rx_1, rx_2, r0) = (rx_1, rx_2, 0)$$

and since the third component is 0, $r\mathbf{x} \in W$.

1. **Community of addition.** Let $\mathbf{x} \in W, \mathbf{y} \in W$. Then

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, 0) = (y_1 + x_1, y_2 + x_2, 0) = \mathbf{y} + \mathbf{x}$$

2. **Associativity of addition**. This is similar to 1; see the problems.

3. **Existence of zero**. We let $\mathbf{x} \in W$ and see whether there is any vector in W which acts as an additive identity. Because $(0, 0, 0) \in W$ (the last component is zero) and $\mathbf{x} + (0, 0, 0) = \mathbf{x}$, $\theta = (0, 0, 0)$ acts as zero for W.

4. **Additive inverse**. Let $\mathbf{x} = (x_1, x_2, 0)$. Since $(-x_1, -x_2, 0)$ is also in W and

$$(x_1, x_2, 0) + (-x_1, -x_2, 0) = (0, 0, 0) = \theta$$

we know that $-\mathbf{x} = (-x_1, -x_2, 0)$ is in W.

5-8. These are straightforward and left to the problems.

A close examination of Example 1 shows that most axioms actually are true by "inheritance." For example, in showing $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$, the fact that $\mathbf{x}$ and $\mathbf{y}$ were in W was not important; since they were in V and commutativity holds for all $\mathbf{x}$ and $\mathbf{y}$ in V, commutativity holds in **any** subset of V. Thus it is the properties of closure and the existence of zero and additive inverse in W that must be checked. However, we can show that closure leads to the existence of zero and an additive inverse in W.

Theorem 3.3.1. *Let V be a real vector space and W a nonempty subset of V. Then W is a subspace of V if and only if (1) $\mathbf{x}, \mathbf{y}$ in W implies $\mathbf{x} + \mathbf{y}$ is in W and (2) r in $\mathbb{R}$ and $\mathbf{x}$ in W implies $r\mathbf{x}$ is in W.*

This theorem contains the solution to the subspace problem. It is emphasized that the theorem holds for complex vector spaces with $\mathbb{R}$ replaced by $\mathbb{C}$. Before proving this theorem, we give some examples.

Example 2. Let $V = E^3$ and $W = \{\mathbf{x}|\mathbf{x} = a(1,0,2) + b(1,-1,3)\}$, where a and b can be any real numbers. Is W a subspace of V?

Solution Let $\mathbf{x} \in W, \mathbf{y} \in W$, and consider $\mathbf{x}+\mathbf{y}$. We have

$$\begin{aligned}\mathbf{x}+\mathbf{y} &= (a(1,0,2)+b(1,-1,3)) + (c(1,0,2)+d(1,-1,3)) \\ &= a(1,0,2)+b(1,-1,3)+c(1,0,2)+d(1,-1,3) \\ &= (a+c)(1,0,2)+(b+d)(1,-1,3)\end{aligned}$$

Hence $\mathbf{x}+\mathbf{y}$ is in the form of an element of W; that is, $\mathbf{x}+\mathbf{y} \in W$. Consider now $r\mathbf{x} = r(a(1,0,2)+b(1,-1,3)) = ra(1,0,2)+rb(1,-1,3)$. Thus $r\mathbf{x} \in W$, and W is a subspace by Theorem 3.3.1.

Example 3. Let $V = E^3$ and $W = \{\mathbf{x}|\mathbf{x} = (x_1, x_2, 1)\}$, and consider $\mathbf{x}+\mathbf{y}$. Now

$$\mathbf{x}+\mathbf{y} = (x_1, x_2, 1) + (y_1, y_2, 1) = (x_1+y_1, x_2+y_2, 2)$$

The last component of $\mathbf{x}+\mathbf{y}$ is not equal to 1. Therefore $\mathbf{x}+\mathbf{y} \notin W$, and so W is **not** a subspace of V.

Example 4. Let $V = \mathcal{C}_{nn}$ and let $W = \{n \times n \text{ hermitian matrices}\}$. Is W a subspace of V?

Solution From the solution of Example 10 of Sec. 3.2. We know that if A is hermitian, then iA need not be hermitian. For example,

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

is hermitian, but

$$iA = \begin{pmatrix} i & -1 \\ 1 & 2i \end{pmatrix}$$

is not hermitian. Therefore, W is not closed under scalar multiplication and is not a subspace of $\mathcal{C}_{nn}$.

Proof of Theorem 3.3.1. This is an "if and only if" theorem: p holds if and only if q holds. To prove it, we must prove two things:

1. If p holds, then q holds.

2. If q holds, then p holds.

Part 1. If W is a subspace, then $\mathbf{x}+\mathbf{y} \in W$ and $r\mathbf{x} \in W$ for all $r \in \mathbb{R}$, $\mathbf{x} \in W$, $\mathbf{y} \in W$.

If W is a subspace of V, then W is a vector space and all axioms hold. In particular, closure holds. Therefore $r\mathbf{x} \in W$ by closure for scalar multiplication, and $\mathbf{x}+\mathbf{y} \in W$ by closure for vector addition.

Part 2. If $\mathbf{x}+\mathbf{y} \in W$ and $r\mathbf{x} \in W$ for all $r \in \mathbb{R}$, $\mathbf{x} \in W$, $\mathbf{y} \in W$, then W is a subspace of V.

We must check all the vector space axioms for W, keeping in mind the hypothesis that $\mathbf{x}+\mathbf{y} \in W$ and $r\mathbf{x} \in W$ for all $r \in \mathbb{R}$, $\mathbf{x} \in W$, $\mathbf{y} \in W$. As we have mentioned, all axioms except closure, zero, and additive inverse follow by the heredity. The closure is just restatement of our hypothesis for this part of the theorem. Now let $\mathbf{x} \in W$ and $r = -1$. Then by hypothesis $(-1)\mathbf{x} \in W$, so $(-1)\mathbf{x}+\mathbf{x} \in W$. But

$$
\begin{aligned}
(-1)\mathbf{x}+\mathbf{x} &= (-1)\mathbf{x}+1\mathbf{x} \\
&= (-1+1)\mathbf{x} && \text{(Distributive law)} \\
&= 0\mathbf{x} \\
&= \theta && \text{(Theorem 3.2.2)}
\end{aligned}
$$

Therefore $\theta \in W$. (We note that Theorem 3.2.2 holds for all of V so we can apply it to W.)

To show additive inverses are in W, let $\mathbf{x} \in W$ and $r = -1$. We have $(-1)\mathbf{x} \in W$ and

$$\mathbf{x}+(-1)\mathbf{x} = [1+(-1)]\mathbf{x} = 0\mathbf{x} = \theta \in W$$

Therefore, additive inverses are in W. $\square$

Example 5. Let V be any vector space. Then V itself is a subspace.

Example 6. Let V be any vector space, and $W = \{\theta\}$. Then W is a subspace because $r\theta = \theta$, $\theta+\theta = \theta$, for r real or complex.

Because any vector space V has V and $\{\theta\}$ as subspaces, these are called the **trivial subspaces** of V. All other subspaces of V are called **proper subspaces**, or nontrivial subspaces, of V.

Example 7. For the real case: Let $\mathbf{x} = a_1\mathbf{u} + a_2\mathbf{z}$, and $\mathbf{y} = b_1\mathbf{u} + b_2\mathbf{z}$, so that $\mathbf{x}$ and $\mathbf{y}$ are in V. Let $W = \{\mathbf{x}|\mathbf{x} = c_1\mathbf{u} + c_2\mathbf{z}$, where c_1 and c_2 can be any real (complex) numbers$\}$. Show that W is a subspace of V.

Solution For the real case: Let $\mathbf{x} = a_1\mathbf{u} + a_2\mathbf{z}$, and $\mathbf{y} = b_1\mathbf{u} + b_2\mathbf{z}$, so that $\mathbf{x}$ and $\mathbf{y}$ are in W. Consider $\mathbf{x} + \mathbf{y}$:

$$\mathbf{x} + \mathbf{y} = \underbrace{(a_1 + b_1)}_{\text{Real numbers}}\ \mathbf{u} + \underbrace{(a_2 + b_2)}_{\text{Real numbers}}\ \mathbf{z} \in W$$

Closure for scalar multiplication is similar. By Theorem 3.3.1, W is a subspace of V.

Example 8. Let $V = \{x \in \mathbb{C}^2 | x$ has purely imaginary components$\}$. Now V is not a subspace of $\mathbb{C}^2$ because $i(i, i) = (-1, -1)$ which is not in V.

Example 9. Let $V = \mathcal{M}_{22}$ and $W = \{$invertible 2×2 matrices$\}$. Determine whether W is a subspace of $\mathcal{M}_{22}$.

Solution 1 If W were to be a subspace, θ would have to be in W. But

$$\theta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

is not invertible and cannot be in W. So W is not a subspace.

Solution 2 Let

$$\mathbf{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

be invertible, and consider $\mathbf{x} + \mathbf{y}$. Now

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} a + e & b + f \\ c + g & d + h \end{pmatrix}$$

which is invertible if and only if its determinant is nonzero, that is, if and only if

$$\begin{aligned} (a + e)(d + h) - (c + g)(b + f) &\neq 0 \\ ad + de + ah + eh - bc - gb - cf - gf &\neq 0 \\ \underbrace{(ad - bc)}_{\neq 0} + (de + ah - gb - cf) + \underbrace{(ef - gh)}_{\neq 0} &\neq 0 \end{aligned}$$

It is not clear that the sum on the left-hand side should be nonzero. This solution shows that testing $\mathbf{x} + \mathbf{y}$ may be tedious. The first solution is preferable in this case.

Example 10. Let $A \in \mathcal{M}_{nn}$. Show that the set $W(\subseteq \mathcal{M}_{n1})$ of all solutions to $A_{n\times n}X_{n\times 1} = 0_{n\times 1}$ is a subspace of $\mathcal{M}_{n1}$.

Solution Let $U_{n\times 1}$ and $V_{n\times 1}$ be solutions of the homogeneous equation $A_{n\times n}X_{n\times 1} = 0_{n\times 1}$. To check the closures, let $r \in \mathbb{R}$ and calculate

$$\begin{aligned} A(U+V) &= AU + AV = 0 + 0 = 0 \\ A(rU) &= rAU = r0 = 0 \end{aligned}$$

Therefore $U+V$ and rU are solutions in $W; W$ is a subspace.

Example 11. Let W be the subset of $\mathcal{C}_{nn}$ defined by $W = \{Z|\overline{Z} = Z\}$. Is W a subspace of $\mathcal{C}_{nn}$? (Note that W is not $\mathcal{W}_{nn}$ because our "scalars" for scalar multiplication come from $\mathbb{C}$ now.)

Solution First we check for closure of addition. Let Y and Z be in W. We have

$$\overline{Y+Z} \underset{\substack{\text{Previous}\\ \text{property}}}{=} \overline{Y} + \overline{Z} \underset{Y,Z\in W}{=} Y + Z$$

so W is closed under addition. Now check multiplication by letting $c \in \mathbb{C}$ and $Z \in W$; we must see whether $\overline{cZ} = cZ$. However $\overline{cZ} = \bar{c}\overline{Z} = \bar{c}Z$. Because $\bar{c} \neq c$ unless c is real, closure of multiplication fails and W is not a subspace.

Table 3.3.1

ADDITION AND MULTIPLICATION FOR $F = \{0, 1\}$

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

The idea of subspace has important applications. In fact, subspaces are used to define certain concepts in coding theory.

Example 12. (Linear codes) This example is of an unusual vector space for which subspaces have applications in coding theory. First, consider $F = \{0, 1\}$ with the operations of multiplication and addition defined as in $\mathbb{R}$ except that we define $1+1=0$ (see Table 3.3.1). The set F along with these operations is called a **commutative field**. This just means that it enjoys the properties of the real number system. For the vector space let F^n be the set of n-tuples of elements of F with operations

$$(a_1, \dots , a_n) + (b_1, \dots , b_n) = (a_1 + b_1, \dots , a_n + b_n)$$
$$r(a_1, \dots , a_n) = (ra_1, \dots , ra_n)$$

Note that a_k, b_k, and r can be only 0 or 1. F^n with these operations is a vector space. How is it used in coding theory? Well, in this case we have what is called a **binary channel** because the components of the vectors in F^n can be chosen from a set of only **two** symbols. A subset of F^n is called a **linear code** if and only if it is a **subspace** of F^n. For example, F^3 contains eight possible vectors. The subset $V = \{(0, 0, 0), (1, 0, 0)\}$ is a linear code because we have closure: we can just check all possible sums and scalar products. Note that $(1, 0, 0) + (1, 0, 0) = (0, 0, 0)$.

Problems 3.3

In Probs. 1 to 14, a vector space V and subset W are given. Determined whether W is a subspace.

1. $V = \mathcal{M}_{mn}, W = \{x \in \mathcal{M}_{mn}$ with nonnegative entries$\}$
2. $V = \mathcal{M}_{nn}, W = \{$symmetric $n \times n$ matrices$\}$
3. (Define the trace of an $n \times n$ matrix A as tr $A = a_{11} + a_{22} + \cdots + a_{nn}$.) Let $V = \mathcal{M}_{nn}$ and $W = \{A \in \mathcal{M}_{nn}$ with tr $A = 0\}$
4. $V = E^3, W = \{(x_1, x_2, x_3)|ax_1 + bx_2 + cx_3 = 0$, where a, b, c are fixed numbers$\}$
5. $V = \mathcal{M}_{22}, W = \{$noninvertible matrices$\}$
6. $V = E^2, W = \{(x_1, x_2)|{x_1}^2 + {x_2}^2 = 1\}$
7. $V = \mathcal{M}_{nn}, W = \{A \in \mathcal{M}_{nn}|A = -A^T\}$

8. $V = E^3, W = \{\mathbf{x} \in E^3 | \mathbf{x} \text{ is perpendicular to } (a, b, c)\}$
9. $V = \mathcal{C}_{nn}, W = \{n \times n \text{ matrices with real entries}\}$
10. $V = \mathbb{C}^n, W = \{n\text{-tuples with real entries}\}$
11. $V = \mathcal{P}_n, W = \{f \text{ in } \mathcal{P}_n | f(0) = 0\}$
12. $V = \mathcal{P}_n, W = \{f \text{ in } \mathcal{P}_n | f(0)0 \neq 0\}$
13. $V = \mathcal{P}_n, W = \{f \text{ in } \mathcal{P}_n | f(1) = 0\}$
14. $V = \mathcal{P}_n, W = \{f \text{ in } \mathcal{P}_n | f(1) \neq 0\}$
15. Consider the system of equations $AX = 0$, where A is in $\mathcal{C}_{nn}$. Show that the set of all solutions of $AX = 0$ is a vector space under the usual operations.
16. If $B_{n\times 1} \neq 0$ in the system of equations $AX = B$, where A is $n \times n$ show that the set of solutions cannot be a vector space, given the standard matrix operations.
17. Let A and B be square matrices. Show that $\text{tr}(A + B) = \text{tr } A + \text{tr } B$, $\text{tr}(rA) = r(\text{tr } A)$, $\text{tr}(AB) = \text{tr}(BA)$. (See Prob. 3 for the definition of $\text{tr } A$.)
18. Using the terminology of Example 12, show that

$$V = \{(0, 0, 0, 0), (1, 1, 0, 1), (1, 0, 0, 1), (0, 1, 0, 0)\}$$

is a linear code.
19. Let $V = E^2$, which can be associated with the plane. Show that $W = \{(x_1, x_2) | x_2 = mx_1\}$ is a subspace of V. Show that $U = \{(x_1 x_2) | x_2 = mx_1 + b, b \neq 0\}$ is not a subspace of V. This shows that if a straight line is to represent a subspace of the plane, the line must pass through the origin.
20. Let $V = E^3$. Show that $W = \{(x_1, x_2, x_3) | ax_1 + bx_2 + cx_3 = d\}$ is a subspace of V if and only if $d = 0$. This shows that if a plane is to represent a subspace of three-space, it must pass through the origin.

21. Let $V = E^3$. Show that $W = \{(x_1, x_2, x_3)|x_1 = at = k_1,\ x_2 = bt + k_2,\ x_3 = ct + k_3, t \text{ real}\}$ is a proper subspace of V if and only if there exists T such that $aT + k_1 = bT + k_2 = cT + k_3 = 0$. This shows that a line represents a subspace of three-space if and only if the line passes through the origin.

22. Fill in the details for the solution of Example 1.

3.4 LINEARLY INDEPENDENT SETS OF VECTORS

The equation $x + 2y - z = 0$ has general solution $(x, y, z) = (s - 2r, r, s)$, where r and s are any numbers. Any solution can be written in terms of the vectors (1,0,1) and $(-2, 1, 0)$ from E^3 as

$$(s - 2r, r, s) = s(1, 0, 1) + r(-2, 1, 0)$$

That is, an infinite number of solutions can be constructed in terms of just two vectors, and analysis of the solutions can be performed by considering just these two vectors. To use similar methods of analysis in vector spaces, we will need the concepts of span and linear independence of sets of vectors. Both concepts involve linear combinations of vectors.

Definition 3.4.1. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be vectors in a vector space V. A **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is any sum of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

where the numbers $c_1, c_2, \ldots, c_n$ are called the **coefficient** of the linear combination.

Example 1. Write five linear combinations of the vectors $(1, -1)$, (1,2), and (3,0) in E^2.

Solution Five possibilities are

$$\begin{aligned}
0(1, -1) + 0(1, 2) + 0(3, 0) &= (0, 0)\\
3(1, -1) - 1(1, 2) + 7(3, 0) &= (23, -5)\\
1(1, -1) + 0(1, 2) + 0(3, 0) &= (1, -1)\\
2(1, -1) + 1(1, 2) + 5(3, 0) &= (18, 0)\\
2(1, -1) + 1(1, 2) - 1(3, 0) &= (0, 0)
\end{aligned}$$

Note that the first and last linear combinations yield the same vector (0,0), even though the coefficients are not the same. The last four linear combinations are called nontrivial because in each at least one coefficient is nonzero.

Example 2. Write $(7, -2, 2)$ in E^3 as a linear combination of $(1, -1, 0), (0, 1, 1)$, and (2,0,1).

Solution We want to find c_1, c_2, c_3 so that

$$(7, -2, 2) = c_1(1, -1, 0) + c_2(0, 1, 1) + c_3(2, 0, 1)$$

or

$$(7, -2, 2) = (c_1 + 2c_3, -c_1 + c_2, c_2 + c_3)$$

which yields equations

$$\begin{aligned} 7 &= c_1 + 2c_3 \\ -2 &= -c_1 + c_2 \\ 2 &= c_2 + c_3 \end{aligned}$$

The solution is $c_1 = 1, c_2 = -1, c_3 = 3$, so

$$(7, -2, 2) = (1, -1, 0) - (0, 1, 1) + 3(2, 0, 1)$$

Example 3. Can $(3, -1, 4)$ be written as a linear combination of $(1, -1, 0)$, (0,1,1), and $(3, -5, -2)$?

Solution We check to see whether the equation

$$(3, -1, 4) = c_1(1, -1, 0) + c_2(0, 1, 1) + c_3(3, -5, -2)$$

has a solution. This is equivalent to

$$\begin{aligned} 3 &= c_1 + 3c_3 \\ -1 &= -c_1 + c_2 - 5c_3 \\ 4 &= c_2 - 2c_3 \end{aligned}$$

In reduced form this is

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

and there is no solution. Hence $(3, -1, 4)$ cannot be written as a linear combination of the given vectors.

Definition 3.4.2. Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a set of vectors in the vector space V. The **span** of $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is the set of all possible linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. The notation is span $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$.

The span of a set of vectors from V is actually a subspace of V.

Theorem 3.4.1. *If $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subset V$ with V a vector space, then span $\mathcal{S}$ is a subspace of V.*

Proof. Let $\mathbf{x}$ and $\mathbf{y}$ be in span $\mathcal{S}$. Then $\mathbf{x}$ and $\mathbf{y}$ are linear combinations:

$$\begin{aligned}\mathbf{x} &= a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \\ \mathbf{y} &= b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n\end{aligned}$$

Now

$$\mathbf{x} + \mathbf{y} = (a_1 + b_1)\mathbf{v}_1 + \cdots + (a_n + b_n)\mathbf{v}_n$$

and

$$r\mathbf{x} = (ra_1)\mathbf{v}_1 + \cdots + (ra_n)\mathbf{v}_n \quad r \text{ a scalar}^1$$

which are also in span $\mathcal{S}$. Therefore, span $\mathcal{S}$ is a subspace of V. It is called the **subspace of V spanned by $\mathcal{S}$.** □

Example 4. Example 2 could have been worded as follows: Show that $(7, -2, 2)$ is in span $\{(1, -1, 0),\ (0,1,1),\ (2,0,1)\}$. The solution would be exactly the same.

Example 5. Example 3 could have been worded as follows: Is $(3, -1, 4)$ in span $\{(1, -1, 0),\ (0,1,1),\ (3, -5, -2)\}$? The solution would be exactly the same. The answer is that $(3, -1, 4)$ is not in span $\{(1, -1, 0)\ (0,1,1),$ $(3, -5, -2)\}$.

Example 6. We have already seen that the solutions of $x + 2y - z = 0$ can be written as

$$s(1, 0, 1) + r(-2, 1, 0)$$

for all complex s and r. Another way to say this is that all solutions form the subspace span $\{(1, 0, 1), (-2, 1, 0)\}$.

In some instances span $\mathcal{S}$ may be all of V.

[1] The statement of the theorem is for any vector space. In the real case the scalars would come from $\mathbb{R}$; in the complex case, from $\mathbb{C}$.

Example 7. Let $S = \{1, x, x^2\}$, which is a set of vectors in $\mathcal{P}_2$. Then span $S = \{a_0 1 + a_1 x + a_2 x^2$, where a_0, a_1, a_2 can be any real numbers$\}$. Thus span $S = \mathcal{P}_2$.

Example 8. Let $S = \{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1)\} \subseteq E^4$. Span S is all E^4, since span S is all vectors of the form

$$x_1(1,0,0,0) + x_2(0,1,0,0) + x_3(0,0,1,0) + x_4(0,0,0,1) = (x_1, x_2 x_3, x_4)$$

where x_1, x_2, x_3, and x_4 can be any real numbers.

Example 9. Show the span $\{(1,0,1), (-1,2,3), (0,1,-1)\}$ is all of E^3.

Solution We must show that any vector (a,b,c) in E^3 can be written as a linear combination of the three given vectors. That is, we must show that there are constants c_1, c_2, c_3 so that

$$(a,b,c) = c_1(1,0,1) + c_2(-1,2,3) + c_3(0,1,-1)$$

regardless of what real values a, b, and c take. The last equation is equivalent to

$$\begin{aligned} a &= c_1 - c_2 \\ b &= \phantom{c_1 + {}} 2c_2 + c_3 \\ c &= c_1 + 3c_2 - c_3 \end{aligned}$$

which has solutions

$$c_1 = \frac{5a + b + c}{6} \quad c_2 = \frac{b + c - a}{6} \quad c_3 = \frac{a + 2b - c}{3}$$

Therefore $(a,b,c) \in$ span $\{(1,0,1), (-1,2,3), (0,1,-1)\}$, and the span is all of E^3.

Example 10. Determine whether span $\{(1,-1,0), (0,1,1), (3,-5,-2)\}$, is all of E^3.

Solution Let (a,b,c) be an arbitrary vector in E^3. We want to know whether it is possible to write

$$(a,b,c) = c_1(1,-1,0) + c_2(0,1,1) + c_3(3,-5,-2)$$

The last equation is equivalent to

$$\left(\begin{array}{rrr|c} 1 & 0 & 3 & a \\ -1 & 1 & -5 & b \\ 0 & 1 & -2 & c \end{array}\right)$$

which reduces to

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

Therefore a solution exists only if $c - a - b = 0$; but this places a restriction on (a, b, c), and so the very first equation cannot be solved for an arbitrary vector (a, b, c). Therefore the span of the given vectors is not all E^3.

The question in Example 10 could have been asked in a slightly different way.

Example 11. Describe span S, where $S = \{(1, -1, 0), (0, 1, 1), (3, -5, -2)\}$.

Solution Suppose (a, b, c) is in span S. Then the equation

$$(a, b, c) = c_1(1, -1, 0) + c_2(0, 1, 1) + c_3(3, -5, -2)$$

must be solvable. Working as in Example 10, we conclude that $c - a - b = 0$. Thus span $S = \{(a, b, c)|c = a+b\}$. That is, the span of S is all vectors whose third component is the sum of the first two components. So, for example, $(1, 3, 5) \notin$ span S and $(1, 3, 4) \in$ span S.

The vector space E^2 is spanned by $S = \{(1, 0), (0, 1)\}$. It is also spanned by a larger set $S' = \{(1, 0), (0, 1), (1, 1)\}$. As we will see later, E^2 and functions on E^2 can be analyzed by using spanning sets; hence for economy's sake, we want to be able to find the smallest possible spanning sets for vector spaces. To do this, the idea of **linear independence** is required.

Definition 3.4.3. A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of vectors in a vector space V is called **linearly independent** if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \theta$$

is $c_1 = c_2 = \cdots = c_n = 0$. If the set is not linearly independent, it is called **linearly dependent**.

To determine whether a set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent or linearly dependent, we need to find out about the solution of

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \theta$$

If we find (by actually solving the resulting system or by any other technique) that only the trivial solution $c_1 = c_2 = \cdots = c_n = 0$ exists, then S is linearly independent. However, if one or more of the c_k's is nonzero, then the set S is linearly dependent.

Example 12. Determine whether $S = \{(1,0),(0,1)\}$ is linearly independent.

Solution Consider

$$c_1(1,0) + c_2(0,1) = \theta = (0,0)$$

This equation is equivalent to

$$(c_1, c_2) = (0,0)$$

which has only $c_1 = 0, c_2 = 0$ as a solution. Therefore, S is linearly independent.

Example 13. Is $S = \{(1,0),(0,1),(1,-1)\}$ linearly independent?

Solution Consider

$$c_1(1,0) + c_2(0,1) + c_3(1,-1) = \theta = (0,0)$$

which is equivalent to

$$\begin{aligned} c_1 + c_3 &= 0 \\ c_2 - c_3 &= 0 \end{aligned}$$

This system has solution $c_3 = k, c_1 = -k, c_2 = k$; if $k \neq 0$, then we have a nontrivial solution, and so S is not linearly independent—it is linearly dependent.

Example 14. Determine whether $S = \{1+x, x+x^2, 1+x^2\}$ is linearly independent in $\mathcal{P}_2$.

Solution Consider

$$c_1(1+x)+c_2(x+x^2)+c_3(1+x^2)=\theta=0+0x+0x^2$$

By collecting terms on the left-hand side, this equation can be rewritten

$$(c_1+c_3)+(c_1+c_2)x+(c_2+c_3)x^2=0+0x+0x^2=\theta$$

From algebra we know that a polynomial is identically zero only when all the coefficients are zero. So we have

$$\begin{aligned} c_1 \quad\;\; + c_3 &= 0 \\ c_1 + c_2 \quad\;\; &= 0 \\ c_2 + c_3 &= 0 \end{aligned}$$

which has only the trivial solution. Therefore, S is linearly independent.

Example 15. The set $S=\{\theta\}$ is linearly dependent in any real or complex vector space because $c_1\theta=\theta$ has nontrivial solution $c_1=1$.

Linear dependence of a set of two or more vectors means that at least one of the vectors in the set can be written as a linear combination of the others. Recall Example 13 and the set $S=\{(1,0),(0,1),(1,-1)\}$. In Fig. 3.4.1 we have shown geometrically the dependence of the vectors in S. A general statement of this situation is as follows:

Theorem 3.4.2. *Let $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ be a set of at least two vectors $(n\geq 2)$ in a vector space V. Then S is linearly dependent if and only if one of the vectors in S can be written as a linear combination of the rest.*

Proof. $(\Rightarrow)$ If S is linearly dependent, then there are constants $c_1,c_2,\ldots,c_n$, some of which are nonzero, such that

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_n\mathbf{v}_n=\theta$$

Suppose c_k $(1\leq k\leq n)$ is a nonzero coefficient in the linear combination. Then

$$c_k\mathbf{v}_k=-c_1\mathbf{v}_1-c_2\mathbf{v}_2-\cdots-c_{k-1}\mathbf{v}_{k-1}-c_{k+1}\mathbf{v}_{k+1}-\cdots-c_n\mathbf{v}_n$$

and since $c_k\neq 0$,

$$\mathbf{v}_k=-\frac{c_1}{c_k}\mathbf{v}_1-\frac{c_2}{c_k}\mathbf{v}_2-\cdots-\frac{c_{k-1}}{c_k}\mathbf{v}_{k-1}-\frac{c_{k+1}}{c_k}\mathbf{v}_{k+1}-\cdots-\frac{c_n}{c_k}\mathbf{v}_n$$

Therefore, $\mathbf{v}_k$ is a linear combination of the other vectors in S.

($\Leftarrow$) Suppose $\mathbf{v}_k = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_{k-1}\mathbf{v}_{k-1} + d_{k+1}\mathbf{v}_{k+1} + \cdots + d_n\mathbf{v}_n$. Then, adding $(-1)\mathbf{v}_k$ to both sides, we have

$$\theta = d_1\mathbf{v}_1 + \cdots + (-1)\mathbf{v}_k + \cdots + d_n\mathbf{v}_n$$

Because the coefficient of $\mathbf{v}_k$ is nonzero, the set S is linearly dependent $\square$

Example 16. Show that

$$S = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 6 & 9 \end{pmatrix} \right\}$$

is linearly dependent in $\mathcal{M}_{22}$. Write one of the vectors as a linear combination of the others.

Solution Consider

$$c_1 \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} + c_2 \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} + c_3 \begin{pmatrix} -1 & 2 \\ 6 & 9 \end{pmatrix} = \theta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

This equation is equivalent to

$$\begin{aligned} c_1 - c_2 - c_3 &= 0 \\ c_1 - 2c_3 &= 0 \\ 2c_2 + 6c_3 &= 0 \\ 3c_1 + c_2 + 9c_3 &= 0 \end{aligned}$$

which reduces to

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

Therefore $c_1 = -2k, c_2 = -3k, c_3 = k$ is a solution, where k is arbitrary. Thus the set S is linearly dependent. Choosing $k = 1$, we have

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

and we can write

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

Of course, we could also write

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

or

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

Some Geometry of Spanning Sets in E^2 and E^3 The span of a single nonzero vector is a line containing the origin. Span $\{\mathbf{v}\}$ is all multiples of $\mathbf{v}$, which is all position vectors in the same direction as $\mathbf{v}$ (see Fig. 3.4.2). The terminal points of these vectors form the line with vector equation

$$\mathbf{r} = t\mathbf{v} + \theta$$

The span of two independent vectors is a plane containing the origin. To see this in E^3, let $\mathbf{v}$ and $\mathbf{w}$ be given by (a, b, c) and (d, e, f), respectively. The plane containing $\mathbf{v}$ and $\mathbf{w}$ has normal vector $\mathbf{v} \times \mathbf{w}$ and vector equation

$$(\mathbf{v} \times \mathbf{w}) \cdot (x - 0, y - 0, z - 0) = 0$$

If we calculate $\mathbf{v} \times \mathbf{w}$ and write the vector equation, we find

$$(bf - ce)x + (cd - af)y + (ae - bd)z = 0$$

where (x, y, z) is a vector in the plane. However, if (x, y, z) is to be a linear combination of $\mathbf{v}$ and $\mathbf{w}$, we must have

$$\begin{aligned} x &= c_1 a + c_2 d \\ y &= c_1 b + c_2 e \\ z &= c_1 c + c_2 f \end{aligned}$$

This system reduces to $(a \neq 0)$

$$\left(\begin{array}{cc|c} 1 & d/a & x/a \\ 0 & 1 & (ay - bx)(ax - cd) \\ 0 & 0 & (az - cx)(ae - bd) - (ay - bx)(af - cd) \end{array}\right)$$

which has a solution if and only if

$$(az - cx)(ae - bd) - (ay - bx)(af - cd) = 0$$

which holds if and only if the vector equation holds. So the span of two independent vectors is the plane containing the vectors. See Fig. 3.4.3.

The span of three nonzero vectors in E^3 can be a line, a plane, or all of E^3, depending on the degree of dependence of the three vectors. If all three are multiples of each other, we have only a line. If two of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are independent but the entire set is linearly dependent, then $\mathbf{v}_3$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ and $\mathbf{v}_3$ lies in the plane defined by $\mathbf{v}_1$ and $\mathbf{v}_2$. That is, the vectors are **coplanar**. Lay three pencils on a tabletop with erasers joined for a graphic example of coplanar vectors. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent, then the span is all E^3. This can be verified directly in individual cases; to show it in general requires methods of the next section.

Linear combinations in complex vector spaces have important applications, as the next examples illustrate.

Example 17. The set of Pauli spin matrices, used in the study of electron spin in quantum chemistry, is

$$S = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

Show that S is linearly independent in $\mathcal{C}_{22}$. Discuss the importance of the independence.

Solution We consider the equation

$$c_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \theta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

because the zero for $\mathcal{C}_{22}$ is

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

just as for $\mathcal{M}_{22}$. The matrix equation gives us

$$\begin{aligned} c_3 &= 0 \\ c_1 - ic_2 \qquad &= 0 \\ c_1 + ic_2 \qquad &= 0 \\ c_3 &= 0 \end{aligned}$$

which reduces to

$$\left(\begin{array}{ccc|c} 1 & -i & 0 & 0 \\ 1 & i & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

And

$$\det\begin{pmatrix} 1 & -i & 0 \\ 1 & i & 0 \\ 0 & 0 & 1 \end{pmatrix} = 2i \neq 0$$

so we have only the trivial solution, and S is a linearly independent set. The importance of the independence is that none of the matrices can be written in terms of the others; so the study of electron spin by Pauli matrices cannot, in general, be conducted with a proper subset of S.

Example 18. Let

$$\mathcal{S} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

Show that $\mathcal{S}$ spans all $\mathcal{C}_{22}$. Show that $\mathcal{S}$ is linearly independent. Note that $\mathcal{S}$ is the set of Pauli spin matrices with I_2 adjoined.

Solution Consider the equation

$$\begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} = c_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + c_4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which is equivalent to

$$\begin{aligned} z_{11} &= \qquad\qquad\quad c_3 + c_4 \\ z_{12} &= c_1 - ic_2 \\ z_{21} &= c_1 + ic_2 \\ z_{22} &= \qquad\qquad -c_3 + c_4 \end{aligned}$$

which has solution $c_1 = (z_{12} + z_{21})/2, c_2 = (z_{21} - z_{12})/(2i)$, $c_3 = (z_{11} - z_{22})/2$, and $c_4 = (z_{11} + z_{22})/2$. Therefore $\mathcal{S}$ spans $\mathcal{C}_{22}$. For the independence, note that if $z_{11} = z_{12} = z_{21} = z_{22} = 0$, we have only the trivial solution.

Problems 3.4

In Probs. 1 to 9, a set S of vectors in a vector space is given. (**a**) Describe span S (see Example 11); (**b**) determine whether S is linearly independent; and (**c**) if S is linearly dependent, write one vector as a linear combination of the others.

1. $S = \{(1,2,0), (0,3,2)\}$ in E^3

2. $S = \{1, x^2\}$ in $\mathcal{P}_2$

3. $S = \{(1,i,0), (0,1,i), (i, i-1, -1)\}$ in $\mathbb{C}^3$

4. $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$ in $\mathcal{M}_{22}$

5. $S = \{(1,1), (2,6), (7,12)\}$ in E^2

6. $S = \{(1,0,2), (3,2,-7), (-1,-2,11)\}$ in E^3

7. $S = \{(1,-1,2), (0,0,0)\}$ in E^3

8. $S = \{x+1, x^2-2, x-1, 3\}$ in $\mathcal{P}_2$

9. $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$ in $\mathcal{M}_{22}$

10. Which of the following sets of vectors span E^2?

    (**a**) $\{(1,-1,1), (2,0,3), (3,-1,4)\}$
    (**b**) $\{(0,0,0), (1,0,0), (1,1,0)\}$
    (**c**) $\{(1,0,0), (1,1,0), (1,1,1)\}$
    (**d**) $\{(1,-1,1), (1,0,1), (3,-1,2)\}$

11. Show that a set of two vectors from E^3 cannot span E^3.

12. Write any three distinct vectors from E^2. Show that they are linearly dependent.

13. Show that, in general, any three vectors from E^2 must be linearly dependent.

14. Show that any set S of vectors which contains θ is linearly dependent.

15. A polynomial is called **even** if its terms are constants and constants times **even** powers of x. Show that span $\{1, x^2\}$ is all the even polynomials in $\mathcal{P}_3$.

16. A polynomial is called **odd** if its terms are only constants times **odd** powers of x. Show that span $\{x, x^3\}$ is all the odd polynomials in $\mathcal{P}_3$.

17. The set
$$S = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$
is the set of real Pauli spin matrices used in the study of electron spin. Show that span S is the set of all 2×2 symmetric matrices with trace zero.

18. Let S be a linearly independent set of vectors from a vector space V. Show that any subset of S is linearly independent.

19. Let S be a linearly dependent set of vectors from a vector space V. Show that any set T with $S \subseteq T$ is a linearly dependent set.

20. Matrices A and B are said to **anticommute** if $AB = -BA$. Show that any pair of the Pauli spin matrices (see Example 17) anticommute.

21. The **commutator** of two square matrices of the same size is defined to be $AB - BA$ and is denoted $[A, B]$. Let
$$A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 5 \\ 4 & 1 \end{pmatrix}$$
Calculate $[A, I], [A, B], [B, A]$, and $[A, A]$.

22. For any square matrix A, what are $[A, I], [A, A], [A, O]$, and $[A, A^2]$? If A is invertible, what is $[A, A^{-1}]$?

23. Show that A and B commute for multiplication if and only if the commutator of A and B is 0.

24. Show, for $n \times n$ matrices A, B, and C and any scalar r, that

    (a) $[A, B + C] = [A, B] + [A, C]$

(**b**) $[rA, B] = r[A, B] = [A, rB]$

(**c**) $[A, B] = -[B, A]$

(**d**) $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$

(**e**) $[A, B]^T = [B^T, A^T]$

(**f**) $\mathrm{tr}[A, B] = 0$

3.5 BASES OF VECTOR SPACES; THE BASIS PROBLEM

The set of vectors $S = \{(1,1), (1,-1)\}$ spans E^2. That is, any vector in E^2 is a linear combination of $(1,1)$ and $(1,-1)$. The set of vectors $T = \{(1,1), (1,-1), (1,0)\}$ also spans E^2. Sets S and T differ in that S is linearly independent while T is linearly dependent. This makes a difference in writing a vector as a linear combination of vectors in the set. For example, writing (2,4) in terms of the vectors in S, we have for the only possibility

$$(2,4) = 3(1,1) - 1(1,-1)$$

However, in terms of vectors from T, we have several possibilities:

$$\begin{aligned}(2,4) &= 3(1,1) - 1(1,-1) + 0(1,0)\\(2,4) &= 0(1,1) - 4(1,-1) + 6(1,0)\\(2,4) &= 4(1,1) + 0(1,-1) - 2(1,0)\end{aligned}$$

Or, in general

$$(2,4) = (k+4)(1,1) + k(1,-1) + (-2-2k)(1,0)$$

The point is: If a set S of vectors spans V and S is linearly dependent, then representation of a vector $\mathbf{x}$ in terms of vectors in S is not unique. If we want uniqueness, the spanning set must also be linearly independent. Such a set is called a **basis** for V. Bases are used in coding theory, as we see later in this section.

Definition 3.5.1. A vector space V is said to be **finitely generated** if there exists a finite set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in V such that span $S = V$. If the set S is also linearly independent, then S is called a **basis** for V.

This definitions say the following

A finite set is a basis of V if it (1) spans V and (2) is linearly independent.

Example 1. $S = \{(1,2),(3,-1)\}$ is a basis of E^2.

Solution We must show that the set is linearly independent and spans E^2. That is, we must show that

$$c_1(1,2) + c_2(3,-1) = (a,b)$$

has a solution for any (a,b) and that

$$c_1(1,2) + c_2(3,-1) = (0,0)$$

has only the solution $c_1 = c_2 = 0$. These equations, respectively, in augmented form are

$$\left(\begin{array}{rr|c} 1 & 3 & a \\ 2 & -1 & b \end{array}\right) \quad \text{and} \quad \left(\begin{array}{rr|c} 1 & 3 & 0 \\ 2 & -1 & 0 \end{array}\right)$$

Instead of solving both sets of equations separately, we solve both at once by working with the doubly augmented matrix

$$\left(\begin{array}{rr|c|c} 1 & 3 & a & 0 \\ 2 & -1 & a & 0 \end{array}\right)$$

Doing this, we find that this matrix reduces to

Check linear independence with this part

$$\left(\begin{array}{cc|c|c} 1 & 3 & a & 0 \\ 0 & 1 & \dfrac{-b+2a}{7} & 0 \end{array}\right)$$

Check span with this part

We find for the span that $c_2 = (2a-b)/7, c_1 = (a+3b)/7$; for linear independence we find that $c_1 = c_2 = 0$. Since S is linearly independent and spans E^2, it is a basis for E^2.

In Example 1, the coefficients in the linear combination of basis elements were unique for any given vector (a,b). This is true in general.

Theorem 3.5.1. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V. Let $\mathbf{v}$ be in V. The coefficients in the representation*

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

are unique.

Proof. Suppose we have two representations

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$
$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$$

for $\mathbf{v}$; we will show that the coefficients are actually equal. To do this, form $\mathbf{v} + (-\mathbf{v})$, which equals θ, and combine terms to obtain

$$\theta = (a_1 - b_1)\mathbf{v}_1 + \cdots + (a_n - b_n)\mathbf{v}_n$$

Since S is a basis, it is a linearly independent set. Thus, the coefficients in the last linear combination must all be zero. That is, $a_1 = b_1, \ldots, a_n = b_n$, and the original linear combinations are the same. □

Example 2. $S = \{(i, 1+i), (2, 1-i)\}$ is a basis for $\mathbb{C}^2$.

Solution Proceeding as in Example 1, we form the doubly augmented matrix and row-reduce:

$$\left(\begin{array}{cc|c|c} i & 2 & a & 0 \\ 1+i & 1-i & b & 0 \end{array}\right) \xrightarrow{-i\text{R}1} \left(\begin{array}{cc|c|c} 1 & -2i & -ai & 0 \\ 1+i & 1-i & b & 0 \end{array}\right)$$
$$\xrightarrow{-(1+i)\text{R}1+\text{R}2} \left(\begin{array}{cc|c|c} 1 & -2i & -ai & 0 \\ 0 & -1+i & b+ai-a & 0 \end{array}\right)$$

This system has a unique solution. Therefore, S is linearly independent and spans $\mathbb{C}^2$; it is a basis for $\mathbb{C}^2$.

Example 3. Show that the set $S = \{(1, 2), (3, -1), (1, 0)\}$ is not a basis for E^2.

Solution The set S is linearly dependent because, for example,

$$(1, 2) + 2(3, -1) - 7(1, 0) = (0, 0)$$

So S cannot be a basis for E^2.

Example 4. The zero vector space has no basis, because any subset contains the zero vector and must be linearly dependent.

Example 4 shows that a vector space may fail to have a basis. We need to decide whether a given vector space has a basis or not. Spanning sets of vectors help us answer the question.

Theorem 3.5.2. *If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is a set of nonzero vectors which spans a subspace W of a vector space V, then some subset of S is a basis for W. (***Note***: This means V itself, being a trivial subspace, has a basis if it is spanned by S.)*

Proof. If S is a linearly independent set, then by definition S is a basis for W. If S is linearly dependent, then one of the vectors can be written as a linear combination of the others. Suppose $\mathbf{v}_m$ is such a vector (if not, shift the vectors in S around and relabel so that this is true). We claim that $S' = \{\mathbf{v}_1, \dots, \mathbf{v}_{m-1}\}$ still spans W. To see this, let $\mathbf{x}$ be in W with

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_{m-1}\mathbf{v}_{m-1} + c_m\mathbf{v}_m$$

Now $\mathbf{v}_m = d_1\mathbf{v}_1 + \cdots + d_{m-1}\mathbf{v}_{m-1}$, so we can substitute this expression into the former linear combination to obtain

$$\mathbf{x} = (c_1 + c_m d_1)\mathbf{v}_1 + \cdots + (c_{m-1} + c_m d_{m-1})\mathbf{v}_{m-1}$$

Thus S' spans W. If S' is linearly independent, S' is a basis for W. If S' is linearly dependent, one of the vectors in S' is a linear combination of the others. Now we argue as before. In this way we must arrive eventually at a linearly independent set which spans W. (If we reduce to a set with a single vector, that set is linearly independent because S was a set of nonzero vectors.) The resulting set is a basis of W. □

Thus we have the following fundamental result:

> Any finitely generated vector space, generated by a set of nonzero vectors, has a basis.

Example 5. Let V be the set of all polynomials with the usual operations. The vector space V is not finitely generated. In fact, if we take any finite subset S of V, then there will be a term of maximum degree, say x^p, in the set. The polynomial x^{p+1} is not in span S, and S cannot span V.

We are particularly interested in bases of finitely generated vector spaces. Example 3 illustrated the fact that any set of three or more vectors from E^2 cannot be a basis for E^2. After all, these vectors would be coplanar and form a linearly dependent set. A set of only one vector cannot be a basis for E^2 because it spans only a line through the origin. Thus it appears that any basis for E^2 must contain exactly two vectors. This follows from Theorem 3.5.3.

Theorem 3.5.3. *If $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for V, then (a) any set of $n+1$ (or more) vectors is linearly dependent and therefore is not a basis for V and (b) any set of $n-1$ (or less) vectors fails to span V and therefore is not a basis for V.*

This theorem means that the number of vectors in a basis is unique. If we find a basis $\mathcal{S}$ for V and $\mathcal{S}$ has eight vectors in it, then **every** *basis has eight vectors in it. Because of this we can define the* **dimension** *of a vector space V to be the number of vectors in a basis for V. If a basis $\mathcal{S}$ has n vectors in it, the dimension of V $(\dim V)$ is n, we write $\dim V = n$, and we say V is* **finite-dimensional**. *More particularly, V is called an* **n-dimensional vector space** *when a basis for V has n vectors in it. Example 1 shows that $\dim E^2 = 2$. The dimension of the zero vector space is defined to be zero.*

Proof. (**a**) Let $T = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n, \mathbf{w}_{n+1}\}$. That is, let T contain exactly $n+1$ vectors. We will show that T cannot be a basis by showing that T is linearly dependent. To do this, we consider

$$c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n + c_{n+1}\mathbf{w}_{n+1} = \theta \tag{3.5.1}$$

Now each $\mathbf{w}_k$ can be written as

$$\mathbf{w}_k = a_{1k}\mathbf{v}_1 + a_{2k}\mathbf{v}_2 + \cdots + a_{nk}\mathbf{v}_n \quad k = 1, 2, \ldots, n+1$$

since $\mathcal{S}$ spans V. Substituting this into the Eq. (3.5.1), we have

$$\begin{aligned}\sum_{k=1}^{n+1} c_k\mathbf{w}_k &= \sum_{k=1}^{n+1} c_k \sum_{j=1}^{n} a_{jk}\mathbf{v}_j \\ &= \sum_{j=1}^{n}\left(\sum_{k=1}^{n+1} a_{jk}c_k\right)\mathbf{v}_j = \theta\end{aligned}$$

Since $\mathcal{S}$ is linearly independent

$$\sum_{k=1}^{n+1} a_{jk}c_k = 0 \qquad \text{for all } j = 1, \ldots, n$$

That is,

$$\begin{array}{c} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1,n+1}c_{n+1} = 0 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2,n+1}c_{n+1} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{n,n+1}c_{n+1} = 0 \end{array}$$

This last homogeneous system has fewer equations than unknowns; so there exists a nontrivial solution for $c_1, c_2, \ldots, c_{n+1}$. This means that T is linearly dependent. Now suppose T contains **more than** $n+1$ vectors. Let $\widetilde{T}$ be a subset of $n+1$ vectors. It must be linearly dependent (as was just shown). Since $\widetilde{T} \subset T$, set T contains a linearly dependent subset and must itself be linearly dependent (see Prob. 19 of Sec. 3.4).

(**b**) Suppose T contains $n-1$ vectors and spans V. Then by Theorem 3.5.2 T must contain a basis $\mathcal{T}$ for V. If $\mathcal{T}$ contains r vectors, we must have $r \leq n-1$. Since $\mathcal{T}$ is a basis and $\mathcal{S}$ has $r+1$ or more vectors in it, it follows from part (**a**) that $\mathcal{S}$ is linearly dependent. This contradicts the fact that $\mathcal{S}$ is a basis.

We now know that finitely generated vector spaces which are generated by a set of nonzero vectors have bases and are finite-dimensional. Clearly, finite-dimensional vector spaces are finitely generated by a basis. Now we give several examples. $\square$

Example 6. (Standard basis) Show that E^n has the (standard) basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, where

$$\mathbf{e}_j = (0, 0, \ldots, 0, \underset{\substack{\nearrow \\ \mathbf{j}\text{th component}}}{1}, 0, \ldots, 0)$$

Solution For the span consider any $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in E^n$ and note that

$$\begin{aligned} \mathbf{x} &= x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \cdots + x_n(0, \ldots, 0, 1) \\ &= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \end{aligned}$$

For linear independence consider

$$c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n = \theta = (0, 0, \ldots, 0)$$

so that $c_1 = c_2 = \cdots = c_n = 0$. Therefore, $\mathcal{E}$ is a basis for E^n, and $\dim E^n = n$. This is reasonable since we associate (see Fig. 3.5.1)

E^1	Line	One-dimensional object
E^2	Plane	Two-dimensional object
E^3	Space	Three-dimensional object

Comparing Examples 1 and 4, we see that E^2 has more than one possible basis. In general, a vector space (nonzero) has an infinite number of bases. However, the number of elements in any basis is always the same: Remember, this number is the dimension of the space

Example 7. Show that $\mathcal{M}_{23}$ has dimension 6.

Solution A basis is (in fact, this is the standard basis)

$$S = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

so $\dim \mathcal{M}_{23} =$ number of vectors in $S = 6$.

Example 8. The vector space $\mathcal{P}_n$ has dimension $n + 1$.

Solution A basis is

$$S = \{1, x, \ldots, x^n\}$$

To see this, we check first for linear independence. The equation

$$a_1 1 + a_2 x + a_3 x^2 + \cdots + a_{n+1} x^{n+1} = 0$$

holds only if the polynomial on the left is zero for all real x. From algebra this occurs only if all the coefficients are zero, that is, only if $a_1 = a_2 = \cdots = a_{n+1} = 0$. Therefore, S is linearly independent. That S spans $\mathcal{P}_n$ follows from the fact that any polynomial in $\mathcal{P}_n$ is of the form

$$a_0 + a_1 x + \cdots + a_n x^n$$

Example 9.

$$\mathcal{S} = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

is a basis for $\mathcal{C}_{22}$. Example 18 of Sec. 3.4 showed that $\mathcal{S}$ is linearly independent and spans $\mathcal{C}_{22}$.

Now that we know what a basis of a vector space is, we can state **the second fundamental problem of linear algebra**.

> **Basis Problem**
> Let V be a vector space. The basis problem may take one of the following forms.
>
> > **Problem 1.** Construct a basis for V, by choosing vectors from V.
> > **Problem 2.** Given a set S of vectors in V, construct a basis for V by enlarging S, or deleting some (but not all) vectors from S, or both.

Before we try to solve this problem, we might ask, Is a solution even possible? Theorem 3.5.2, which tells us to "throw" out dependent vectors from a spanning set" to get a basis, helps us here.

> **Problem 1.** If we can pick a set of vectors from V which spans V, then by throwing out dependent vectors we will arrive at a basis for V.
> **Problem 2.** If the given set S spans V, we proceed as in problem 1. If not, we enlarge S by putting in more vectors until a spanning set is achieved. Then we proceed as in problem 1.

Example 10. (Problem 2 of the basis problem) Let $S = \{(1,0,3),(2,1,4)\}$. Find a basis T for E^3 containing S.

Solution 1 Since E^3 has dimension 3, we know that T must contain exactly three vectors. Set S is already linearly independent, so we have to add only one more vector to S. However, we must be careful. The new vector we join to the set S must not make the set T linearly dependent. So the new vector must not be in the span of the vectors already in S.

Now span $\{(1,0,3),(2,1,4)\} = \{\mathbf{x}|\mathbf{x} = a(1,0,3) + b(2,1,4)\} = \{\mathbf{x}|\mathbf{x} = (a+2b, b, 3a+4b)\}$. We must make sure that the new vector is not of the form $(a+2b, b, 3a+4b)$. To do this, we suppose our new vector is (x_1, x_2, x_3), and we force the equation

$$(a+2b, b, 3a+4b) = (x_1, x_2, x_3)$$

to have no solution for a and b. Matching components yields the equations

$$\left(\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 3 & 4 & x_3 \end{array}\right)$$

which reduces to

$$\left(\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 0 & 0 & x_3 - 3x_1 + 2x_2 \end{array}\right)$$

So if we choose x_1, x_2, x_3 with $x_3 - 3x_1 + 2x_2 \neq 0$, we have a vector not in span $\{(1,0,3),(2,1,4)\}$. Therefore, $\mathbf{x} = (0,1,0)$ works and

$$T = \{(1,0,3),(2,1,4),(0,1,0)\}$$

is a basis for E^3.

Solution 2 If the third vector (x_1, x_2, x_3) were to make $\{(1,0,3),(2,1,4),(x_1,x_2,x_3)\}$ linearly dependent, then

$$\det\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{pmatrix} = 0$$

since one of the rows would be a linear combination of the others. So we require

$$\det\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{pmatrix} \neq 0$$

Calculating the determinant, we have

$$x_3 + 2x_2 - 3x_1 \neq 0$$

which is the same condition obtained in solution 1. Possible third vectors are (1,0,0), (0,1,0), (0,0,1), (0,1,1), or (1,1,0); actually there are an infinite number of choices.

Solution 3 (Trial and error) We just try standard basis vectors until one works. Try first (1,0,0) and check for linear dependence:

$$c_1(1,0,3) + c_2(2,1,4) + c_2(1,0,0) = (0,0,0)$$

$$\Rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 4 & 0 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Therefore, the vectors are linearly independent and form a basis
$T = \{(1,0,3),(2,1,4),(1,0,0)\}$.

In Example 10 we have implicitly used the following theorem.

Theorem 3.5.4. *Let* $\dim V = n$, *and let* $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *be a subset of* V. *The following are equivalent:*

1. *Set* S *is a basis for* V,
2. *Set* S *is linearly independent,*
3. *Set* S *spans* V.

Proof. $(1 \Rightarrow 2)$ This follows from the definition of basis.
$(2 \Rightarrow 3)$ Suppose S is linearly independent and S does **not** span V. Then there is a vector $\mathbf{v}_{n+1} \in V$ which is not in span S. That is,

$$T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}\}$$

is a linearly independent set from V. But then $\dim V \geq n+1$, which contradicts the hypothesis that $\dim V = n$.
$(3 \Rightarrow 1)$ Suppose S spans V and is not a basis for V. Then it must be linearly dependent. By Theorem 3.5.2 there is a subset of S which is a basis of V. However, this subset must have less than n vectors in it, which implies that $\dim V < n$, a contradiction. □

Note that had Theorem 3.5.4 been available prior to Examples 1 and 2, the solutions would have required half the work. Showing linear independence would have sufficed.

Example 11. (Problem 1 of the basis problem) Find a basis for the solution space of

$$\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 0 \\ x_2 + x_3 - x_4 &= 0 \\ 3x_1 + 4x_2 - 2x_3 + 5x_4 &= 0 \end{aligned}$$

Solution The equation in augmented matrix form are

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 3 & 4 & -2 & 5 & 0 \end{array}\right) \xrightarrow{\text{Row reduction}} \left(\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

and we can let $x_4 = k, x_3 = j$ to find the solutions

$$\begin{aligned}(x_1, x_2, x_3, x_4) &= (2j - 3k, k - j, k, j) \\ &= j(2, -1, 0, 1) + k(-3, 1, 1, 0)\end{aligned}$$

Since $S = \{(2, -1, 0, 1), (-3, 1, 1, 0)\}$ spans the solution space and is linearly independent, S is a basis for the solution space. Therefore the dimension of the solution space is 2.

Examples 10 and 11 show that solution of the basis problem is usually not found by formula or rote methods. It requires thought, versatility, and the ability to use almost all the preceding material.

Example 12. The set

$$\text{span } S = \text{span}\{(1, -1, 2), (0, 5, -8), (3, 2, -2), (8, 2, 0)\}$$

is a vector space. Find a basis for it.

Solution We delete vectors which are linear combinations of the others. To see the dependencies, we consider

$$c_1(1, -1, 2) + c_2(0, 5, -8) + c_3(3, 2, -2) + c_4(8, 2, 0) = (0, 0, 0)$$

which is equivalent to

$$\left(\begin{array}{rrrr|r} 1 & 0 & 3 & 8 & 0 \\ -1 & 5 & 2 & 2 & 0 \\ 2 & -8 & -2 & 0 & 0 \end{array}\right) \xrightarrow{\text{Row reduction}} \left(\begin{array}{rrrr|r} 1 & 0 & 3 & 8 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{3.5.2}$$

Solutions are $c_4 = k, c_3 = j, c_2 = -j - 2k, c_1 = -3j - 8k$. Choosing $k = 1$, $j = 0$, we have

$$8(1, -1, 2), +2(0, 5, -8) = (8, 2, 0)$$

Choosing $k = 0$ and $j = 1$, we have

$$3(1, -1, 2) + (0, 5, -8) = (3, 2, -2)$$

Thus (8,2,0) and $(3, 2, -2)$ depend on $(1, -1, 2)$ and $(0, 5, -8)$, and

$$T = \{(1, -1, 2), (0, 5, -8)\}$$

is a basis for span S. The dimension of span S is 2.

In Example 12, the augmented matrix on the left-hand side of Eq. (3.5.2) has columns consisting of the vectors from S. Also the number of rows (two) in the row-reduced form is equal to the dimension of span S.

In general, for problems in E^n of the type in Example 12, we have a helpful theorem. Before stating it, we note that for an $m \times n$ matrix A, if we consider the rows as vectors from E^n then the span of those vectors is called the **row space of** A.

Theorem 3.5.5. *If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a set of vectors from E^n and A is the matrix formed by putting $\mathbf{v}_1$ in row $1, \mathbf{v}_2$ in row 2, and so on, and if B is the reduced row echelon form of A, then the nonzero rows of B form a basis for the row space of A. That is, the nonzero rows of B form a basis for span S.*

Proof. Let the matrix be

$$A_{m\times n} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{pmatrix}$$

By the definition of row operations, if a row of zeros is obtained, that row was equal to a linear combination of other vectors in the set. The remaining rows are therefore all linear combinations of the independent vectors from the original set. Thus the span of the nonzero rows is equal to span S. Thus the nonzero rows, being independent, form a basis for span S and dim(span S) = number of nonzero rows. $\square$

Example 13. Work Example 12 by using Theorem 3.5.5.

Solution Form A and row-reduce.

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -8 \\ 3 & 2 & -2 \\ 8 & 2 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 2 \\ 3 & 2 & -2 \\ 8 & 2 & 0 \\ 0 & 5 & -8 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -8 \\ 0 & 10 & -16 \\ 0 & 5 & -8 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -\frac{8}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

By Theorem 3.5.5, $T = \{(1, -1, 2), (0, 1, -\frac{8}{5})\}$ is a basis for span S, and dim(span
$S) = 2$.

Theorem 3.5.5 can be stated in terms of the **rank** of a matrix A.

Definition 3.5.2. Let A be an $m \times n$ matrix. The **row rank** of a matrix is the number of nonzero rows in the reduced row echelon form of A. The **column rank** of a matrix is the number of nonzero rows in the reduced row echelon form of A^T. The row and column ranks of the zero matrix are defined to be zero.

Example 14. Calculate the row and column ranks of

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

Solution In Example 13 we found the row rank to be 2. For the column rank we can do column operations or form A^T, do row operations, and transpose. We will use column operations.

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -8 \\ 3 & 2 & -2 \\ 8 & 2 & 0 \end{pmatrix} \xrightarrow[-2\text{C1}+\text{C3}]{\text{C1}+\text{C2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & -8 \\ 3 & 5 & -8 \\ 8 & 10 & -16 \end{pmatrix} \xrightarrow{\frac{8}{5}\text{C2}+\text{C3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 3 & 5 & 0 \\ 8 & 10 & 0 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{5}\text{C2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 0 \\ 8 & 2 & 0 \end{pmatrix}$$

The column rank of A is 2 also.

The column rank and row rank of A in Example 14 were equal. This is always true.

Theorem 3.5.6. *For any matrix A*

$$\textbf{row rank } A = \textbf{column rank } A$$

Proof. Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ be the rows of $A_{m\times n}$. After reduction a basis $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$, $k \le m$, is found for span $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$. Thus

$$\begin{array}{c} \mathbf{v}_1 = a_{11}\mathbf{w}_1 + a_{12}\mathbf{w}_2 + \cdots + a_{1k}\mathbf{w}_k \\ \mathbf{v}_2 = a_{21}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{2k}\mathbf{w}_k \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{v}_2 = a_{m1}\mathbf{w}_1 + a_{m2}\mathbf{w}_2 + \cdots + a_{mk}\mathbf{w}_k \end{array} \tag{3.5.3}$$

Now writing

$$\mathbf{v}_j = (b_{j1}, b_{j2}, \dots, b_{jn}) \qquad 1 \le j \le n$$

and

$$\mathbf{w}_i = (c_{i1}, c_{i2}, \dots, c_{in}) \qquad 1 \le i \le n$$

we find after substitution into Eq. (3.5.3) that

$$\underbrace{\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{pmatrix}}_{\substack{\text{This is } j\text{th} \\ \text{column of } A}} = \underbrace{c_{1j}\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + c_{kj}\begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}}_{k \text{ of these}}$$

This means that the transpose of each column is a linear combination of k vectors; therefore the column rank of A is less than or equal to k. That is,

$$\text{column rank } A \le \text{row rank } A$$

In the same way, we find now rank $A \le$ column rank A. Therefore the ranks are equal. $\square$

As a result of this theorem, we define the **rank of a matrix** A as just the row rank of A.

Now results regarding linear equations and matrices can be concisely stated.

1. $A_{m\times n}X_{n\times 1} = B_{m\times 1}$ has a solution if and only if rank $A =$ rank $(A|B)$

2. Matrix $A_{n\times n}$ is invertible if and only if rank $A = n$. Also $\det A_{n\times n} \neq 0$ if and only if rank $A = n$.
3. $A_{m\times n}X_{n\times 1} = 0_{n\times 1}$ has a nontrivial solution if and only if rank $A < n$.

Now let us return to coding theory to see another way that bases are used.

Example 15. As defined in Sec. 3.3, a binary channel linear code V is a subspace of F^n. Code V, being finitely generated by nonzero vectors, has a basis. When the elements of a basis of V are arranged in rows of a matrix, the matrix is called the **generator matrix** of V. The span of the rows, therefore, is the entire code; the matrix furnishes a compact storage mechanism for a code. For example, the matrix

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

generates the code in this table:

VECTOR IN CODE	GENERATION
(0,0,0)	$0(1,0,1) + 0(0,1,1)$
(0,1,1)	$0(1,0,1) + 1(0,1,1)$
(1,0,1)	$1(1,0,1) + 0(0,1,1)$
(1,1,0)	$1(1,0,1) + 1(0,1,1)$

A 30×20 matrix (30 vectors of 20 elements each) generates a code of more than 10^9 code vectors. After all, each element of the code is a sum of 30 basis vectors, and there are 2 choices for each of the 30 coefficients. Thus there are $2^{30} = 1,073,741,824 > 10^9$ code vectors in V.

Example 15 raises an important question for decoding: Is each code vector a unique linear combination of the rows of the generator matrix? The answer is yes and was proved in Theorem 3.5.1.

Another question generated by the coding example is: If a code C in F^n is "smaller" than another code D in F^n, in the sense that C is a proper subspace of D, then is a basis for C "smaller" than a basis for D? This seems reasonable, and the answer is yes, as shown in Theorem 3.5.7.

Theorem 3.5.7. *If V is finite-dimensional vector space and W is a subspace of V, then* $\dim W \leq \dim V$.

Proof. If $W = \{\theta\}$, then the result is true. Suppose $W \neq \{\theta\}$. We first show that W is finitely generated. Since $W \neq \{\theta\}$, there is a nonzero vector in $\mathbf{v}_1$ in W. Now either span $\{\mathbf{v}_1\} = W$ or not. If so, W is finitely generated by $\{\mathbf{v}_1\}$. If not, there is a vector $\mathbf{v}_2$ not in span $\{\mathbf{v}_1\}$ which is still in W. The set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent; otherwise, $\mathbf{v}_2$ would be in span $\{\mathbf{v}_1\}$. Now if span $\{\mathbf{v}_1, \mathbf{v}_2\} = W$, then we are done; if not, we continue to add vectors and check linear independence. This process must stop at some point because if V has dimension n, then any set of $n+1$ vectors from V (and therefore from W) must be linearly dependent. Say that the process stops after k steps. Then we have $W = \text{span}\,\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ and $k \leq n$. □

Problems 3.5

1. Which of the following sets are bases for E^2?

   (**a**) $\{(1,-3),(-17,54)\}$ (**b**) $\{(1,0),(0,0)\}$ (**c**) $\{(1,2)\}$
   (**d**) $\{(1,-1),(2,-2)\}$ (**e**) $\{(1,0),(0,1),(1,-1)\}$

2. Which of the following sets are bases for E^3?

   (**a**) $S = \{(1,-1,2),(2,3,5),(-3,0,2)\}$

   (**b**) $S = \{(1,2,4),(-6,2,3),(-4,6,11)\}$

   (**c**) $S = \{(1,-1,2),(0,3,6)\}$

   (**d**) $S = \{(1,0,1),(0,1,1),(1,1,0),(1,0,0)\}$

In Probs. 3 and 7, a set S and a vector space V are given. Find a basis for V containing S.

3. $V = E^3, S = \{(1,-1,1),(0,1,-1)\}$

4. $V = E^4, S = \{(1,0,2,2),(1,1,0,0)\}$

5. $V = \mathcal{M}_{22}, S = \left\{\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$

6. $V = \mathcal{P}_2, S = \{1 - x + x^2, x - x^2\}$

7. $V = \mathbb{C}^3, S = \{(i,1,0),(0,i,1)\}$

In Probs. 8 to 12, a set S and a vector space V are given. Find a basis for V by deleting vectors from S.

8. $V = E^2, S = \{(1,-1),(1,2),(3,4)\}$
9. $V = \text{span}\{(1,2,0),(1,5,6),(2,13,18)\} = \text{span } S$
10. $V = E^3, S = \{(1,-1,2),(4,-3,7),(2,0,5),(1,2,6)\}$
11. $V = \mathcal{P}_2, S = \{1+x, x+x^2, 1+x^2, x-x^2\}$
12. $V = \mathbb{C}^3, S = \{(1,0,0),(i,1,0),(1,i,0),(0,0,1)\}$

In Probs. 13 to 18, a vector space V is given. Find a basis and state $\dim V$.

13. $V = E^5$
14. $V = \mathcal{P}_4$
15. $V = \mathcal{M}_{24}$
16. $V = \text{span}\{(1,-1,2),(5,5,5),(0,6,3)\}$
17. $V = \mathcal{C}_{32}$
18. $V =$ solution space of

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

19. Show that $\dim \mathcal{M}_{nn} = n^2$. What is the dimension of the subspace of symmetric matrices?
20. Show that $\dim \mathcal{P}_n = n+1$. What is the dimension of the subspace of even polynomials (polynomials with even exponents)?
21. What is $\dim \mathcal{M}_{mn}$? What is $\dim \mathcal{C}_{mn}$?
22. Show that
    $\dim \mathcal{P}_n = \dim$ (subspace of even polynomials) $+ \dim$ (subspace of odd polynomials)

23. Calculate the rank of the given matrices.

(**a**) $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ (**b**) $A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 5 & 6 & 2 \\ -1 & 2 & 4 & 3 \\ 1 & 2 & -1 & 2 \end{pmatrix}$

(**c**) $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$

24. Show that the standard basis for E^n is also a basis for $\mathbb{C}^n$.

25. Show that $\dim \mathbb{C}^n = n$.

26. Show that

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

generates the linear code

$$\begin{aligned} V =\{&(0,0,0,0),(1,1,0,1),(1,0,0,1),(1,0,1,0),(0,1,0,0),\\ &(0,1,1,1),(0,0,1,1),(1,1,1,0)\} \end{aligned}$$

Show that

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

generates a subcode W of V. What are $\dim W$ and $\dim V$?

27. Compare rank A and rank $(A|B)$ to determine whether the following systems have solutions

(a)
$$\begin{aligned} x - y + z &= 2 \\ 2x + y - z &= 3 \\ x + 2y + 4z &= 7 \end{aligned}$$

(b)
$$\begin{aligned} x + y + 2z &= 3 \\ -x - 3y + 4z &= 2 \\ -x - 5y + 10z &= 7 \end{aligned}$$

(c)
$$\begin{aligned} x + y + 2z &= 3 \\ -x - 3y + 4z &= 2 \\ -x - 5y + 10z &= 11 \end{aligned}$$

28. Find bases for the row spaces of the matrices in Prob. 23.

3.6 PERPENDICULARITY IN VECTOR SPACES

E^3 has a standard basis of $\mathcal{E} = \{(1,0,0),(0,1,0),(0,0,1)\}$ which can be pictured as in Fig. 3.5.1. These vectors have length 1 and are perpendicular to each other. Such a basis is called **orthonormal: ortho** for orthogonal (perpendicular) and **normal** for normalized (length 1). To discuss orthonormal bases for other vector spaces, we must extend the idea of dot product to what is called an **inner product**.

Definition 3.6.1. Let V be a real or complex vector space. An **inner product** on V is a function which associates with each pair $\mathbf{x}$, $\mathbf{y}$ of vectors a number $\langle \mathbf{x}, \mathbf{y} \rangle$ and satisfies the following properties. (On the right-hand side we list the situation in E^2. The overbar denotes complex conjugate.)

Let r be a scalar, $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$	In $E^2, \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y}$ $\mathbf{x} = (x_1, x_2)$ $\mathbf{y} = (y_1, y_2)$
1. $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$	1. $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 = y_1 x_1 + y_2 x_2 = \mathbf{y} \cdot \mathbf{x} = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$
2. $\langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle$	2. $\langle \mathbf{x} + \mathbf{z}, \mathbf{y} \rangle = (\mathbf{x} + \mathbf{z}) \cdot \mathbf{y} = (x_1 + z_1) y_1 + (x_2 + z_2) y_2 = x_1 y_1 + x_2 y_2 + z_1 y_1 + z_2 y_2 = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle$
3. $\langle r\mathbf{x}, \mathbf{y} \rangle = r \langle \mathbf{x}, \mathbf{y} \rangle$	3. $\langle r\mathbf{x}, \mathbf{y} \rangle = (r\mathbf{x}) \cdot \mathbf{y} = (rx_1, rx_2) \cdot (y_1, y_2) = rx_1 y_1 + rx_2 y_2 = r(x_1 y_1 + x_2 y_2) = r \langle \mathbf{x}, \mathbf{y} \rangle$
4. $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ unless $\mathbf{x} = \theta$	4. $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x} \cdot \mathbf{x} = {x_1}^2 + {x_2}^2$ which is a sum of nonnegative terms. The sum can be zero if and only if $x_1 = 0$ and $x_2 = 0$, that is, when $\mathbf{x} = 0$.

Example 1. In E^n, if $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, then an **inner product** can be defined as

$$\langle \mathbf{x}, \mathbf{y} \rangle \equiv x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

This inner product reduces to the ordinary dot product in E^2 and E^3. It is called the **standard inner product for** E^n.

Example 2. In $\mathbb{C}^n$, if $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, then an **inner product** can be defined as

$$\langle \mathbf{x}, \mathbf{y} \rangle \equiv x_1\bar{y} + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n$$

This is the **standard inner product for** $\mathbb{C}^n$. Note that $\langle c\mathbf{x}, \mathbf{y} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle$, but $\langle \mathbf{x}, c\mathbf{y} \rangle = x_1\overline{cy}_1 + x_2\overline{cy_2} + \cdots + x_n\overline{cy}_n = \bar{c}\langle \mathbf{x}, \mathbf{y} \rangle$.

Example 3. In $\mathcal{M}_{23}$ if

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

then we can define an inner product as

$$\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{13}b_{13} + a_{21}b_{21} + a_{22}b_{22} + a_{23}b_{23}$$

Example 4. In E^2, let $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2)$, and define $\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1y_1 + 2x_2y_2\langle \mathbf{x}, \mathbf{y} \rangle$ is an inner product.

Example 4 shows that there may be more than one inner product defined on a vector space; E^2 also has the standard inner product.

In E^2 and E^3, we have

$$|x| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

so that it is natural to define the length of a vector $\mathbf{x}$ in any vector space with an inner product as $\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. Vector spaces with an inner product are also called **inner product spaces**.

Definition 3.6.2. Let V be a vector space with inner product $\langle \cdot, \cdot \rangle$. If **x,y** are in V,

1. The **norm** (length) of $\mathbf{x}$ is denoted by $\|\mathbf{x}\|$ and is defined as

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

2. The **distance** between $\mathbf{x}$ and $\mathbf{y}$ is defined as $\|\mathbf{x} - \mathbf{y}\|$.

3. Nonzero $\mathbf{x}$ and $\mathbf{y}$ are defined to be **orthogonal** if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Note that $\|x\|$ may be different for different inner products, as the next example shows.

Example 5. In E^2 let $\mathbf{x} = (1,1)$ and $\mathbf{y} = (2,3)$. Calculate $\|\mathbf{x}\|$ and $\|\mathbf{x}-\mathbf{y}\|$ for norms generated by the standard inner product and the inner product in Example 4.

Solution

Standard inner product	Inner product from Example 4
$\langle (1,1),(2,3)\rangle = 5$	$\langle (1,1),(2,3)\rangle = 10$
$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x},\mathbf{x}\rangle} = \sqrt{2}$	$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x},\mathbf{x}\rangle} = \sqrt{2+2} = 2$
$\|\mathbf{x}-\mathbf{y}\| = \|(-1,-2)\| = \sqrt{5}$	$\|\mathbf{x}-\mathbf{y}\| = \|(-1,-2)\| = \sqrt{10}$

Several properties of inner products are generalizations of properties of the dot product in E^2 and E^3. Some of these are stated in Theorem 3.6.1.

Theorem 3.6.1. *Let V be a vector space with inner product $\langle \cdot,\cdot\rangle$. The following properties hold for all $\mathbf{x}$ and $\mathbf{y}$ in V.*

*(**a**)* $\langle \mathbf{x}, \theta\rangle = 0$

*(**b**)* $\|\mathbf{x}\| \geq 0, \|\mathbf{x}\| = 0$ *if and only if* $\mathbf{x} = \theta$

*(**c**)* $\langle \mathbf{x}, \mathbf{y}+\mathbf{z}\rangle = \langle \mathbf{x},\mathbf{y}\rangle + \langle \mathbf{x},\mathbf{z}\rangle$ *and* $\langle \mathbf{x}, r\mathbf{y}\rangle = \bar{r}\langle \mathbf{x},\mathbf{y}\rangle$

*(**d**)* $\|k\mathbf{x}\| = |k|\ \|\mathbf{x}\|$

*(**e**)* $|\langle \mathbf{x},\mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ *(Cauchy-Schwarz inequality)*

*(**f**)* $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *(triangle inequality)*

Proof. (**a**) We have

$$\langle \mathbf{x},\theta\rangle = \langle \mathbf{x},\theta+\theta\rangle = \langle \mathbf{x},\theta\rangle + \langle \mathbf{x},\theta\rangle$$

The only solution to the equation $q = q + q$ is $q = 0$. Thus $\langle \mathbf{x},\theta\rangle = 0$.
(**b**), (**c**) See the problems .
(**d**) $\|k\mathbf{x}\| = \overline{\langle k\mathbf{x}, k\mathbf{x}\rangle} = \sqrt{k\bar{k}\langle \mathbf{x},\mathbf{x}\rangle} = \sqrt{|k|^2}\sqrt{\langle \mathbf{x},\mathbf{x}\rangle} = \sqrt{|k|^2}\|\mathbf{x}\|$.
Since the radical signifies the positive square root, $\sqrt{k^2} = |k|$.
(**e**) See the problems.

$$
\begin{aligned}
(f) \quad \|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{x}+\mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}+\mathbf{y}\rangle \\
&= \langle \mathbf{x}, \mathbf{x}\rangle + \langle x, y\rangle + \overline{\langle \mathbf{x}, \mathbf{y}\rangle} + \langle \mathbf{y}, \mathbf{y}\rangle \\
&\le \langle \mathbf{x}, \mathbf{x}\rangle + 2\|\mathbf{x}\|\|\mathbf{y}\| + \langle \mathbf{y}, \mathbf{y}\rangle \qquad \text{by } (\mathbf{e}) \\
&= \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 \\
&= (\|\mathbf{x}\| + \|\mathbf{y}\|^2)
\end{aligned}
$$

Taking the square root of both sides gives

$$\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$$

□

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in a real vector space, we can define the angle between $\mathbf{x}$ and $\mathbf{y}$. Notice that the Cauchy-Schwarz inequality allows us to define

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}$$

because $|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|\|\mathbf{y}\|$ implies that

$$-1 \le \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \le 1$$

Now that orthogonality and norm have been defined, we can discuss bases of

orthonormal vectors. A set $\mathcal{O} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of vectors is an **orthonormal set** in a vector space V with inner product $\langle\cdot,\cdot\rangle$ if

$$\|\mathbf{v}_k\| = 1 \qquad k = 1, 2, \ldots, n$$

and

$$\langle \mathbf{v}_k, \mathbf{v}_j\rangle = 0 \qquad k \ne j$$

That is,

$$\langle \mathbf{v}_k, \mathbf{v}_j\rangle = \delta_{kj}$$

One of the most important results of linear algebra is that **orthonormal sets are always linearly independent**. Consider

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + \cdots + c_n\mathbf{v}_n = \theta$$

Even though we do not know what V is or what the $\mathbf{v}_k$'s look like, we can solve for the coefficients. We use the important technique of taking the inner product of both sides with a particular vector, in this case $\mathbf{v}_k$. We have

$$c_1\langle \mathbf{v}_1, \mathbf{v}_k\rangle + c_2\langle \mathbf{v}_2, \mathbf{v}_k\rangle + \cdots + c_k\langle \mathbf{v}_k, \mathbf{v}_k\rangle + \cdots + c_n\langle \mathbf{v}_k, \mathbf{v}_n\rangle = \langle \theta, \mathbf{v}_k\rangle = 0 \tag{3.6.1}$$

and since $\langle \mathbf{v}_k, \mathbf{v}_j\rangle = 0$, for $j \neq k$, Eq. (3.6.1) becomes

$$c_k = 0$$

Since this works for any $k = 1, 2, \ldots, n$, $c_1 = c_2 = \cdots = c_n = 0$, and the set $\mathcal{O}$ is linearly independent.

The fact that orthonormal sets are always linearly independent makes them nice to work with when we are finding bases for vector spaces with inner products. All we have to worry about is spanning V. Thus in E^n the standard basis is an **orthonormal basis**. If a basis is made up of orthogonal vectors but not all are of norm 1, then the basis is an **orthogonal** basis.

Example 6. Let $V = \mathcal{P}_2, p(x) = a_0 + a_1x + a_2x^2$, and $q(x) = b_0 + b_1x + b_2x^2$. Define $\langle p, q\rangle = a_0b_0 + a_1b_1 + a_2b_2$. Show that $\langle \cdot, \cdot \rangle$ is an inner product. Give an example of two orthogonal vectors from $\mathcal{P}_2$. Find an orthonormal basis for $\mathcal{P}_2$.

Solution Verifying properties 1 to 4 of the definition of inner product (Definition 3.6.1) is straightforward. The vectors

$$p(x) = 1 + x^2 \quad \text{and} \quad q(x) = x$$

are orthogonal. The vector $r(x) = 1 - x^2$ is orthogonal to both p and q. The set

$$S = \{1 + x^2, x, 1 - x^2\}$$

is orthogonal and linearly independent. Since $\dim \mathcal{P}_2 = 3$, set S must be a basis. However, S is only an orthogonal basis. To make it an orthonormal basis, the vectors must normalized. We put

$$\begin{aligned}
u_1 &= \frac{p}{\|p\|} = \frac{1 + x^2}{\sqrt{\langle 1 + x^2, 1 + x^2\rangle}} = \frac{1 + x^2}{\sqrt{2}} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}x^2 \\
u_2 &= q = x \qquad (q \text{ is already of norm } 1) \\
u_3 &= \frac{r}{\|r\|} = \frac{1 - x^2}{\sqrt{\langle 1 - x^2, 1 - x^2\rangle}} = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}x^2
\end{aligned}$$

Now

$$\mathcal{O} = \{u_1, u_2, u_3\} = \left\{ \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}x^2, x, \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}x^2 \right\}$$

is an orthonormal basis for $\mathcal{P}_2$. $\mathcal{O}$ is not the only orthonormal basis for $\mathcal{P}_2$ with this inner product; the standard basis $\mathcal{S} = \{1, x, x^2\}$ is also an orthonormal basis.

Problems 3.6

1. Which of the following are inner products on E^2? Let $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$

   (**a**) $\langle \mathbf{x}, \mathbf{y} \rangle = 4x_1y_1 + x_2y_2$ (**b**) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 + y_1$
   (**c**) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 - x_2y_2$ (**d**) $\langle \mathbf{x}, \mathbf{y} \rangle = {x_1}^2y_1 + x_2y_2$
   (**e**) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_2 + x_2y_1$ (**f**) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1x_2y_1y_2$
   (**g**) $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 + x_2 + y_1 + y_2$ (**h**) $\langle \mathbf{x}, \mathbf{y} \rangle = |x_1y_1 + x_2y_2|$

2. Using the inner product from Prob. 1(**a**) on E^2, show that $\mathcal{O} = \{(\frac{1}{2}, 0), (0, 1)\}$ is an orthonormal basis for E^2 with that inner product.

3. In E^2 let $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$. Show that the function

   $$\langle \mathbf{x}, \mathbf{y} \rangle = ax_1y_1 + bx_2y_2 \qquad \text{where } a > 0, b > 0$$

   is an inner product. What happens if either $a < 0$ or $b < 0$?

4. **Difficulty norm**. Consider a way of "norming" a hike up a mountain. Difficulty indices are as follows.

CATEGORY	ANGLE OF CLIMB	DIFFICULTY INDEX
1	$0° - 10°$	1
2	$10° - 20°$	2
3	$25° - 45°$	4

   A difficulty norm for a trip of x_1 mi in category $1, x_2$ mi in category 2, and x_3 mi in category 3 is

   $$\|(x_1, x_2, x_3)\| = \sqrt{{x_1}^2 + 2{x_2}^2 + 4{x_3}^2}$$

which is generated by the inner product

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3) \rangle = x_1 y_1 + 2x_2 y_2 = 4x_3 y_3$$

Which of the following hikes is most difficult (has the largest difficulty norm)?

	MILES IN		
TRIP	**CATEGORY 1**	**CATEGORY 2**	**CATEGORY 3**
T_1	3	2	1
T_2	0	6	0
T_3	5	0	1

Note that all hikes are 6 mi long.

5. Let V be a vector space with inner product $\langle \mathbf{x}, \mathbf{y} \rangle$. Let a new function $\langle\ ,\ \rangle_k$ be defined on V by $\langle \mathbf{x}, \mathbf{y} \rangle_k = k \langle \mathbf{x}, \mathbf{y} \rangle$, where $k > 0$. Show that $\langle\ ,\ \rangle_k$ is also an inner product. Define $\|\mathbf{x}\|_k = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_k}$. Compare $\|\mathbf{x}\|_k$ with $\|\mathbf{x}\|_1$ (the original norm) for $k > 1$ and $k < 1$.

6. Let
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$
in $\mathcal{M}_{22}$ with standard inner product. Calculate $\langle A, B \rangle$, $\|A\|$, $\|B\|$, $\|A - B\|$, and $\cos\theta$, where θ is the angle between A and B.

7. Consider E^3 with the standard inner product.
   (**a**) Find a vector orthogonal to both $(3, -1, 0)$ and (1,2,3).
   (**b**) Find vectors x and y orthogonal to $(1, -1, 2)$ **and** independent of each other.

8. Let V be a vector space with inner product $\langle\ ,\ \rangle$. Let $\mathbf{x} \perp \mathbf{y}$ mean that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Show that
   (**a**) If $\mathbf{x} \perp \mathbf{y}$ and $\mathbf{x} \perp \mathbf{z}$, then $\mathbf{x} \perp (a\mathbf{y} + b\mathbf{z})$ for all a and b.
   (**b**) If $\mathbf{x} \perp \mathbf{v}_1, \mathbf{x} \perp \mathbf{x}_2, \ldots, \mathbf{x} \perp \mathbf{v}_n$, then $\mathbf{x}$ is orthogonal to any vector in span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

9. Furnish details for verifying the inner product in Example 1.

10. Furnish details for verifying the inner product in Example 2.

11. Prove part (**b**) and (**c**) of Theorem 3.6.1.

12. Prove part (**e**) of Theorem 3.6.1 for real vector spaces by following these steps

    (**a**) Show that the inequality is true for $\mathbf{x} = \theta$.

    (**b**) Assume $\mathbf{x} \neq \theta$ and use

    $$0 \leq \langle r\mathbf{x} + \mathbf{y}, r\mathbf{x} + \mathbf{y} \rangle \quad r \text{ real}$$

    to find

    $$0 \leq \underbrace{\|x\|^2}_{\text{Call } A} + \underbrace{2\langle \mathbf{x}, \mathbf{y} \rangle}_{\text{Call } B} r + \underbrace{\|y\|^2}_{\text{Call } C}$$

    (**c**) Recall that a quadratic

    $$Ar^2 + Br + C \tag{1}$$

    which is greater than or equal to 0 has either no or one real root. Calculate the discriminant (from the quadratic formula) for Eq. (1). It must be less than or equal to zero.

    (**d**) Use $B^2 - 4AC \leq 0$ to obtain the Cauchy-Schwarz inequality.

13. Show that if $\mathbf{x} \perp \mathbf{y}$, then $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

14. Show that if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set, then

    $$\|\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n\|^2 = \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 + \cdots + \|\mathbf{v}_n\|^2$$

15. In $\mathcal{M}_{n1}, \langle X, Y \rangle \equiv X^T Y$ is an inner product. Calculate $X^T Y$ for

    $$X = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \qquad Y = \begin{pmatrix} -2 \\ 4 \\ -3 \end{pmatrix}$$

16. In $\mathcal{C}_{n1}, \langle X, Y \rangle \equiv X^* Y$ is an inner product (recall $X^* = \overline{X^T}$). Calculate $X^* Y$ for

    $$X = \begin{pmatrix} 1 \\ i \end{pmatrix} \qquad Y = \begin{pmatrix} 1 - i \\ 2 \end{pmatrix}.$$

Calculate U^*V for

$$U = \begin{pmatrix} 1-i \\ 1 \end{pmatrix} \qquad V = \begin{pmatrix} 1-i \\ 2 \end{pmatrix}$$

17. Let A and B be matrices in $\mathcal{M}_{mn}$ and define $\langle A, B\rangle = \text{tr } A^T B$.
    (**a**) What choice of m and n makes $\langle A, B\rangle$ equivalent to the inner product in Example 3?
    (**b**) Show that $\langle A, B\rangle$ as defined is an inner product on $\mathcal{M}_{mn}$.

3.7 THE BASIS PROBLEM REVISITED

We consider in this section the following form of the basis problem:

> Given a basis $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of a vector space V with inner product $\langle \ , \ \rangle$ find an orthonormal basis $\mathcal{O} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ of V.

One reason for wanting an orthonormal basis for V is that for any vector $\mathbf{x} \in V$ it is easy to calculate the coefficients of $\mathbf{x}$ in the linear combination

$$\mathbf{x} = c_1\mathbf{u}_1 + c_1\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$$

In fact,

$$\begin{aligned}\langle \mathbf{x}, \mathbf{u}_k\rangle &= c_1\langle \mathbf{u}_1, \mathbf{u}_k\rangle + \cdots + c_k\langle \mathbf{u}_k, \mathbf{u}_k\rangle + \cdots + c_n\langle \mathbf{u}_n, \mathbf{u}_k\rangle \\ &= c_k\end{aligned}$$

Example 1. The set $\mathcal{O} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \{(1/\sqrt{2}, 0, 1/\sqrt{2}), (0, 1, 0), (1/\sqrt{2}, 0, -1/\sqrt{2})\}$ is an orthonormal basis for E^3. Write $(2, -1, 4)$ as a linear combination of the basis elements.

Solution We have

$$\begin{aligned}\mathbf{x} &= \langle \mathbf{x}, \mathbf{u}_1\rangle\mathbf{u}_1 + \langle \mathbf{x}, \mathbf{u}_2\rangle\mathbf{u}_2 + \langle \mathbf{x}, \mathbf{u}_3\rangle\mathbf{u}_3 \\ &= \frac{6}{\sqrt{2}}\mathbf{u}_1 + (-1)\mathbf{u}_2 + \frac{-2}{\sqrt{2}}\mathbf{u}_3\end{aligned}$$

A solution of the basis problem stated above involves what is known as the **Gram-Schmidt procedure**.

To get an idea of this procedure, consider the problem in E^3 and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, as pictured in Fig. 3.7.1. What we can do is to first obtain an orthogonal basis $T = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ and then to normalize all the vectors to obtain $\mathcal{O}$. For the first vector $\mathbf{w}_1$ in T, choose $\mathbf{w}_1 = \mathbf{v}_1$. Now $\mathbf{w}_2$ must be perpendicular to $\mathbf{w}_1$. That is, it must lie in the plane P_1 shown in Fig. 3.7.2. We see that there are an infinite number of choices of $\mathbf{w}_2$. The idea in the Gram-Schmidt process is to choose $\mathbf{w}_2$ as a linear combination of $\mathbf{w}_1$ and $\mathbf{v}_2$. Since span $\{\mathbf{w}_1, \mathbf{v}_2\}$ is the plane containing $\mathbf{w}_1$ and $\mathbf{v}_2$, we know that $\mathbf{w}_2$ must lie in the intersection of the plane P_1 and span $\{\mathbf{w}_1, \mathbf{v}_2\}$ (see Fig. 3.7.3). Once $\mathbf{w}_2$ is chosen, $\mathbf{w}_3$ must be found perpendicular to $\mathbf{w}_1$ and $\mathbf{w}_2$. In the Gram-Schmidt process, $\mathbf{w}_3$ is required to lie in span $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}_3\}$ and to be perpendicular to $\mathbf{w}_1$ and $\mathbf{w}_2$ (see Fig. 3.7.4). Finally then

$$\mathcal{O} = \left\{ \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} \right\}$$

Let us actually carry out this process in E^3 before stating the general result.

Example 2. Using the procedure just outlined, construct an orthogonal basis from $S = \{(1,1,0), (1,3,1), (2,2,3)\} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in E^3 with the standard inner product.

Solution We choose $\mathbf{w}_1 = (1,1,0)$ and require of $\mathbf{w}_2$ that

$$\begin{aligned} \langle \mathbf{w}_2, \mathbf{w}_1 \rangle &= 0 \\ \mathbf{w}_2 &= c_1 \mathbf{w}_1 + c_2 \mathbf{v}_2 \end{aligned} \tag{3.7.1}$$

Now take the inner product of both sides of Eq. (3.7.1) with $\mathbf{w}_1$ to find

$$0 = \langle \mathbf{w}_2, \mathbf{w}_1 \rangle = c_1 \langle \mathbf{w}_1, \mathbf{w}_1 \rangle + c_2 \langle \mathbf{v}_2, \mathbf{w}_1 \rangle$$

which is one homogeneous equation in two unknowns. The solutions are $c_1 = -c_2 \langle \mathbf{v}_2, \mathbf{w}_1 \rangle / \|\mathbf{w}_1\|^2$. Choosing $c_2 = 1$, we have

$$\mathbf{w}_2 = \frac{-\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2} \mathbf{w}_1 + \mathbf{v}_2 = \mathbf{v}_2 - \underbrace{\frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2} \mathbf{w}_1}_{\text{Projection of } v_2 \text{ on } w_1} \tag{3.7.2}$$

$$= (1,3,1) - \tfrac{4}{2}(1,1,0) = (-1,1,1)$$

Now having $\mathbf{w}_1 = (1, 1, 0)$ and $\mathbf{w}_2 = (-1, 1, 1)$, we require of $\mathbf{w}_3$ that

$$\begin{aligned} \langle \mathbf{w}_1, \mathbf{w}_3 \rangle &= 0 \\ \langle \mathbf{w}_2, \mathbf{w}_3 \rangle &= 0 \\ \mathbf{w}_3 &= c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + c_3\mathbf{v}_3 \end{aligned} \tag{3.7.3}$$

Now take inner products of both sides of Eqs. (3.7.3) with $\mathbf{w}_1$ and $\mathbf{w}_2$ to find equations for c_1, c_2, c_3:

$$\begin{aligned} 0 = \langle \mathbf{w}_1, \mathbf{w}_3 \rangle &= c_1\langle \mathbf{w}_1, \mathbf{w}_1 \rangle + c_2\langle \mathbf{w}_2, \mathbf{w}_1 \rangle + c_3\langle \mathbf{w}_1, \mathbf{v}_3 \rangle \\ 0 = \langle \mathbf{w}_2, \mathbf{w}_3 \rangle &= c_1\langle \mathbf{w}_2, \mathbf{w}_1 \rangle + c_2\langle \mathbf{w}_2, \mathbf{w}_2 \rangle + c_3\langle \mathbf{w}_2, \mathbf{v}_3 \rangle \end{aligned}$$

Since $\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 0$, these last equations reduce to

$$\begin{aligned} 0 &= c_1\|\mathbf{w}_1\|^2 && + c_3\langle \mathbf{w}_1, \mathbf{v}_3 \rangle \\ 0 &= & c_2\|\mathbf{w}_2\|^2 & + c_3\langle \mathbf{w}_2, \mathbf{v}_3 \rangle \end{aligned}$$

Putting $c_3 = 1$, we have

$$c_1 = \frac{-\langle \mathbf{w}_1, \mathbf{v}_3 \rangle}{\|\mathbf{w}_1\|^2} \qquad c_2 = \frac{-\langle \mathbf{w}_2, \mathbf{v}_3 \rangle}{\|\mathbf{w}_2\|^2}$$

$$\begin{aligned} \mathbf{w}_3 &= \mathbf{v}_3 - \overset{\substack{\text{Projection of} \\ v_3 \text{ on } w_2 \ \downarrow}}{\frac{\langle \mathbf{w}_2, \mathbf{v}_3 \rangle}{\|\mathbf{w}_2\|^2}}\mathbf{w}_2 - \overset{\substack{\text{Projection of} \\ v_3 \text{ on } w_1 \ \downarrow}}{\frac{\langle \mathbf{w}_1, \mathbf{v}_3 \rangle}{\|\mathbf{w}_1\|^2}}\mathbf{w}_1 \\ &= (2, 2, 3), -\tfrac{3}{3}(-1, 1, 1) - \tfrac{4}{2}(1, 1, 0) \\ &= (1, -1, 2) \end{aligned} \tag{3.7.4}$$

Thus an orthogonal basis is

$$T = \{(1, 1, 0), (-1, 1, 1), (1, -1, 2)\}$$

By normalizing each vector in T, we find an orthogonal basis

$$\left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) \right\}$$

Graphs of S and $\mathcal{O}$ are shown in Fig. 3.7.5.

The method of Example 2 actually furnishes a method of proof for a general result.

Theorem 3.7.1. *(Gram-Schmidt procedure) Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V with inner product $\langle\ ,\ \rangle$. The set $T = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$, where*

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 \\
\mathbf{w}_n &= \mathbf{v}_n - \frac{\langle \mathbf{v}_n, \mathbf{w}_{n-1}\rangle}{\|\mathbf{w}_{n-1}\|^2}\mathbf{w}_{n-1} - \cdots - \frac{\langle \mathbf{v}_n, \mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \frac{\langle \mathbf{v}_n, \mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1
\end{aligned} \tag{3.7.5}$$

is an orthogonal basis for V. The set

$$\mathcal{O} = \left\{ \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \ldots, \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|} \right\}$$

is an orthonormal basis for V.

Equation (3.7.5) are a generalization of Eqs. (3.7.2) and (3.7.4) in Example 2.

Example 3. Find an orthonormal basis for $V = \text{span}\{(1,-1,0,2), (3{,}2{,}1{,}2), (-2,1,0,1)\}$ in E^4 with the standard inner product.

Solution Let $\mathbf{v}_1 = (1,-1,0,2), \mathbf{v}_2 = (3,2,1,2)$, and $\mathbf{v}_3 = (-2,1,0,1)$. Using the formulas from Theorem 3.7.1, we have

$$\begin{aligned}
\mathbf{w}_1 &= (1,-1,0,2) \\
\mathbf{w}_2 &= (3,2,1,2) - \tfrac{5}{6}(1,-1,0,2) = (\tfrac{13}{6}, \tfrac{17}{6}, 1, \tfrac{1}{3}) = \tfrac{1}{6}(13,17,6,2) \\
\mathbf{w}_3 &= (-2,1,0,1) - \frac{-\frac{7}{6}}{\frac{498}{36}}\left(\frac{13}{6}, \frac{17}{6}, 1, \frac{1}{3}\right) - \frac{-1}{6}(1,-1,0,2) \\
&= \tfrac{1}{498}(-822, 534, 42, 678)
\end{aligned}$$

An orthonormal basis is

$$\begin{aligned}
&\left\{ \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}, \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} \right\} \\
&\quad = \left\{ \frac{1}{\sqrt{6}}(1,-1,0,2), \frac{1}{\sqrt{498}}(13,17,6,2), \frac{1}{\sqrt{355,572}}(-411,267,21,339) \right\}
\end{aligned}$$

Now that we know what an orthonormal basis is, we can describe an important class of matrices which we will use heavily later on.

Definition 3.7.1. A square matrix A is called **orthogonal** if $A^T = A^{-1}$, that is, if $AA^T = A^TA = I$.

Example 4. Show that
$$A = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}$$
is orthogonal.

Solution
$$A^TA = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
Therefore $A^T = A^{-1}$ and A is orthogonal.

The rows of A in Example 4 are, as vectors, $\mathbf{v}_1 = (\frac{3}{5}, \frac{4}{5})$ and $\mathbf{v}_2 = (-\frac{4}{5}, \frac{3}{5})$. Notice that in the standard inner product on E^2, $\mathbf{v}_1 \perp \mathbf{v}_2$, $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$. That is, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for E^2. In fact, the same can be said for the columns of A. This observation carries over to all orthogonal matrices, as the next theorems show.

Theorem 3.7.2. *A matrix $A_{n\times n}$ is orthogonal if and only if the rows of A form an orthonormal basis for E^n with the standard inner product.*

Proof. ($\Leftarrow$) Suppose the rows of A form an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Then
$$AA^T = \begin{pmatrix} - & \mathbf{v}_1 & - \\ - & \mathbf{v}_2 & - \\ & \vdots & \\ - & \mathbf{v}_n & - \end{pmatrix} \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{pmatrix}$$
Given the definition of matrix multiplication, the ij entry of the product is $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$, which is 1 if $i = j$ and 0 if $i \neq j$. Therefore $AA^T = I$.

($\Rightarrow$) Suppose A is orthogonal. Then $AA^T = I$. Then c_{ij}, the ij entry of AA^T, is 1 if $i = j$ and 0 if $i \neq j$. By the definitions of matrix multiplication and the transpose of a matrix, we see that $\langle \text{row } i, \text{row } j \rangle$ is 1 if $i = j$ and is 0 if $i \neq j$. Therefore the rows of $A_{n\times n}$ form an orthonormal basis for E^n. □

Theorem 3.7.3. *Matrix A is orthogonal if and only if A^T is orthogonal.*

Proof. If A is orthogonal, then $AA^T = A^T A = I$. Thus

$$A^T(A^T)^T = A^T A = I$$

and

$$(A^T)^T A^T = AA^T = I$$

Therefore A^T is orthogonal. The argument is reversible (see the problems). $\square$

Theorem 3.7.4. *A matrix $A_{n\times n}$ is orthogonal if and only if the columns of A form an orthogonal basis for E^n with the standard inner product.*

Proof. Matrix A is orthogonal if and only if A^T is orthogonal. The columns of A are the rows of A^T, which form an orthonormal basis for E^n. $\square$

Example 5. The identity matrix is an orthogonal matrix.

Example 6. Give an example of a 3×3 orthogonal matrix (not I).

Solution There are an infinite number of possibilities. By using the standard basis for E^3 we can obtain

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

Another possibility is

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice that $A^T = A$, so that $A^2 = I$. Such matrices are called **involutory**.

In the complex case, a square matrix A is called **unitary** if $A^* A = I$. For example,

$$A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

is unitary. We have virtually identical theorems for unitary matrices, and their proofs are virtually the same.

Theorem 3.7.5. *A square matrix A is unitary if and only if its rows and columns are mutually orthogonal with respect to the standard inner product in $\mathbb{C}^n$.*

Example 7. The matrix

$$\begin{pmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

is unitary.

Orthogonal Projections If W is a subspace of E^n with an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, then the **orthogonal projection of v onto** W is defined as

$$\text{proj}_W\ \mathbf{v} = \langle \mathbf{v} \cdot \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{v} \cdot \mathbf{v}_m \rangle \mathbf{v}_m$$

where the dot product is the standard dot product on E^n. For example, if $\mathbf{v} = (1, 2, 2)$ and $W = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) | \mathbf{x}_3 = 0\}$ with basis $S = \{(1, 0, 0), (0, 1, 0)\}$, then $\text{proj}_W(1, 2, 2) = (1, 2, 0)$ (see Fig. 3.7.6). From Fig. 3.7.6 it is not too hard to see that $\text{proj}_W\ \mathbf{v}$ is the vector in W "closest" to $\mathbf{v}$, in the sense that the distance from the "end of $\mathbf{v}$" to the "end of $\text{proj}_W\ \mathbf{v}$" is the smallest compared to all other vectors from W. That is,

$$\|\mathbf{v} - \text{proj}_W\ \mathbf{v}\| \leq \|\mathbf{w} - \mathbf{v}\|$$

for all vectors $\mathbf{w}$ in W. Although this is true in general, we do not prove it here. We note that if nonstandard inner products are used, the norms generated need not represent distance in the usual way.

Problems 3.7

1. Verify that the following are orthogonal bases of E^3. Convert them to orthonormal bases.

   (**a**) $\{(1, 1, 0), (0, 0, 1), (-1, 1, 0)\}$
   (**b**) $\{(1, 1, 1), (0, 1, -1), (1, -\frac{1}{2}, -\frac{1}{2})\}$

2. Use the Gram-Schmidt procedure to construct orthonormal bases for the following subspaces of E^4.

   (**a**) span $\{(1, 0, 0, 0), (1, 0, 1, 0), (1, 0, 1, 1)\}$
   (**b**) span $\{(1, 0, -1, 0), (1, 0, 1, 1), (1, -1, 0, 0)\}$

3. Which of the following are orthogonal matrices?

(**a**) $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ (**b**) $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(**c**) $\begin{pmatrix} \frac{5}{13} & -\frac{12}{13} & 0 \\ \frac{12}{13} & \frac{5}{13} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (**d**) $\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$

4. Which of the following are unitary matrices?

(**a**) $\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ (**b**) $\begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix}$

(**c**) $\begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}$ (**d**) $\begin{pmatrix} 1 & i & 3+i \\ -i & -1 & 2-2i \\ 3-i & 2+2i & 2 \end{pmatrix}$

5. Let $\mathcal{O} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an orthonormal basis for V. Let $\mathbf{x} \in V$. Show that if $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, then $\|\mathbf{x}\|^2 = {c_1}^2 {c_2}^2 + \cdots + {c_n}^2$.

6. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a linearly independent set of vectors in a vector space V of dimension $n, n > k$, with inner product. Show that if $\mathbf{x}$ is orthogonal to all the vectors in S, then $T = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{x}\}$ is also linearly independent ($\mathbf{x} \neq \theta$).

7. Show that

$$\mathcal{O} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{pmatrix} \right\}$$

is an orthonormal basis for $\mathcal{M}_{22}$, with the inner product

$$\left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right\rangle = ae + bf + cg + dh$$

Write

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

as a linear combination of the basis vectors.

8. Use the Gram-Schmidt procedure to transform the given bases to orthonormal bases in the spaces given. (Use the standard inner product.)

(**a**) $\{(1,1),(1,2)\}$ in E^2

(**b**) $\{(1,0),(3,7)\}$ in E^2

(**c**) $\{(1,1,-1),(0,1,-1),(1,1,0)\}$ in E^3

(**d**) $\{(1,2,3),(4,5,6),(1,1,0)\}$ in E^3

9. Reverse the argument in Theorem 3.7.3 to complete the proof.

10. Use the Gram-Schmidt procedure to find an orthonormal basis for span $\{(1,3,1),(1,2,1)\}$ in E^3 with the standard inner product. Do the same in E^3 with the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = 3x_1y_1 + 2x_2y_2 + x_3y_3$.

11. Find an orthonormal basis for E^3 which includes $\mathbf{v} = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ (do this by inspection).

12. Find an orthogonal basis for E^3 which includes $\mathbf{v}_1 = (1,1,1)$. First find any basis which includes $\mathbf{y}_1$ as the first vector; then use the Gram-Schmidt procedure, but do not normalize $\mathbf{v}_1$.

13. (**a**) Let A be a unitary matrix. Show that $|\det A| = 1$.

    (**b**) Show by example that for a unitary matrix A, $\det A$ need not be 1 or -1. (**Note**: Remember that A can have complete entries.)

    (**c**) Let A be an orthogonal matrix. Show that $\det A = \pm 1$. In general, a matrix A with $\det A = 1$ is called **unimodular**.

14. What conditions on a and b guarantee that

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

    is orthogonal? Is there a condition which will make A unitary?

15. Is the square of an orthogonal matrix orthogonal? Is the square of a unitary matrix also unitary?

16. If A is orthogonal, is A^n orthogonal ($n \geq 2$)?

17. (**a**) Show that if U is orthogonal and $B = U^TAU$,then $\det B = \det A$.

(**b**) Show that if U is unitary and $B = U^*AU$,then $\det B \det A$.

18. If A and B are orthogonal what about $A + B$? What about $-A$?

19. Let A be involutory. Show that if n is odd, then $A^n = A$.

20. If A is unitary, is $-A$ unitary?

21. If A is idempotent, must A be symmetric?

22. Show that if $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for V, then for any $\mathbf{x}$ in V,

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{x}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n$$

23. Develop a proof of Theorem 3.7.5 by mimicking the proofs of Theorems 3.7.2, 3.7.3, and 3.7.4.

24. Show that the rotation matrix

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

is orthogonal. Is the matrix unitary?

25. Let a and b be real numbers. Under what conditions is

$$\begin{pmatrix} 0 & a & 0 & ib \\ a & 0 & ib & 0 \\ 0 & ib & 0 & a \\ ib & 0 & a & 0 \end{pmatrix}$$

unitary? (This is an example of a **scattering matrix**.)

26. The set $S = \{(1/\sqrt{2}, -1/\sqrt{2}, 0), (0, 1, 0)\}$ does not span E^3. Find the orthogonal projections of the following vectors on span S. Sketch span S and the vector. (**a**) (1,0,0) (**b**) $(1, -1, 0)$ (**c**) (1,1,2)

3.8 CHANGING BASES IN VECTOR SPACES

Ask several people directions to a certain location, and the instructions may sound different yet be equivalent. In Fig. 3.8.1 we see a map of a small town with an origin 0, where we ask directions to the corner of McDaniel and Walker streets, marked by an X. Two possible sets of directions are

1. Go 8 blocks east and 6 blocks north.

2. Go 6 blocks north and 8 blocks east.

Both sets of directions are correct. What is the difference? Each set is related to a different basis of E^2! In Fig. 3.8.2 we have drawn two bases S_1 and S_2

FIGURES

which correspond to cases 1 and 2 respectively, above. We have in units of blocks

$$S_1 = \{(1,0),(0,1)\} \qquad S_2 = \{(0,1),(1,0)\}$$

The order of the basis vectors is switched in S_1 and S_2. The corner is at (8,6), yet in terms of the different bases (8,6) has different coefficients, in terms of order

$$(8,6) = 8(1,0) + 6(0,1)$$
$$(8,7) = 6(0,1) + 8(1,0))$$

Note that the coefficients in the linear combinations above are actually the "directions" in terms of the given basis. Comparing S_1 and S_2, we see that **order** is important in a basis when we are using the basis to describe vectors. We also note that the "direction" vectors

$$\begin{pmatrix} 8 \\ 6 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 6 \\ 8 \end{pmatrix}$$

can be related by means of matrices:

$$\begin{pmatrix} 8 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 6 \\ 8 \end{pmatrix}$$

In this section we analyze the **problem of describing a given vector in terms of different bases**. This type of problem is of interest in mechanics.

Consider a guy wire, as shown in Fig. 3.8.3. The vector $\mathbf{F}$ represents a force acting horizontally on the wire. In terms of the standard basis (with x_1 and x_2 axes as shown), $\mathbf{F} = 2(1,0) + 0(0,1)$. Of interest are the normal and tangential forces, $\mathbf{T}$ and $\mathbf{N}$, on the wire. In terms of the basis $\{\mathbf{U}_1, \mathbf{U}_2\} = \{(1/\sqrt{2}, -1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$, however, $\mathbf{F} = \sqrt{2}\mathbf{U}_1 + \sqrt{2}\mathbf{U}_2$. That is, the tangential and normal components are both $\sqrt{2}$.

Definition 3.8.1. An **ordered basis** is a basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ of vectors along with an ordering[2] of the vectors in the basis.

Definition 3.8.2. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an ordered basis for a vector space V, and let $\mathbf{x} \in V$. If

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \tag{3.8.1}$$

then the coefficients $c_1, c_2, \dots, c_n$ are called the **coordinates of x with respect to** S. The matrix

$$(\mathbf{x})_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

is called the **coordinate matrix of** x **with respect to** S.

A coordinate matrix of x for a given ordered basis is unique, because the linear combination in Eq. (3.8.1) is unique (by Theorem 3.5.1).

Example 1. Let $\mathbf{x} = (2,3)$ in E^2. Find the coordinate matrix for $\mathbf{x}$ with respect to (**a**) $S =$ standard basis, (**b**) $S = \{(1,1), (1,-1)\}$, (**c**) $S = \{(1,0), (1,1)\}$, and (**d**) $S = \{(1,1), (1,0)\}$.

Solution

(**a**) Since $(2,3) = 2(1,0) = 3(0,1)$,

$$((2,3))_S = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

[2] An ordering of a set S of n objects is a function from the set $\{1, 2, \dots, n\}$ onto S. It is simply a rule for telling which object of S is to be called the first, the second, and so on.

(**b**) Since $\{(1,1),(1,-1)\}$ is an orthogonal basis,

$$(2,3) = \frac{\langle (2,3),(1,1)\rangle}{\|(1,1)\|^2}(1,1) + \frac{\langle (2,3),(1,-1)\rangle}{\|(1,-1)\|^2}(1,-1)$$

(see Prob. 22 of Sec. 3.7). Thus

$$((2,3))_S = \begin{pmatrix} \frac{5}{2} \\ -\frac{1}{2} \end{pmatrix}$$

(**c**) We solve $(2,3) = c_1(1,0) + c_2(1,1)$ to find $c_1 = -1$ and $c_2 = 3$, so

$$((2,3))_S = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

(**d**) In this case $(2,3) = c_1(1,1) + c_2(1,0)$. Solving, we find $c_1 = 3$ and $c_2 = -1$.

Therefore

$$((2,3))_S = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

Notice in (**c**) and (**d**) that the reversal of the order of the basis elements resulted in a reversal of the elements of the coordinate matrix.

Example 2. $S = \{(1,0,1),(0,1,1),(1,2,4)\}$ is a basis for E^3. Suppose

$$(\mathbf{x})_S = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$$

Find $\mathbf{x}$.

Solution In terms of the basis S,

$$\begin{aligned} \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} &\to -1(1,0,1) + 2(0,1,1) + 4(1,2,4) \\ &= (3,10,17) \end{aligned}$$

The arrow simply replaces the word "represents."

From Example 1 we see that if we change the basis for a vector space V, then the coordinate matrix for a vector changes. This poses the problem:

> Given a vector space V with basis S, we can calculate $(\mathbf{x})_S$. If we give V a new basis T, can we calculate $(\mathbf{x})_T$ by using $(\mathbf{x})_S$? That is, is there a simple relationship between $(\mathbf{x})_S$ and $(\mathbf{x})_T$?

Let us consider the problem in E^3. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $T = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ be two bases for E^3. Suppose $\mathbf{x} \in E^3$ and

$$(\mathbf{x})_S = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

This is,

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \tag{3.8.2}$$

Now we want to "convert the $\mathbf{v}$'s to $\mathbf{w}$'s." Since T is a basis, there are coefficients $a_{11}, a_{12}, \ldots, a_{33}$ such that

$$\begin{aligned} \mathbf{v}_1 &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + a_{31}\mathbf{w}_3 \\ \mathbf{v}_2 &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + a_{32}\mathbf{w}_3 \\ \mathbf{v}_3 &= a_{13}\mathbf{w}_1 + a_{23}\mathbf{w}_2 + a_{33}\mathbf{w}_3 \end{aligned} \tag{3.8.3}$$

Substituting these into (3.8.2), we find

$$\begin{aligned} \mathbf{x} = {} & (a_{11}c_1 + a_{12}c_2 + a_{13}c_3)\mathbf{w}_1 + (a_{21}c_1 + a_{22}c_2 + a_{23}c_3)\mathbf{w}_2 \\ & + (a_{31}c_1 + a_{32}c_2 + a_{33}c_3)\mathbf{w}_3 \end{aligned}$$

and so

$$(\mathbf{x})_T = \begin{pmatrix} a_{11}c_1 + a_{12}c_2 + a_{13}c_3 \\ a_{21}c_1 + a_{22}c_2 + a_{23}c_3 \\ a_{31}c_1 + a_{32}c_2 + a_{33}c_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

By Eq. (3.8.3) $= ((\mathbf{v}_1)_T(\mathbf{v}_2)_T(\mathbf{v}_3)_T)(\mathbf{x})_S$
Therefore,

$$(\mathbf{x})_T = ((\mathbf{v}_1)_T(\mathbf{v}_2)_T(\mathbf{v}_3)_T)(\mathbf{x})_S$$

Coordinate matrix of x in new basis T

Columns are coordinate matrices of old basis vectors with respect to new basis

Coordinate matrix of x in old basis

In an n-dimensional vector space V, if $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is the old basis and T is the new basis, then

$$(\mathbf{x})_T = ((\mathbf{v}_1)_T(\mathbf{v}_2)_T \cdots (\mathbf{v}_n)_T)(\mathbf{x})_S \tag{3.8.4}$$

This can be derived in the same way as in the E^3 case. The matrix in Eq. (3.8.4) is called the **transition matrix from** S **to** T, and it is customarily denoted P. Thus we have $(\mathbf{x})_T = P(\mathbf{x})_S$. Occasionally we write $P_{T\leftarrow S}$ when it is necessary to emphasize the fact that P is the transition matrix **from** S **to** T.

These are two important facts about transition matrices:

1. Matrix P is invertible, and P^{-1} is the transition matrix from T to S.
2. If S and T are orthonormal bases, then P is an orthogonal matrix.

The proofs of these facts are outlined in the problems.

Example 3. Consider E^2 and bases $S = \{(1,1),(1,-1)\}$ and $T = \{(1,0),(1,1)\}$.

(**a**) Find the transition matrix from basis S to basis T.

(**b**) Find the transition matrix from basis T to basis S.

(**c**) Show that the matrices found in (**a**) and (**b**) are inverses of each other.

(**d**) Find $((3,-5))_T$.

(**e**) Find $((3,-5))_T$ by using the transition matrix from (**a**).

Solution

(**a**)

$$P = ((\mathbf{v}_1)_T(\mathbf{v}_2)_T) = (((1,1))_T((1,-1))_T).$$

Since

$$\mathbf{v}_1 = (1,1) = 0(1,0) + 1(1,1)$$

and

$$\mathbf{v}_2 = (1,-1) = 2(1,0) - 1(1,1)$$

we have

$$P_{T\leftarrow S} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix}$$

(**b**) Let $Q = ((\mathbf{w}_1)_S(\mathbf{w}_2)_S) = (((1,0))_S((1,1))_S)$. Since

$$\mathbf{w}_1 = (1,0) = \tfrac{1}{2}(1,1) + \tfrac{1}{2}(1,-1)$$

and

$$\mathbf{w}_2 = (1,1) = 1(1,1) + 0(1,-1)$$

we have

$$Q_{S \leftarrow T} = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$$

(**c**) Since

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

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} weknowthat\mathrm{P}^{-1} = Q$ and $Q^{-1} = P$.

d) Now $(3,-5) = -1(1,1) + 4(1,-1)$, so

$$((3,-5))_S = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

(**e**) Now

$$((3,-5))_T = P((3,-5))_S = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ -5 \end{pmatrix}$$

We can check our answer by computing

$$8\mathbf{w}_1 - 5\mathbf{w}_2 = 8(1,0) - 5(1,1) = (3,-5)$$

$$\begin{pmatrix} 8 \\ -5 \end{pmatrix}$$

Example 4. Consider E^3 with orthonormal bases (in the standard inner product)

$$S = \{(1/\sqrt{2}, -1/\sqrt{2}, 0), (1/\sqrt{2}, 1/\sqrt{2}, 0)(0,0,1)\}$$
$$T = \{(1/\sqrt{2}, 0, 1/\sqrt{2}), (0,1,0), (-1/\sqrt{2}, 0, 1/\sqrt{2})\}$$

Show that the transition matrix P from S to T is orthogonal. Write the transition matrix Q from T to S.

Solution Using inner products

$$(v_1)_T = \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{w}_1 \rangle \\ \langle \mathbf{v}_1, \mathbf{w}_2 \rangle \\ \langle \mathbf{v}_1, \mathbf{w}_3 \rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -1/\sqrt{2} \\ -\frac{1}{2} \end{pmatrix}$$

$$(v_2)_T = \begin{pmatrix} \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \\ \langle \mathbf{v}_2, \mathbf{w}_2 \rangle \\ \langle \mathbf{v}_2, \mathbf{w}_3 \rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 1/\sqrt{2} \\ -\frac{1}{2} \end{pmatrix}$$

$$(v_3)_T = \begin{pmatrix} \langle \mathbf{v}_3, \mathbf{w}_1 \rangle \\ \langle \mathbf{v}_3, \mathbf{w}_2 \rangle \\ \langle \mathbf{v}_3, \mathbf{w}_3 \rangle \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$$

we know that

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1/\sqrt{2} \end{pmatrix}$$

Matrix P is orthogonal since its columns are orthogonal and have norm 1. Since P^{-1} is the transition matrix from T to S, we know that $Q = P^{-1}$. However, P is orthogonal so $P^{-1} = P^T$ and

$$Q = P^T = \begin{pmatrix} \frac{1}{2} & -1/\sqrt{2} & -\frac{1}{2} \\ \frac{1}{2} & 1/\sqrt{2} & -\frac{1}{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}$$

Example 5. In Example 4, if the first and second vectors in S are interchanged, how is P changed?

Solution Examining the solution to Example 4, we see that the roles of $\mathbf{v}_1$ and $\mathbf{v}_2$ are reversed. So the transition matrix is

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1\sqrt{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1/\sqrt{2} \end{pmatrix}$$

That is, the first and second columns are interchanged.

Example 6. In Example 4, if the first and second vectors in T are interchanged, how is P changed?

Solution Examining the solution of Example 4, we see that since the roles of $\mathbf{w}_1$ and $\mathbf{w}_2$ are reversed, the first two entries in each coordinate matrix $(\mathbf{v}_1)_T, (\mathbf{v}_2)_T$, and $(\mathbf{v}_3)_T$ are reversed. Therefore the transition matrix is

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

That is, the first and second **rows** are interchanged.

Examples 5 and 6 illustrate the general principle that if **vectors in basis S are interchanged**, then **corresponding columns in $P_{T \leftarrow S}$ are interchanged**. If **vectors in basis T are interchanged**, then **corresponding rows in $P_{T \leftarrow S}$ are interchanged**.

Example 7. Consider the basis $T = \{(-1/\sqrt{2}, 1/\sqrt{2})(-1/\sqrt{2}, -1/\sqrt{2})\}$ in E^2 obtained from the basis $S = \{(0,1), (-1,0)\}$ by rotating the vectors in S by $\pi/4$ radians (Fig. 3.8.4). Find the transition matrix from S to T.

Solution Since the bases are orthonormal, we can calculate the coordinate matrices by using inner products:

$$(\mathbf{v}_1)_T = ((0,1))_T = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

$$(\mathbf{v}_2)_T = ((-1,0))_T = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Therefore

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

Incidentally P^T is

$$\begin{pmatrix} \cos\pi/4 & -\sin\pi/4 \\ \sin\pi/4 & \cos\pi/4 \end{pmatrix}$$

Example 7 illustrates the general form of a rotation as discussed in Chap. 2. In general, rotations have orthogonal transitions matrices.

Example 8. Recall the example of the guy wire in Fig. 3.8.3. Find the transition matrix from the standard basis for E^2.

Solution We have

$$(1,0) = \frac{1}{\sqrt{2}}\mathbf{U}_1 + \frac{1}{\sqrt{2}}\mathbf{U}_2 \quad \text{and} \quad (0,1) = \frac{1}{\sqrt{2}}\mathbf{U}_1 - \frac{1}{\sqrt{2}}\mathbf{U}_2$$

so that

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

Thus, if $\mathbf{F} = (a, b)$, its normal and tangential components are given by

$$P\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} (1/\sqrt{2})(a+b) \\ (1/\sqrt{2})(a-b) \end{pmatrix} \begin{array}{l} \longleftarrow \text{Normal} \\ \longleftarrow \text{Tangential} \end{array}$$

PROBLEMS 3.8

In Probs. 1 to 4, let $S = \{(1,3), (-2,1)\}, T = \{(1/\sqrt{2}, -1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$, $U = \{(-2,1), (1,3)\}$, and $Z = \{(\sqrt{3}/2, \frac{1}{2}), (-\frac{1}{2}, \sqrt{3}/2)\}$ be bases for E^2 with the standard inner product.

1. Find the transition matrices (**a**) from S to T, (**b**) from T to U, and (**c**) from S to U. Multiply the matrices from (**a**) and (**b**) in both ways. Is either product related to the matrix from (**c**)?

2. Let $\mathbf{x} = (1,1)$. Find $(\mathbf{x})_S$. Using the matrices from Prob. 1, find $(\mathbf{x})_T$ and $(\mathbf{x})_U$.

3. Find the transition matrices from T to Z and Z to T.

4. (**a**) Find the transition matrices from S to Z. (**b**) Find the transition matrix from U to Z by using interchanges in the matrix from (**a**).

In Probs. 5 to 10, let $S = \{(1,1,0), (0,1,1), (1,0,1)\}$, $T = \{(1/\sqrt{2}, -1/\sqrt{2}, 0), (1/\sqrt{2}, 1/\sqrt{2}, 0), (0,0,1)\}$, $U = \{(1,0,1), (0,1,1), (1,1,0)\}$ and $Z = \{(1,0,0), (0, 1/\sqrt{2}, -1/\sqrt{2}), (0, 1/\sqrt{2}, 1/\sqrt{2})\}$ be bases for E^3 with the standard inner product. Let $\mathbf{x} = (1,-1,2)$.

5. Find the transition matrices (**a**) from S to T, (**b**) from T to U, and (**c**) from S to U. Multiply the matrices from (**a**) and (**b**) in both ways. Is either product related to the matrix from (**c**)?

6. Find $(\mathbf{x})_S$. Using the matrices from Prob. 1, find $(\mathbf{x})_T$ and $(\mathbf{x})_U$.

7. Find the transition matrices from T to Z and from Z to T.

8. (**a**) Find the transition matrix from S to Z. (**b**) Find the transition matrix from U to Z by using interchanges in the matrix from (**a**).

9. If
$$(\mathbf{x})_S = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$$
what is $\mathbf{x}$?

10. If
$$(\mathbf{x})_T = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
what is $\mathbf{x}$?

11. If S is any basis for V and
$$(\mathbf{x})_S = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
what is $\mathbf{x}$?

12. If S is any basis for V and $\mathbf{x} = \theta$, what is $(\mathbf{x})_S$?

13. Show that if P is the transition matrix from S to T in the vector space V, then P^{-1} exists and P^{-1} is the transition matrix from T to S, by completing the following steps.

    (**a**) Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $T = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ and Q be the transition matrix from T to S. Write QP is
$$QP = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

    (**b**) Multiply $(\mathbf{x})_T = P(\mathbf{x})_S$ by Q.

(**c**) Substitute the result from (**b**) into $(\mathbf{x})_S = Q(\mathbf{x})_T$ to obtain

$$(\mathbf{x})_S = QP(\mathbf{x})_S \tag{1}$$

(**d**) Let $\mathbf{x} = \mathbf{v}_1$ in Eq. (1), and show

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$$

(**e**) Let $\mathbf{x} = \mathbf{v}_k$ in Eq. (1), and show

$$k\text{th place} \longrightarrow \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{kk} \\ \vdots \\ a_{nk} \end{pmatrix}$$

(**f**) Conclude that $QP = I$.

14. Show that if $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $T = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ are orthonormal bases of a real vector space V, then the transition matrix from S to T is orthogonal by completing the following steps.

(**a**) Show that transition matrix from S to T is

$$P = \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{w}_1\rangle & \langle \mathbf{v}_2, \mathbf{w}_1\rangle & \cdots & \langle \mathbf{v}_n, \mathbf{w}_1\rangle \\ \langle \mathbf{v}_1, \mathbf{w}_2\rangle & \langle \mathbf{v}_2, \mathbf{w}_2\rangle & \cdots & \langle \mathbf{v}_n, \mathbf{w}_2\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \mathbf{v}_1, \mathbf{w}_n\rangle & \langle \mathbf{v}_2, \mathbf{w}_n\rangle & \cdots & \langle \mathbf{v}_n, \mathbf{w}_n\rangle \end{pmatrix}$$

(**b**) Show that the transition matrix from T to S is

$$Q = \begin{pmatrix} \langle \mathbf{w}_1, \mathbf{v}_1\rangle & \langle \mathbf{w}_2, \mathbf{v}_1\rangle & \cdots & \langle \mathbf{w}_n, \mathbf{v}_1\rangle \\ \langle \mathbf{w}_1, \mathbf{v}_2\rangle & \langle \mathbf{w}_2, \mathbf{v}_2\rangle & \cdots & \langle \mathbf{w}_n, \mathbf{v}_2\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \mathbf{w}_1, \mathbf{v}_n\rangle & \langle \mathbf{w}_2, \mathbf{v}_n\rangle & \cdots & \langle \mathbf{w}_n, \mathbf{v}_n\rangle \end{pmatrix}$$

(**c**) From Prob. 13 we know that $Q = P^{-1}$. Use (**a**) and (**b**) to show that $Q = P^T$.

(**d**) Conclude that $P^{-1} = P^T$ and so P is orthogonal.

15. Show that if P is the transition matrix from S to T, and Q is the transition matrix from T to U, where S, T, and U are bases for V, then QP is the transition matrix from S to U.

16. Let $S = \{1+x, 1+x^2, 1-x^2\}$ and $T = \{1, 1+x^2, 1+x\}$ be bases for P_2. Find the transition matrix from S to T.

17. Show that a transition matrix for an n-dimensional vector space V is $n \times n$.

18. In $\mathbb{C}^2$ let $S = \{(i,0),(0,i)\}$ and $T = \{(1,0),(0,1)\}$. Calculate the transition matrices from S to T and from T to S.

19. Consider the guy wire as shown. Compute the transition matrix from the standard basis for E^2 to $\{\mathbf{U}_1, \mathbf{U}_2\}$, where $\mathbf{U}_1$ and $\mathbf{U}_2$ have unit length.

20. Consider the guy wire at angle α as shown. Compute the transition matrix from the standard basis for E^2 to $\{\mathbf{U}_1, \mathbf{U}_2\}$, where $\mathbf{U}_1$ and $\mathbf{U}_2$ have unit length.

3.9 CALCULUS REVISITED

Vector spaces, bases and orthogonality can be used to discuss many topics from calculus. We mention a few here.

Vector Spaces

Recall the theorem about sums and multiples of continuous functions: If f and g are continuous on an interval $[a,b]$ and c is any real number, then

$$f + g \text{ is a continuous function on } [a,b]$$

and

$$cf \text{ is a continuous function on } [a,b]$$

This theorem says that the set of functions continuous on $[a,b]$ is closed under sums and constant multiples. Thus we have a situation which may yield a vector space. With this in mind, we define $C[a,b]$ as the set of functions continuous on $[a,b]$ along with addition of functions and constant multiplication of functions defined in the usual way. To see whether $C[a,b]$ is a vector space, we must check the vector space axioms. Closure holds by the theorem stated above. The properties

$$f+g=g+f \quad (f+g)+h=f+(g+h) \quad c(f+g)=cf+cg$$
$$(b+c)f=bf+cf \quad 1f=f \quad (bc)f=b(cf)$$

are easily checked. The zero vector is the constant function (it is continuous) which is identically zero on $[a,b]$. Given f, the function $-f$ is continuous and is the additive inverse of f. Therefore, $C[a,b]$ is a vector space.

Differentiable functions are another important set of functions in calculus. If we define $C^1[a,b]$ as the set of functions f defined on $[a,b]$ with $f' \in C[a,b]$ (that is, f' is continuous) and give $C^1[a,b]$ the same operations as $C[a,b]$, then $C^1[a,b]$ is a subspace of $C[a,b]$. To see this, we need only check closure. Since a theorem from calculus tells us that the sum and constant multiples of differentiable functions are differentiable, we have the necessary closure. So $C^1[a,b]$ is a vector space in its own right as well as a subspace of $C[a,b]$.

Example 1. Give an example of a function in $C[-1,1]$ which is not in $C^1[-1,1]$. This shows that $C^1[-1,1]$ is a proper subspace of $C[-1,1]$.

Solution A continuous function which is not differentiable will be sufficient. The function $y=|x|$ is continuous on $[-1,1]$ but fails to have a derivative at $x=0$.

Several functions from calculus are differentiable an infinite number of times. For example, $y=e^x$, $y=\sin x$, $y=\cos x$, and all polynomials are infinitely differentiable over all $\mathbb{R}$. Thus we can define vector spaces

$$C^2[a,b], C^3[a,b], \ldots, C^n[a,b], \ldots$$

and even

$$C^\infty[a,b]$$

Definition 3.9.1. $C^n[a,b]$ is the set of all functions f defined on $[a,b]$ with $f, f', f'', \ldots, f^{(n)}$ being continuous on $[a,b]$. (Right- and left-hand derivatives are used at a and b.) With the standard operations of addition and multiplication by a constant, $C^n[a,b]$ is a vector space. $C^\infty[a,b]$ is the set of all f in $C^n[a,b]$ for all n. $C^\infty[a,b]$ is a vector space.

Example 2. Give a example of a function which is in $C^1[-1,1]$ but is not in $C^2[-1,1]$.

Solution Let $y = x^{5/3}$. Then $y' = \frac{5}{3}x^{2/3}$ and $y'' = \frac{10}{9}x^{-1/3}$. Although y' is continuous on $[-1,1], y''$ is not, because $y''(0)$ is undefined.

Examples 1 and 2 illustrate the subspace relationship

$$C[a,b] \supset C^1[a,b] \supset C^2[a,b] \supset \cdots \supset C^\infty[a,b]$$

among these vector spaces.

The vector space $C[a,b]$ differs from the vector spaces we have been studying in that $C[a,b]$ is **not finite-dimensional**. To see this, suppose that $\dim(C[a,b]) = n$. Then a finite basis $S = \{f_1(x), f_2(x), \dots, f_n(x)\}$ of continuous functions in $C[a,b]$ would exist, and any set of $n+1$ functions would have to be linearly dependent. However, $T = \{1, x, x^2, \dots, x^n\}$ is a set of $n+1$ functions in $C[a,b]$ which are linearly independent. So it is impossible to have $\dim(C[a,b]) = n$ for any n. Since polynomials are infinitely differentiable, the same argument shows that $C^k[a,b]$ is not finite-dimensional for any k.

Although $C[a,b]$ is not finite-dimensional, there are important finite-dimensional subspaces of $C[a,b]$.

Example 3. Show that the set of all solutions to the equation[3] $y' = \alpha y$ is a one-dimensional subspace of $C(\mathbb{R})$

Solution We recall that the functions[4] satisfying $y' = \alpha y$ are $y(x) = Ce^{\alpha x}$, where C is arbitrary. That is, $W = \{y | y(x) = Ce^{\alpha x}\}$. To show that W is a subspace of $C(\mathbb{R})$, we first note that since e^x is a continuous function, $W \subseteq C(\mathbb{R})$. Now closure of W under the operations must be shown. Since $c_1e^{\alpha x} + c_2e^{\alpha x} = (c_1 + c_2)e^{\alpha x} = c_3e^{\alpha x}$ and $c_1(c_2e^{\alpha x}) = (c_1c_2)e^{\alpha x}$, we have the closure, so W is a subspace of $C(\mathbb{R})$. A basis for W is $S = \{e^{\alpha x}\}$. Thus $\dim W = 1$.

[3]Recall that the equation $y' = \alpha y$ is important in the study of population growth, radioactive decay, and other areas.

[4]That these are all the functions satisfying $y' = \alpha y$ is a theorem of ordinary differential equations.

Example 4. Show that the set of all solutions to the equation $y'' = -\alpha^2 y$ is a two-dimensional subspace of $C(\mathbb{R})$.

Solution The solutions of the equation are of the form $A\cos\alpha x + B\sin\alpha x$. Showing that $W = \{y|y(x) = A\cos\alpha x + B\sin\alpha x\}$ is a subspace is done in the same way as Example 3. A basis for W is $S = \{\cos\alpha x, \sin\alpha x\}$. That S spans W is easy. For the linear independence consider $c_1\cos\alpha x + c_2\sin\alpha x = 0$. The equation must hold for all x. Substitute $x = 0$ to find $c_1 = 0$. Substitute $x = \pi/(2\alpha)$ to obtain $c_2 = 0$.

Linear Independence

Linear independence of a set of functions in $C[a, b]$ is not always easy to show.

Example 5. Consider the set $S = \{\sin 2x, \sin x\cos x\}$ in $C(\mathbb{R})$. Is S a linearly independent set?

Solution We could look at

$$c_1\sin 2x + c_2\sin x\cos x = \theta = 0 \tag{3.9.1}$$

as in the definition. The difficulty here is that since the vectors are functions, we do not obtain a set of linear equations with constant coefficients. The coefficients of the linear equation involve the variable x, and the equation must hold for all x. In this example, we rely on our knowledge of trigonometric identities and recall that

$$\frac{1}{2}\sin 2x = \sin x\cos x$$

So the set is linearly dependent.

There is another method available if the functions in a set of functions are differentiable enough times, that is, the method of the wronskian. For two functions the basic result is as follows:

Functions f and g in $C^1[a, b]$ are linearly independent if

$$\det\begin{pmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{pmatrix} \neq 0 \quad \text{for some } x \text{ in } [a, b]$$

The determinant is called the **wronskian** of f and g and is denoted $W(f, g)$.

Example 6. Show that $S = \{\sin x, \cos x\}$ is linearly independent in $C^1(\mathbb{R})$.

Solution We have

$$W(\sin x, \cos x) = \det \begin{pmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{pmatrix} = -\sin^2 x - \cos^2 x = -1$$

Since $W(\sin x, \cos x) \neq 0$ for all x, set S is linearly independent.

Let us show why $W(f, g) \neq 0$ is sufficient for linear independence. Consider the usual equation in checking for independence

$$c_1 f(x) + c_2 g(x) = 0$$

Now we have, after differentiating the equation, the system

$$\begin{pmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and by Cramer's rule the system has only the zero solution when

$$\det \begin{pmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{pmatrix} \neq 0$$

For larger sets of functions we have this result:
The set $\{f_1, f_2, \ldots, f_n\}$ in $C^n[a, b]$ is linearly independent if

$$W(f_1, \ldots, f_n) = \det \begin{pmatrix} f_1 & f_2 & \cdots & f_n \\ f_1' & f_2' & \cdots & f_n' \\ f_1'' & f_2'' & \cdots & f_n'' \\ \cdots & \cdots & \cdots & \cdots \\ {f_1}^{(n)} & {f_2}^{(n)} & \cdots & {f_n}^{(n)} \end{pmatrix} \neq 0$$

for some x in $[a, b]$.

Orthogonality

An inner product on $C[a, b]$ can be defined by

$$\langle f, g \rangle = \int_a^b f(x) g(x) w(x) dx$$

where $w(x)$ is a fixed continuous function, positive on $[a, b]$; it is called a **weight function**. The simplest choice for $w(x)$ is $w(x) \equiv 1$. Once $\langle f, g \rangle$ is defined, we can say, "f is orthogonal to g if $\langle f, g\rangle = 0$" and "The norm of f is $\sqrt{\langle f, f\rangle}$," or we can make any other statement which makes sense in a vector space with inner product. We can even speak of the "angle" between two functions.

Example 7. If $C[-1,1]$, let $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$. Calculate $\langle x, x^2\rangle$, $\langle 1+x^2, x-x^2\rangle$, $\|1+x\|$, and $\langle x^n, x^m\rangle$, where n is even and m is odd.

Solution

$$\begin{aligned}
\langle x, x^2\rangle &= \int_{-1}^{1} x^3 \, dx = \left.\frac{x^4}{4}\right|_{-1}^{1} = 0 \\
\langle 1+x^2, x-x^2\rangle &= \int_{-1}^{1} (1+x^2)(x-x^2)dx \\
&= \int_{-1}^{1} (x - x^2 + x^3 - x^4)dx = -\frac{16}{15} \\
\|1+x\| &= \sqrt{\langle 1+x, 1+x\rangle} = \sqrt{\int_{-1}^{1} (1+x)^2 dx} \\
&= \sqrt{\left.\frac{(1+x)^3}{3}\right|_{-1}^{1}} = \sqrt{\frac{8}{3}} = \frac{2\sqrt{6}}{3} \\
\langle x^n, x^m\rangle &= \int_{-1}^{1} x^{n+m} dx = \left.\frac{x^{n+m+1}}{n+m+1}\right|_{-1}^{1} = 0
\end{aligned}$$

since $n+m+1$ is even. With this inner product x^{2n+1} and x^{2m} are orthogonal.

Example 8. In $C[0,1]$ with inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$, find an orthonormal basis for span $\{1, x, x^2\}$.

Solution Since the Gram-Schmidt procedure works for any vector space with an inner product, we use it. We have

$$
\begin{aligned}
w_1(x) &= 1 \\
w_2(x) &= x - \frac{\langle x, 1\rangle}{\|1\|^2}1 \\
&= x\frac{\int_0^1 x\ dx}{\int_1^1 1\ dx}1 = x - \frac{1}{2} \\
w_3(x) &= x^2 - \frac{\langle x^2, x - \frac{1}{2}\rangle}{\|x - \frac{1}{2}\|^2}\left(x - \frac{1}{2}\right) - \frac{\langle x^2, 1\rangle}{\|1\|^2}1 \\
&= x^2 - \frac{\frac{1}{12}}{\frac{1}{12}}\left(x - \frac{1}{2}\right) - \frac{1}{3} \\
&= x^2 - x + \frac{1}{6}
\end{aligned}
$$

Finally, we normalize and obtain

$$
\begin{aligned}
u_1 &= \frac{1}{\|1\|^2} = \frac{1}{\int_0^1 1\ dx} = 1 \\
u_2 &= \frac{x - \frac{1}{2}}{\|x - \frac{1}{2}\|} = \frac{x - \frac{1}{2}}{\sqrt{\frac{1}{12}}} = 2\sqrt{3}(x - \tfrac{1}{2}) \\
u_3 &= \frac{x^2 - x + \frac{1}{6}}{\|x^2 - x + \frac{1}{6}\|} = \frac{x^2 - x + \frac{1}{6}}{\sqrt{\int_0^1 (x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{1}{3}x + \frac{1}{36})dx}} \\
&= 6\sqrt{5}(x^2 - x + \tfrac{1}{6})
\end{aligned}
$$

We still have not verified that $\langle f, g\rangle = \int_a^b f(x)g(x)w(x)dx$ is an inner product on $C[a, b]$. We do that now, by showing the four axioms for an inner product hold.

1.
$$\begin{aligned}\langle f, g\rangle &= \int_a^b f(x)g(x)w(x)\ dx \\ &= \int_a^b g(x)f(x)w(x)\ dx = \langle g, f\rangle\end{aligned}$$

2.
$$\begin{aligned}\langle cf, g\rangle &= \int_a^b cf(x)g(x)w(x)\ dx \\ &= c\int_a^b f(x)g(x)w(x)\ dx = c\langle f, g\rangle\end{aligned}$$

3.
$$\begin{aligned}\langle f+h, g\rangle &= \int_a^b [f(x)+h(x)]g(x)w(x)\ dx \\ &= \int_a^b [f(x)g(x)w(x)+h(x)g(x)w(x)]\ dx \\ &= \int_a^b f(x)g(x)w(x)\ dx + \int_a^b h(x)g(x)w(x)dx \\ &= \langle f, g\rangle + \langle h, g\rangle\end{aligned}$$

4. **Positivity**
$$\langle f, f\rangle = \int_a^b [f(x)]^2 w(x)\ dx$$

Since $w > 0$ and $f^2(x) \geq 0$ on $[a, b]$, $\int_a^b [f(x)]^2 w(x)dx \geq 0$. Thus $\langle f, f\rangle \geq 0$. The definite integral of a nonnegative continuous function is zero if and only if the function is identically zero. Therefore, $\langle f, f\rangle = 0$ if and only if $f(x) \equiv 0$ on $[a, b]$.

Although the vector space structure of spaces such as $C[a, b]$ is not absolutely necessary in beginning calculus courses, it is an important tool in advanced calculus and advanced engineering mathematics. In particular $C[-\pi, \pi]$ has an infinite orthonormal basis

$$\mathcal{O} = \left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\sin x, \frac{1}{\sqrt{\pi}}\cos x, \frac{1}{\sqrt{\pi}}\sin 2x, \frac{1}{\sqrt{\pi}}\cos 2x, \dots\right\}.$$

These wave-type functions can then be used to represent functions in $C[-\pi, \pi]$ in a definite way. This fact is the basis for Fourier analysis, which is successful in analyzing problems in heat conduction, wave propagation, and electrostatics.

PROBLEMS 3.9

1. Let $\langle f, g\rangle = \int_0^1 f(x)g(x)\ dx$ n $C[0,1]$. For the given pairs of functions calculate $\langle f, g\rangle$.

   (**a**) $f(x) = x, g(x) = \sqrt{x}$

   (**b**) $f(x) = \sin \pi x, g(x) = \cos \pi x$

   (**c**) $f(x) = e^x, g(x) = e^{-x}$

   (**d**) $f(x) = x^2, g(x) = 1 - x$

2. Let $\langle f, g\rangle = \int_0^1 f(x)g(x)x^2\ dx$ in $C[0,1]$. For the given pairs of functions calculate $\langle f, g\rangle$.

   (**a**) $f(x) = 1, g(x) = x$

   (**b**) $f(x) = 1, g(x) = \sqrt{x}$

   (**c**) $f(x) = 1, g(x) = 1/(x^3 + 1)$

3. For the inner product in Prob. 1, calculate $\|f\|$ for $f(x) = x$.

4. For the inner product in Prob. 2, calculate $\|f\|$ for $f(x) = x$.

5. In $C^1[a,b]$ define

$$\langle f, g\rangle = \int_a^b [f(x)g(x) + f'(x)g'(x)]\ dx$$

   Show that this is an inner product.

6. In $C^1[a,b]$ define $\langle f, g\rangle = \int_a^b f'(x)g'(x)\ dx$. Show that this is **not** an inner product.

7. In $C^1(\mathbb{R})$, determine linear independence or dependence of the following sets of functions.

   (**a**) $S = \{1, x\}$

   (**b**) $S = \{e^x, e^{-x}\}$

   (**c**) $S = \{\sin x, 2\sin x\}$

8. Show that $\{e^x, e^{2x}\}$ is a linearly independent set in $C(\mathbb{R})$ by considering $c_1e^x + c_2e^{2x} = 0$ and putting $x = 0$ and $x = 1$.

9. Why can the wronskian not be used to determine the linear dependence or independence of $S = \{x^{2/3}, x^2\}$ in $C^1(\mathbb{R})$? Use the direct method to determine linear independence.

10. Recall the Cauchy-Schwarz and triangle inequalities

$$|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \ \|\mathbf{y}\| \quad \text{and} \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Using the inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)\ dx$ in $C[0, 1]$, rewrite the inequalities above, replacing $\mathbf{x}$ and $\mathbf{y}$ by f and g, respectively.

11. Using the inner product in Prob. 1, find the angle between x and x^3. [**Remember**: If $\mathbf{u}$ and $\mathbf{v}$ are vectors, then $\cos\theta = \langle \mathbf{u}, \mathbf{v} \rangle / (\|\mathbf{u}\| \ \|\mathbf{v}\|)$.]

12. Consider $V = \text{span}\ \{1, x, x^2\}$ in $C[-1, 1]$ with inner product $\langle f, g \rangle = \int_{-1}^1 f(x)g(x)\ dx$. Find an orthonormal basis for V. The basis consists of the first three **Legendre polynomials**.

13. Show that the set of solutions to $y'' = y$ forms a subspace for $C(\mathbb{R})$. Show that the dimension is at least two.

14. Consider the subspace W of $C(0, +\infty)$ consisting of all polynomials. Define the inner product $\langle f, g \rangle = \int_0^\infty f(x)g(x)e^{-x}\ dx$. let $V = \text{span}\ \{1, x\}$. Find an orthonormal basis for V. The basis consists of the first two **Laguerre polynomials**.

SUMMARY

Vector spaces are one of the most important structures in applied mathematics; they were defined and developed in this chapter. Some, such as E^n, are direct generalizations of two- and three-space, while others such as vector spaces of polynomials, functions, or matrices are generated by problems in calculus, differential equations, and applied mathematics. **Regardless of their appearance, finitely generated vector spaces have common attributes** such as the **defining axioms, subspace structure, bases, and dimension**.

The **basis problem** was posed as one of the fundamental problems of linear algebra. Its solution becomes extremely important in Chap. 5 when diagonalization of matrices is discussed. At this point its importance lies in analysis of applied problems (analysis of electron spin and resolution of forces into tangential and normal directions are two examples).

The **fundamental theorem** of this chapter guaranteed the **existence of a basis**[5] **for any finitely generated vector space** V. In fact, the proof demonstrated that a basis could be "built" from any set of vectors that spans V. For a given basis, the representation of a vector $\mathbf{x}$ in V is unique, and the coefficients in the linear combination of basis vectors for $\mathbf{x}$ are called the **coordinates** of $\mathbf{x}$ inn the basis. The development of these concepts depended heavily on the concepts of **linear dependence** and **linear independence**. **Orthonormal bases** have some "nice" properties; we exploit them in Chap. 5.

Bases are not unique. If S and T are two bases for a vector space V and $\mathbf{x}$ is in V, then the coordinates of $\mathbf{x}$ in S and in T are related by the **transition matrix**. Transition matrices are used in Chap. 4 to aid in the representations of important functions, called linear transformations.

In calculus, the structure of the real numbers had to be developed before the concept of a function could be defined. Accordingly, because the linear functions we study have domains and ranges which are vector spaces, we are now ready to proceed to the definition of linear transformations in Chap. 4.

ADDITIONAL PROBLEMS

1. Let $V = \{(x, y, z)$ in E^3 with $3x - y + z \geq 0\}$. Is V a vector space?

2. Let $V = \{(x, y, z)$ in E^3 with $x - y + z > 0\}$. Is V a vector space?

3. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a complex vector space V. Is $\{i\mathbf{v}_1, i\mathbf{v}_2, \ldots, i\mathbf{v}_n\}$ a basis for V?

4. Let $\{\mathbf{v}_1, \mathbf{v}_2\}$ be a basis for E^2. Let $A_{2\times 2}$ be a real matrix. Is $\{A\mathbf{v}_1, A\mathbf{v}_2\}$ a basis for E^2? Does it make any difference whether A is singular?

5. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a complex vector space V. Is $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, \ldots, c_n\mathbf{v}_n\}$ a basis for V? Assume $c_k \neq 0$ for $k = 1, 2, \ldots, n$.

6. If $\{(a, b), (c, d)\}$ is to be an orthonormal basis for E^2, what are c and d in terms of a and b? What must be true about

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

[5] Where the generating set contains nonzero vectors only.

7. Let A be an $m \times n$ matrix, $m \neq n$. Leg $P = A(A^T A)^{-1} A^T$. Show that P is symmetric and that $P^2 = P$.

8. Find a first column that will make

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

an orthogonal matrix.

9. Is the set of all orthogonal $n \times n$ matrices a subspace of $\mathcal{M}_{nn}$?

10. Let A be a fixed $n \times n$ matrix. Let V be the set of all matrices of the form $P^{-1}AP$, where P runs through the set of all $n \times n$ invertible matrices. Is V a subspace of $\mathcal{M}_{nn}$?

11. Let V be the set of solutions of $A_{n\times n} X_{n\times 1} = cX_{n\times 1}$, where c is a complex number. Give V the operations of $\mathcal{C}_{n1}$. Is V a vector space?

12. Show that if A and B are $n \times n$ hermitian matrices, then $i(AB - BA)$ is hermitian.

13. Approximation of a given function by other functions is a common tool of applied mathematics. Let the given function be $g(x)$, x in $[a, b]$. To carry out the approximation, one must set up a finite-dimensional vector space V of functions defined on $[a, b]$ and make sure the vector space has an inner product (which generates a norm). The function g must have $\|g\|$ a finite number. Then if $S = \{f_1, f_2, \ldots, f_n\}$ is a basis for V, we try to find constants $c_1, c_2, \ldots, x_n$ so that

$$\|(c_1 f_1 + c_2 f_2 + \cdots + c_n f_n) - g\|$$

is as small as possible. This process of minimization is simplified if S is an orthonormal basis, for then the best approximation is

$$\langle g, f_1 \rangle f_1 + \cdots + \langle g, f_n \rangle f_n$$

Approximate $g(x) = e^x$ by a quadratic polynomial. Use the vector space $\mathcal{P}_2$ with inner product defined by

$$\langle a_1 x^2 + a_2 x + a_3, b_1 x^2 + b_2 x + b_3 \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3$$

14. Suppose in a real vector space V with basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ we define an inner product for which the basis is not orthogonal. Define $G_{n\times n} = (\langle \mathbf{v}_i, \mathbf{v}_j \rangle)_{n\times n}$. Show that $G \neq I$. Show that G is symmetric. In mathematical physics, G is called the **metric** of the vector space.

15. Regarding Prob. 14, show for vectors $\mathbf{x}$ and $\mathbf{y}$ in V, with $\mathbf{x} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ and $\mathbf{y} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$, that

$$\langle \mathbf{x}, \mathbf{y} \rangle = (\mathbf{x})_S{}^T G(\mathbf{y})_S$$

16. Use the result of Prob. 15 to show that for any nonzero vector X in E^n

$$X^T G X > 0$$

This means that the metric of a vector space is a **positive definite** matrix.