3.3 Basis and dimension

The matrix-vector product $\mathbf{Ax}$ is equivalent to a linear combination of the columns of the matrix $\mathbf{A}$. If $\mathbf{A}$ has columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, then

$$\mathbf{Ax} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n.$$

The reader should write out a specific example if this is not clear (see Exercise 1). The quantities $x_1, x_2, \ldots, x_n$ are scalars and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are vectors; an expression such as $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + cldots + x_n\mathbf{v}_n$ is called a *linear combination* of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ because the vectors are combined using the linear operations of addition and scalar multiplication.

When $\mathbf{A} \in \mathbf{R}^{n\times n}$ is nonsingular, each $\mathbf{b} \in \mathbf{R}^n$ can be written in a unique way as a linear combination of the columns of $\mathbf{A}$ (that is, the equation $\mathbf{Ax} = \mathbf{b}$ has a unique solution). The following definition is related.

Definition 3.23. *Let V be a vector space, and suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are vectors in V with the property that each $\mathbf{v} \in V$ can be written in a unique way as a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is called a* basis *of V. Moreover, we say that n is the* dimension *of V.*

A vector space can have many different bases, but it can be shown that each contains the same number of vectors, so the concept of dimension is well-defined.

We now present several examples of bases.

Example 3.24. *The standard basis for $\mathbf{R}^n$ is $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, where every entry of $\mathbf{e}_j$ is zero except the jth, which is one. Then we obviously have, for any $\mathbf{x} \in \mathbf{R}^n$,*

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n,$$

and it is not hard to see that this representation is unique. For example, for $\mathbf{x} \in \mathbf{R}^3$,

$$\mathbf{x} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Example 3.25. *An alternate basis for $\mathbf{R}^3$ is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where*

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

It may not be obvious to the reader why one would want to use this basis instead of the much simpler basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. However, it is easy to check that

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0, \ \mathbf{v}_1 \cdot \mathbf{v}_3 = 0, \ \mathbf{v}_2 \cdot \mathbf{v}_3 = 0,$$

and this property makes the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ almost as easy to use as $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. We explore this topic in the next section.

Example 3.26. *The set $\mathcal{P}_n$ is the vector space of all polynomials of degree n or less (see Exercise 3.1.3). The standard basis is $\{1, x, x^2, \ldots, x^n\}$. To see that this is indeed a basis, we first note that every polynomial $p \in \mathcal{P}_n$ can be written as a linear combination of $1, x, x^2, \ldots, x^n$:*

$$p(x) = c_0 \cdot 1 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

(this is just the definition of polynomial of degree n). Showing that this representation is unique is a little subtle. If we also had

$$p(x) = d_0 \cdot 1 + d_1 x + d_2 x^2 + \cdots + d_n x^n,$$

then, subtraction would yield

$$(c_0 - d_0) + (c_1 - d_1)x + \cdots + (c_n - d_n)x^n = 0$$

for every x. However, a nonzero polynomial of degree n can have at most n roots, so it must be the case that $(c_0 - d_0) + (c_1 - d_1)x + \cdots + (c_n - d_n)x^n$ is the zero polynomial. That is, $c_0 = d_0, c_1 = d_1, \ldots, c_n = d_n$ must hold.

Example 3.27. *An alternate basis for $\mathcal{P}_2$ is*

$$\left\{1, x - \frac{1}{2}, x^2 - x + \frac{1}{6}\right\}$$

(the advantage of this basis will be discussed in Example 3.39 in the next section). To show that this is indeed a basis, we must show that, given any $p(x) = c_0 + c_1 x + c_2 x^2$, there is a unique choice of the scalars a_0, a_1, a_2 such that

$$a_0 \cdot 1 + a_1\left(x - \frac{1}{2}\right) + a_2\left(x^2 - x + \frac{1}{6}\right) = c_0 + c_1 x + c_2 x^2.$$

This equation is equivalent to the three linear equations

$$\begin{aligned} a_0 - \frac{1}{2}a_1 + \frac{1}{6}a_2 &= c_0, \\ a_1 - a_2 &= c_1, \\ a_2 &= c_2. \end{aligned}$$

The reader can easily verify that this system has a unique solution, regardless of the values of c_0, c_1, c_2.

Example 3.28. *Yet another basis for $\mathcal{P}_2$ is $\{L_1, L_2, L_3\}$, where*

$$\begin{aligned} L_1(x) &= 2\left(x - \frac{1}{2}\right)(x - 1), \\ L_2(x) &= -4x(x - 1), \\ L_3(x) &= 2x\left(x - \frac{1}{2}\right). \end{aligned}$$

If we write $x_1 = 0$, $x_2 = 1/2$, and $x_3 = 1$, then the property

$$L_i(x_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases} \tag{3.9}$$

holds. From this property, the properties of a basis can be verified (see Exercise 5).

There are two essential properties of a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of a vector space V. First, every vector in V can be represented as a linear combination of the basis vectors. Second, this representation is unique. The following two definitions provide concise ways to express these two properties.

Definition 3.29. *Let V be a vector space, and suppose $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a collection of vectors in V. The* span *of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is the set of all linear combinations of these vectors:*

$$span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \{\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n \; : \; \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbf{R}\}.$$

Thus, one of the properties of a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of a vector space V is that

$$V = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}.$$

The reader should also be aware that, for *any* vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in a vector space V, span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a subspace of V (possibly the entire space V, as in the case of a basis).

Definition 3.30. *A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is called* linearly independent *if the only scalars $c_1, c_2, \ldots, c_n$ satisfying*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = 0$$

are $c_1 = c_2 = \cdots = c_n = 0$.

It can be shown that the uniqueness part of the definition of a basis is equivalent to the linear independence of the basis vectors. Therefore, a basis for a vector space is a linearly independent spanning set.

A third quality of a basis is the number of vectors in it—the dimension of the vector space. It can be shown that any two of these properties imply the third. That is, if V has dimension n, then any two of the following statements about $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ imply the third:

- $k = n$;
- $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent;
- $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ spans V.

Thus if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is known to satisfy two of the above properties, then it is a basis for V.

Before leaving the topic of basis, we wish to remind the reader of the fact indicated in the opening paragraphs of this section, which is so fundamental that we express it formally as a theorem.

Theorem 3.31. *Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is nonsingular if and only if the columns of $\mathbf{A}$ form a basis for $\mathbf{R}^n$.*

Thus, when A is nonsingular, its columns form a basis for $\mathbf{R}^n$, and *solving $\mathbf{Ax} = \mathbf{b}$ is equivalent to finding the weights that express $\mathbf{b}$ as a linear combination of this basis.* This fact answers the following important question.[9] Suppose we have a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for $\mathbf{R}^n$ and a vector $\mathbf{b} \in \mathbf{R}^n$. Then, of course, $\mathbf{b}$ is a linear combination of the basis vectors. How do we find the weights in this linear combination? How expensive is it to do so (that is, how much work is required)?

To find the scalars $x_1, x_2, \ldots, x_n$ in the equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \mathbf{b},$$

we define[10]

$$\mathbf{A} = [\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n]$$

and solve

$$\mathbf{Ax} = \mathbf{b}$$

via Gaussian elimination. The expense of computing $\mathbf{x}$ can be measured by counting the number of arithmetic operations—the number of additions, subtractions, multiplications, and divisions—required. The total number of operations required to solve $\mathbf{Ax} = \mathbf{b}$ is a polynomial in n, and it is convenient to report just the leading term in the polynomial, which can be shown to be

$$\frac{2}{3}n^3$$

(the lower-order terms are not very significant when n is large). We usually express this saying that the operation count is

$$O\left(\frac{2}{3}n^3\right)$$

("on the *order* of $(2/3)n^3$").

In the next section, we discuss a certain special type of basis for which it is much easier to express a vector in terms of the basis.

[9] If the importance of this question is not apparent to the reader at this point, it will be after he or she reads the next two sections.

[10] This notation means that A is the matrix whose columns are the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Exercises

1. (a) Let

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

Compute both $\mathbf{Ax}$ and

$$2\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 4 \\ -3 \end{bmatrix},$$

and verify that they are equal.

(b) Let $\mathbf{A} \in \mathbf{R}^{n\times n}$ and $\mathbf{x} \in \mathbf{R}^n$, and suppose the columns of $\mathbf{A}$ are

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbf{R}^n,$$

so that the (i, j)-entry of $\mathbf{A}$ is $(\mathbf{v}_j)_i$. Compute both $(\mathbf{Ax})_i$ and $(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n)_i$, and verify that they are equal.

2. Is

$$\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}\right\}$$

a basis for $\mathbf{R}^3$? (Hint: As explained in the last paragraphs of this section, the three given vectors form a basis for $\mathbf{R}^n$ if and only if $\mathbf{Ax} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbf{R}^n$, where $\mathbf{A}$ is the 3×3 matrix whose columns are the three given vectors.)

3. Is

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

a basis for $\mathbf{R}^3$? (See the hint for the previous exercise.)

4. Show that

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

is a basis for $\mathcal{P}_2$, the space of polynomials of degree 2 or less. (Hint: Verify directly that the definition holds.)

5. Show that $\{L_1, L_2, L_3\}$, defined in Example 3.28, is a basis for $\mathcal{P}_2$. (Hint: Use (3.9) to show that

$$p(x) = p(x_1)L_1(x) + p(x_2)L_2(x) + p(x_3)L_3(x)$$

holds for every $p \in \mathcal{P}_2$.)

6. Let V be the space of all continuous, complex-valued functions defined on the real line:

$$V = \{f : \mathbf{R} \to \mathbf{C} \ : \ f \text{ is continuous}\}.$$

Define W to be the subspace of V spanned by e^{ix} and e^{-ix}, where $i = \sqrt{-1}$. Show that $\{\cos(x), \sin(x)\}$ is another basis for W. (Hint: Use Euler's formula: $e^{i\theta} = \cos(\theta) + i\sin(\theta)$.)

3.4 Orthogonal bases and projections

At the end of the last section, we discussed the question of expressing a vector in terms of a given basis. This question is important for the following reason, which we can only describe in general terms at the moment: Many problems that are posed in vector spaces admit a special basis, in terms of which the problem is easy to solve. That is, for many problems, there exists a special basis with the property that if all vectors are expressed in terms of that basis, then a very simple calculation will produce the final solution. For this reason, it is important to be able to take a vector (perhaps expressed in terms of a standard basis) and express it in terms of a different basis. In the latter part of this section, we will study one type of problem for which it is advantageous to use a special basis, and we will discuss another such problem in the next section.

It is quite easy to express a vector in terms of a basis if that basis is *orthogonal*. We wish to describe the concept of an orthogonal basis and show some important examples. Before we can do so, we must introduce the idea of an inner product, which is a generalization of the Euclidean dot product.

The dot product plays a special role in the geometry of $\mathbf{R}^2$ and $\mathbf{R}^3$. The reason for this is the fact that two vectors $\mathbf{x}, \mathbf{y}$ in $\mathbf{R}^2$ or $\mathbf{R}^3$ are perpendicular if and only if

$$\mathbf{x} \cdot \mathbf{y} = 0.$$

Indeed, one can show that

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta),$$

where θ is the angle between the two vectors (see Figure 3.2).

From elementary Euclidean geometry, we know that, if $\mathbf{x}$ and $\mathbf{y}$ are perpendicular, then

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

(the Pythagorean theorem). Using the dot product, we can give a purely algebraic proof of the Pythagorean theorem. By definition,

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}},$$

so

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x} \cdot \mathbf{x} + \mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} \\ &= \mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} \\ &= \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2. \end{aligned}$$

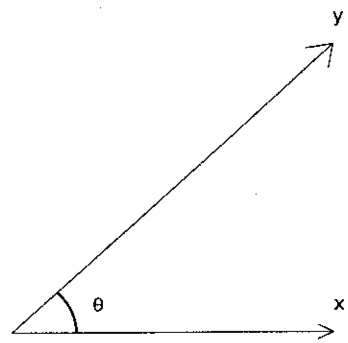

Figure 3.2. *The angle between two vectors.*

This calculation shows that

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

holds if and only if $\mathbf{x}\cdot\mathbf{y} = 0$.

Seen this way, the Pythagorean theorem is an algebraic property that can be deduced in $\mathbf{R}^n$, $n > 3$, even though in those spaces we cannot visualize vectors or what it means for vectors to be perpendicular. We prefer to use the word *orthogonal* instead of perpendicular: Vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}^n$ are orthogonal if $\mathbf{x}\cdot\mathbf{y} = 0$.

In the course of solving differential equations, we deal with function spaces in addition to Euclidean spaces, and our methods are heavily dependent on the existence of an *inner product*—the analogue of the dot product in more general vector spaces. Here is the definition:

Definition 3.32. *Let V be a real vector space. A (real)* inner product *on V is a function, usually denoted $(\cdot,\cdot)$ or $(\cdot,\cdot)_V$, taking two vectors from V and producing a real number. This function must satisfy the following three properties:*

1. $(\mathbf{u},\mathbf{v}) = (\mathbf{v},\mathbf{u})$ *for all vectors* $\mathbf{u}$ *and* $\mathbf{v}$;
2. $(\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{w}) = \alpha(\mathbf{u},\mathbf{w}) + \beta(\mathbf{v},\mathbf{w})$ *and* $(\mathbf{w},\alpha\mathbf{u}+\beta\mathbf{v}) = \alpha(\mathbf{w},\mathbf{u}) + \beta(\mathbf{w},\mathbf{v})$ *for all vectors* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$, *and all real numbers* α *and* β;
3. $(\mathbf{u},\mathbf{u}) \geq 0$ *for all vectors* $\mathbf{u}$, *and* $(\mathbf{u},\mathbf{u}) = 0$ *if and only if* $\mathbf{u}$ *is the zero vector.*

It should be easy to check that these properties hold for the ordinary dot product on Euclidean n-space.

Given an *inner product space* (a vector space with an inner product), we define orthogonality just as in Euclidean space: two vectors are orthogonal if and only if their dot product is zero. It can then be shown that the Pythagorean theorem holds (see Exercise 3).

An *orthogonal basis* for an inner product space V is a basis $\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ with the property that

$$i \neq j \Rightarrow (\mathbf{v}_i,\mathbf{v}_j) = 0$$

(that is, every vector in the basis is orthogonal to every other vector in the basis). We now demonstrate the first special property of an orthogonal basis. Suppose $\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n\}$ is an orthogonal basis for an inner product space V and $\mathbf{x}$ is any vector in V. Then there exist scalars $\alpha_1,\alpha_2,\ldots,\alpha_n$ such that

$$\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n. \tag{3.10}$$

To deduce the value of α_i, we take the inner product of both sides of (3.10) with $\mathbf{v}_i$:

$$\begin{aligned}(\mathbf{v}_i,\mathbf{x}) &= (\mathbf{v}_i,\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n)\\ &= \alpha_1(\mathbf{v}_i,\mathbf{v}_1) + \alpha_2(\mathbf{v}_i,\mathbf{v}_2) + \cdots + \alpha_n(\mathbf{v}_i,\mathbf{v}_n)\\ &= \alpha_i(\mathbf{v}_i,\mathbf{v}_i).\end{aligned}$$

The last step follows from the fact that every inner product $(\mathbf{v}_i,\mathbf{v}_j)$ vanishes except $(\mathbf{v}_i,\mathbf{v}_i)$. We then obtain

$$\alpha_i = \frac{(\mathbf{v}_i,\mathbf{x})}{(\mathbf{v}_i,\mathbf{v}_i)},\ i = 1,2,\ldots,n,$$

and so

$$\mathbf{x} = \frac{(\mathbf{v}_1,\mathbf{x})}{(\mathbf{v}_1,\mathbf{v}_1)}\mathbf{v}_1 + \frac{(\mathbf{v}_2,\mathbf{x})}{(\mathbf{v}_2,\mathbf{v}_2)}\mathbf{v}_2 + \cdots + \frac{(\mathbf{v}_n,\mathbf{x})}{(\mathbf{v}_n,\mathbf{v}_n)}\mathbf{v}_n. \tag{3.11}$$

This formula shows that it is easy to express a vector in terms of an orthogonal basis. Assuming that we compute $(\mathbf{v}_1,\mathbf{v}_1),(\mathbf{v}_2,\mathbf{v}_2),\ldots,(\mathbf{v}_n,\mathbf{v}_n)$ once and for all, it requires just n inner products to find the weights in the linear combination. In the case of Euclidean n-vectors, a dot product requires $2n-1$ arithmetic operations (n multiplications and $n-1$ additions), so the total cost is just

$$O(2n^2).$$

If n is large, this is *much* less costly than the $O(2n^3/3)$ operations required for a nonorthogonal basis. We also remark that if the basis is *orthonormal*—each basis vector is normalized to have length one—then (3.11) simplifies to

$$\mathbf{x} = (\mathbf{v}_1,\mathbf{x})\mathbf{v}_1 + (\mathbf{v}_2,\mathbf{x})\mathbf{v}_2 + \cdots + (\mathbf{v}_n,\mathbf{x})\mathbf{v}_n. \tag{3.12}$$

Example 3.33. *The basis $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\}$ for $\mathbf{R}^3$, where*

$$\mathbf{v}_1 = \begin{bmatrix}\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}\end{bmatrix},\ \mathbf{v}_2 = \begin{bmatrix}\frac{1}{\sqrt{2}}\\ 0\\ -\frac{1}{\sqrt{2}}\end{bmatrix},\ \mathbf{v}_3 = \begin{bmatrix}\frac{1}{\sqrt{6}}\\ -\frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{6}}\end{bmatrix},$$

is orthonormal, as can be verified directly. If

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

then

$$\begin{aligned} \mathbf{x} &= (\mathbf{v}_1 \cdot \mathbf{x})\mathbf{v}_1 + (\mathbf{v}_2 \cdot \mathbf{x})\mathbf{v}_2 + (\mathbf{v}_3 \cdot \mathbf{x})\mathbf{v}_3 \\ &= \frac{6}{\sqrt{3}}\mathbf{v}_1 - \frac{2}{\sqrt{2}}\mathbf{v}_2 + 0\mathbf{v}_3. \end{aligned}$$

3.4.1 The L^2 inner product

We have seen that functions can be regarded as vectors, at least in a formal sense: functions can be added together and multiplied by scalars. (See Example 3.4 in Section 3.1.) We will now show more directly that functions are not so different from Euclidean vectors. In the process, we show that a suitable inner product can be defined for functions.

Suppose we have a function $g \in C[a,b]$—a continuous function defined on the interval $[a,b]$. By sampling g on a grid, we can produce a vector that approximates the function g. Let $x_i = a + i\Delta x$, $\Delta x = (b-a)/N$, and define a vector $G \in \mathbf{R}^N$ by

$$G_i = g(x_i),\ i = 0, 1, \ldots, N-1.$$

Then G can be regarded as an approximation to g (see Figure 3.3). Given another function $f(x)$ and the corresponding vector $F \in \mathbf{R}^N$, we have

$$\begin{aligned} F \cdot G &= \sum_{i=0}^{N-1} F_i G_i \\ &= \sum_{i=0}^{N-1} f(x_i) g(x_i). \end{aligned}$$

Refining the discretization (increasing N) leads to a sampled function that obviously represents the original function more accurately. Therefore, we ask: What happens to $F \cdot G$ as $N \to \infty$? The dot product

$$F \cdot G = \sum_{i=0}^{N-1} f(x_i)\, g(x_i)$$

does not converge to any value as $N \to \infty$, but a simple modification induces convergence. We replace the ordinary dot product by the following scaled dot product, for which we introduce a new notation:

$$(F, G) = \sum_{i=0}^{N-1} F_i G_i \Delta x.$$

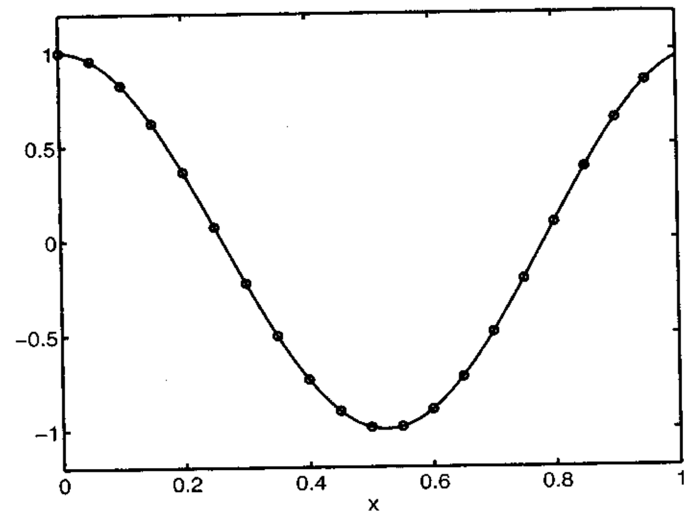

Figure 3.3. *Approximating a function $g(x)$ by a vector $G \in \mathbf{R}^N$.*

Then, when F and G are sampled functions as above, we have

$$(F,G) = \sum_{i=0}^{N-1} f(x_i) g(x_i) \Delta x \to \int_a^b f(x) g(x)\, dx \quad \text{as } N \to \infty.$$

Based on this observation, we argue that a natural inner product $(\cdot,\cdot)$ on $C[a,b]$ is

$$(f,g) = \int_a^b f(x) g(x)\, dx. \tag{3.13}$$

Just as the dot product defines a norm on Euclidean n-space ($\|x\| = \sqrt{x \cdot x}$), so the inner product (3.13) defines a norm for functions:

$$\|f\| = \sqrt{(f,f)} = \sqrt{\int_0^\ell |f(x)|^2\, dx}. \tag{3.14}$$

For completeness, we give the definition of norm. Norms measure the size or magnitude of vectors, and the definition is intended to describe abstractly the properties that any reasonable notion of size ought to have.

Definition 3.34. *Let V be a vector space. A* norm *on V is a real-valued function with domain V, usually denoted by $\|\cdot\|$ or $\|\cdot\|_V$, and satisfying the following properties:*