6B Orthonormal Bases

Orthonormal Lists and the Gram–Schmidt Procedure

6.22 definition: *orthonormal*

- A list of vectors is called *orthonormal* if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.
- In other words, a list $e_1, \dots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

for all $j, k \in \{1, \dots, m\}$.

6.23 example: *orthonormal lists*

(a) The standard basis of $\mathbf{F}^n$ is an orthonormal list.

(b) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ is an orthonormal list in $\mathbf{F}^3$.

(c) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ is an orthonormal list in $\mathbf{F}^3$.

(d) Suppose n is a positive integer. Then, as Exercise 4 asks you to verify,

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

The orthonormal list above is often used for modeling periodic phenomena, such as tides.

(e) Suppose we make $\mathcal{P}_2(\mathbf{R})$ into an inner product space using the inner product given by

$$\langle p, q \rangle = \int_{-1}^{1} pq$$

for all $p, q \in \mathcal{P}_2(\mathbf{R})$. The standard basis $1, x, x^2$ of $\mathcal{P}_2(\mathbf{R})$ is not an orthonormal list because the vectors in that list do not have norm 1. Dividing each vector by its norm gives the list $1/\sqrt{2}, \sqrt{3/2}x, \sqrt{5/2}x^2$, in which each vector has norm 1, and the second vector is orthogonal to the first and third vectors. However, the first and third vectors are not orthogonal. Thus this is not an orthonormal list. Soon we will see how to construct an orthonormal list from the standard basis $1, x, x^2$ (see Example 6.34).

Orthonormal lists are particularly easy to work with, as illustrated by the next result.

6.24 *norm of an orthonormal linear combination*

Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. Then

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \ldots, a_m \in \mathbf{F}$.

Proof Because each e_k has norm 1, this follows from repeated applications of the Pythagorean theorem (6.12). ∎

The result above has the following important corollary.

6.25 *orthonormal lists are linearly independent*

Every orthonormal list of vectors is linearly independent.

Proof Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V and $a_1, \ldots, a_m \in \mathbf{F}$ are such that

$$a_1 e_1 + \cdots + a_m e_m = 0.$$

Then $|a_1|^2 + \cdots + |a_m|^2 = 0$ (by 6.24), which means that all the a_k's are 0. Thus $e_1, \ldots, e_m$ is linearly independent. ∎

Now we come to an important inequality.

6.26 *Bessel's inequality*

Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. If $v \in V$ then

$$\bigl|\langle v, e_1\rangle\bigr|^2 + \cdots + \bigl|\langle v, e_m\rangle\bigr|^2 \leq \|v\|^2.$$

Proof Suppose $v \in V$. Then

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}.$$

Let u and w be defined as in the equation above. If $k \in \{1, \ldots, m\}$, then $\langle w, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle\langle e_k, e_k\rangle = 0$. This implies that $\langle w, u\rangle = 0$. The Pythagorean theorem now implies that

$$\begin{aligned} \|v\|^2 &= \|u\|^2 + \|w\|^2 \\ &\geq \|u\|^2 \\ &= \bigl|\langle v, e_1\rangle\bigr|^2 + \cdots + \bigl|\langle v, e_m\rangle\bigr|^2, \end{aligned}$$

where the last line comes from 6.24. ∎

The next definition introduces one of the most useful concepts in the study of inner product spaces.

6.27 definition: *orthonormal basis*

An *orthonormal basis* of V is an orthonormal list of vectors in V that is also a basis of V.

For example, the standard basis is an orthonormal basis of $\mathbf{F}^n$.

6.28 *orthonormal lists of the right length are orthonormal bases*

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V of length $\dim V$ is an orthonormal basis of V.

Proof By 6.25, every orthonormal list of vectors in V is linearly independent. Thus every such list of the right length is a basis—see 2.38. ■

6.29 example: *an orthonormal basis of* $\mathbf{F}^4$

As mentioned above, the standard basis is an orthonormal basis of $\mathbf{F}^4$. We now show that

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

is also an orthonormal basis of $\mathbf{F}^4$.

We have

$$\left\|\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right)\right\| = \sqrt{\frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{2^2}} = 1.$$

Similarly, the other three vectors in the list above also have norm 1.

Note that

$$\left\langle \left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right) \right\rangle = \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2} \cdot \left(-\tfrac{1}{2}\right) + \tfrac{1}{2} \cdot \left(-\tfrac{1}{2}\right) = 0.$$

Similarly, the inner product of any two distinct vectors in the list above also equals 0.

Thus the list above is orthonormal. Because we have an orthonormal list of length four in the four-dimensional vector space $\mathbf{F}^4$, this list is an orthonormal basis of $\mathbf{F}^4$ (by 6.28).

In general, given a basis $e_1, \dots, e_n$ of V and a vector $v \in V$, we know that there is some choice of scalars $a_1, \dots, a_n \in \mathbf{F}$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

Computing the numbers $a_1, \dots, a_n$ that satisfy the equation above can be a long computation for an arbitrary basis of V. The next result shows, however, that this is easy for an orthonormal basis—just take $a_k = \langle v, e_k \rangle$.

Notice how the next result makes each inner product space of dimension n behave like $\mathbf{F}^n$, with the role of the coordinates of a vector in $\mathbf{F}^n$ played by $\langle v, e_1 \rangle, \dots, \langle v, e_n \rangle$.

The formula below for $\|v\|$ is called Parseval's identity. It was published in 1799 in the context of Fourier series.

6.30 *writing a vector as a linear combination of an orthonormal basis*

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $u, v \in V$. Then

(a) $v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$;

(b) $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$;

(c) $\langle u, v \rangle = \langle u, e_1 \rangle \overline{\langle v, e_1 \rangle} + \cdots + \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$.

Proof Because $e_1, \dots, e_n$ is a basis of V, there exist scalars $a_1, \dots, a_n$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

Because $e_1, \dots, e_n$ is orthonormal, taking the inner product of both sides of this equation with e_k gives $\langle v, e_k \rangle = a_k$. Thus (a) holds.

Now (b) follows immediately from (a) and 6.24.

Take the inner product of u with each side of (a) and then get (c) by using conjugate linearity [6.6(d) and 6.6(e)] in the second slot of the inner product. ■

6.31 example: *finding coefficients for a linear combination*

Suppose we want to write the vector $(1, 2, 4, 7) \in \mathbf{F}^4$ as a linear combination of the orthonormal basis

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

of $\mathbf{F}^4$ from Example 6.29. Instead of solving a system of four linear equations in four unknowns, as typically would be required if we were working with a nonorthonormal basis, we simply evaluate four inner products and use 6.30(a), getting that $(1, 2, 4, 7)$ equals

$$7\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right) - 4\left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right) + \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right) + 2\left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right).$$

Now that we understand the usefulness of orthonormal bases, how do we go about finding them? For example, does $\mathcal{P}_m(\mathbf{R})$ with inner product as in 6.3(c) have an orthonormal basis? The next result will lead to answers to these questions.

The algorithm used in the next proof is called the *Gram–Schmidt procedure*. It gives a method for turning a linearly independent list into an orthonormal list with the same span as the original list.

Jørgen Gram (1850–1916) and Erhard Schmidt (1876–1959) popularized this algorithm that constructs orthonormal lists.

6.32 *Gram–Schmidt procedure*

Suppose $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Let $f_1 = v_1$. For $k = 2, \ldots, m$, define f_k inductively by

$$f_k = v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}.$$

For each $k = 1, \ldots, m$, let $e_k = \dfrac{f_k}{\|f_k\|}$. Then $e_1, \ldots, e_m$ is an orthonormal list of vectors in V such that

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$$

for each $k = 1, \ldots, m$.

Proof We will show by induction on k that the desired conclusion holds. To get started with $k = 1$, note that because $e_1 = f_1/\|f_1\|$, we have $\|e_1\| = 1$; also, $\operatorname{span}(v_1) = \operatorname{span}(e_1)$ because e_1 is a nonzero multiple of v_1.

Suppose $1 < k \leq m$ and the list $e_1, \ldots, e_{k-1}$ generated by 6.32 is an orthonormal list such that

6.33 $$\operatorname{span}(v_1, \ldots, v_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1}).$$

Because $v_1, \ldots, v_m$ is linearly independent, we have $v_k \notin \operatorname{span}(v_1, \ldots, v_{k-1})$. Thus $v_k \notin \operatorname{span}(e_1, \ldots, e_{k-1}) = \operatorname{span}(f_1, \ldots, f_{k-1})$, which implies that $f_k \neq 0$. Hence we are not dividing by 0 in the definition of e_k given in 6.32. Dividing a vector by its norm produces a new vector with norm 1; thus $\|e_k\| = 1$.

Let $j \in \{1, \ldots, k-1\}$. Then

$$\begin{aligned}
\langle e_k, e_j \rangle &= \frac{1}{\|f_k\| \, \|f_j\|} \langle f_k, f_j \rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \left\langle v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}, f_j \right\rangle \\
&= \frac{1}{\|f_k\| \, \|f_j\|} \big(\langle v_k, f_j \rangle - \langle v_k, f_j \rangle \big) \\
&= 0.
\end{aligned}$$

Thus $e_1, \ldots, e_k$ is an orthonormal list.

From the definition of e_k given in 6.32, we see that $v_k \in \operatorname{span}(e_1, \ldots, e_k)$. Combining this information with 6.33 shows that

$$\operatorname{span}(v_1, \ldots, v_k) \subseteq \operatorname{span}(e_1, \ldots, e_k).$$

Both lists above are linearly independent (the v's by hypothesis, and the e's by orthonormality and 6.25). Thus both subspaces above have dimension k, and hence they are equal, completing the induction step and thus completing the proof. ∎

6.34 example: *an orthonormal basis of* $\mathcal{P}_2(\mathbf{R})$

Suppose we make $\mathcal{P}_2(\mathbf{R})$ into an inner product space using the inner product given by

$$\langle p, q \rangle = \int_{-1}^{1} pq$$

for all $p, q \in \mathcal{P}_2(\mathbf{R})$. We know that $1, x, x^2$ is a basis of $\mathcal{P}_2(\mathbf{R})$, but it is not an orthonormal basis. We will find an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ by applying the Gram–Schmidt procedure with $v_1 = 1$, $v_2 = x$, and $v_3 = x^2$.

To get started, take $f_1 = v_1 = 1$. Thus $\|f_1\|^2 = \int_{-1}^{1} 1 = 2$. Hence the formula in 6.32 tells us that

$$f_2 = v_2 - \frac{\langle v_2, f_1 \rangle}{\|f_1\|^2} f_1 = x - \frac{\langle x, 1 \rangle}{\|f_1\|^2} = x,$$

where the last equality holds because $\langle x, 1 \rangle = \int_{-1}^{1} t\, dt = 0$.

The formula above for f_2 implies that $\|f_2\|^2 = \int_{-1}^{1} t^2\, dt = \frac{2}{3}$. Now the formula in 6.32 tells us that

$$f_3 = v_3 - \frac{\langle v_3, f_1 \rangle}{\|f_1\|^2} f_1 - \frac{\langle v_3, f_2 \rangle}{\|f_2\|^2} f_2 = x^2 - \tfrac{1}{2}\langle x^2, 1 \rangle - \tfrac{3}{2}\langle x^2, x \rangle x = x^2 - \tfrac{1}{3}.$$

The formula above for f_3 implies that

$$\|f_3\|^2 = \int_{-1}^{1} \left(t^2 - \tfrac{1}{3}\right)^2 dt = \int_{-1}^{1} \left(t^4 - \tfrac{2}{3}t^2 + \tfrac{1}{9}\right) dt = \tfrac{8}{45}.$$

Now dividing each of f_1, f_2, f_3 by its norm gives us the orthonormal list

$$\sqrt{\tfrac{1}{2}}, \sqrt{\tfrac{3}{2}}\, x, \sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right).$$

The orthonormal list above has length three, which is the dimension of $\mathcal{P}_2(\mathbf{R})$. Hence this orthonormal list is an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ [by 6.28].

Now we can answer the question about the existence of orthonormal bases.

6.35 *existence of orthonormal basis*

Every finite-dimensional inner product space has an orthonormal basis.

Proof Suppose V is finite-dimensional. Choose a basis of V. Apply the Gram–Schmidt procedure (6.32) to it, producing an orthonormal list of length $\dim V$. By 6.28, this orthonormal list is an orthonormal basis of V. ■

Sometimes we need to know not only that an orthonormal basis exists, but also that every orthonormal list can be extended to an orthonormal basis. In the next corollary, the Gram–Schmidt procedure shows that such an extension is always possible.

6.36 *every orthonormal list extends to an orthonormal basis*

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Proof Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. Then $e_1, \dots, e_m$ is linearly independent (by 6.25). Hence this list can be extended to a basis $e_1, \dots, e_m, v_1, \dots, v_n$ of V (see 2.32). Now apply the Gram–Schmidt procedure (6.32) to $e_1, \dots, e_m, v_1, \dots, v_n$, producing an orthonormal list

$$e_1, \dots, e_m, f_1, \dots, f_n;$$

here the formula given by the Gram–Schmidt procedure leaves the first m vectors unchanged because they are already orthonormal. The list above is an orthonormal basis of V by 6.28. ■

Recall that a matrix is called upper triangular if it looks like this:

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix},$$

where the 0 in the matrix above indicates that all entries below the diagonal equal 0, and asterisks are used to denote entries on and above the diagonal.

In the last chapter, we gave a necessary and sufficient condition for an operator to have an upper-triangular matrix with respect to some basis (see 5.44). Now that we are dealing with inner product spaces, we would like to know whether there exists an *orthonormal* basis with respect to which we have an upper-triangular matrix. The next result shows that the condition for an operator to have an upper-triangular matrix with respect to some orthonormal basis is the same as the condition to have an upper-triangular matrix with respect to an arbitrary basis.

6.37 *upper-triangular matrix with respect to some orthonormal basis*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \dots, \lambda_m \in \mathbf{F}$.

Proof Suppose T has an upper-triangular matrix with respect to some basis $v_1, \dots, v_n$ of V. Thus $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, \dots, n$ (see 5.39).

Apply the Gram–Schmidt procedure to $v_1, \dots, v_n$, producing an orthonormal basis $e_1, \dots, e_n$ of V. Because

$$\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k)$$

for each k (see 6.32), we conclude that $\operatorname{span}(e_1, \dots, e_k)$ is invariant under T for each $k = 1, \dots, n$. Thus, by 5.39, T has an upper-triangular matrix with respect to the orthonormal basis $e_1, \dots, e_n$. Now use 5.44 to complete the proof. ■

For complex vector spaces, the next result is an important application of the result above. See Exercise 20 for a version of Schur's theorem that applies simultaneously to more than one operator.

Issai Schur (1875–1941) published a proof of the next result in 1909.

6.38 *Schur's theorem*

Every operator on a finite-dimensional complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.

Proof The desired result follows from the second version of the fundamental theorem of algebra (4.13) and 6.37. ∎

Linear Functionals on Inner Product Spaces

Because linear maps into the scalar field $\mathbf{F}$ play a special role, we defined a special name for them and their vector space in Section 3F. Those definitions are repeated below in case you skipped Section 3F.

6.39 definition: *linear functional, dual space, V'*

- A *linear functional* on V is a linear map from V to $\mathbf{F}$.
- The *dual space* of V, denoted by V', is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V, \mathbf{F})$.

6.40 example: *linear functional on $\mathbf{F}^3$*

The function $\varphi\colon \mathbf{F}^3 \to \mathbf{F}$ defined by

$$\varphi(z_1, z_2, z_3) = 2z_1 - 5z_2 + z_3$$

is a linear functional on $\mathbf{F}^3$. We could write this linear functional in the form

$$\varphi(z) = \langle z, w \rangle$$

for every $z \in \mathbf{F}^3$, where $w = (2, -5, 1)$.

6.41 example: *linear functional on $\mathcal{P}_5(\mathbf{R})$*

The function $\varphi\colon \mathcal{P}_5(\mathbf{R}) \to \mathbf{R}$ defined by

$$\varphi(p) = \int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\, dt$$

is a linear functional on $\mathcal{P}_5(\mathbf{R})$.

The next result is named in honor of Frigyes Riesz (1880–1956), who proved several theorems early in the twentieth century that look very much like the result below.

If $v \in V$, then the map that sends u to $\langle u, v \rangle$ is a linear functional on V. The next result states that every linear functional on V is of this form. For example, we can take $v = (2, -5, 1)$ in Example 6.40.

Suppose we make the vector space $\mathcal{P}_5(\mathbf{R})$ into an inner product space by defining $\langle p, q \rangle = \int_{-1}^{1} pq$. Let φ be as in Example 6.41. It is not obvious that there exists $q \in \mathcal{P}_5(\mathbf{R})$ such that

$$\int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\, dt = \langle p, q \rangle$$

for every $p \in \mathcal{P}_5(\mathbf{R})$ [we cannot take $q(t) = \cos(\pi t)$ because that choice of q is not an element of $\mathcal{P}_5(\mathbf{R})$]. The next result tells us the somewhat surprising result that there indeed exists a polynomial $q \in \mathcal{P}_5(\mathbf{R})$ such that the equation above holds for all $p \in \mathcal{P}_5(\mathbf{R})$.

6.42 *Riesz representation theorem*

Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $v \in V$ such that

$$\varphi(u) = \langle u, v \rangle$$

for every $u \in V$.

Proof First we show that there exists a vector $v \in V$ such that $\varphi(u) = \langle u, v \rangle$ for every $u \in V$. Let $e_1, \ldots, e_n$ be an orthonormal basis of V. Then

$$\begin{aligned} \varphi(u) &= \varphi\big(\langle u, e_1 \rangle e_1 + \cdots + \langle u, e_n \rangle e_n\big) \\ &= \langle u, e_1 \rangle \varphi(e_1) + \cdots + \langle u, e_n \rangle \varphi(e_n) \\ &= \Big\langle u, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n \Big\rangle \end{aligned}$$

for every $u \in V$, where the first equality comes from 6.30(a). Thus setting

6.43 $$v = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n,$$

we have $\varphi(u) = \langle u, v \rangle$ for every $u \in V$, as desired.

Now we prove that only one vector $v \in V$ has the desired behavior. Suppose $v_1, v_2 \in V$ are such that

$$\varphi(u) = \langle u, v_1 \rangle = \langle u, v_2 \rangle$$

for every $u \in V$. Then

$$0 = \langle u, v_1 \rangle - \langle u, v_2 \rangle = \langle u, v_1 - v_2 \rangle$$

for every $u \in V$. Taking $u = v_1 - v_2$ shows that $v_1 - v_2 = 0$. Thus $v_1 = v_2$, completing the proof of the uniqueness part of the result. ■

6.44 example: *computation illustrating Riesz representation theorem*

Suppose we want to find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

6.45
$$\int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\,dt = \int_{-1}^{1} pq$$

for every polynomial $p \in \mathcal{P}_2(\mathbf{R})$. To do this, we make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining $\langle p, q\rangle$ to be the right side of the equation above for $p, q \in \mathcal{P}_2(\mathbf{R})$. Note that the left side of the equation above does not equal the inner product in $\mathcal{P}_2(\mathbf{R})$ of p and the function $t \mapsto \cos(\pi t)$ because this last function is not a polynomial.

Define a linear functional φ on $\mathcal{P}_2(\mathbf{R})$ by letting

$$\varphi(p) = \int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\,dt$$

for each $p \in \mathcal{P}_2(\mathbf{R})$. Now use the orthonormal basis from Example 6.34 and apply formula 6.43 from the proof of the Riesz representation theorem to see that if $p \in \mathcal{P}_2(\mathbf{R})$, then $\varphi(p) = \langle p, q\rangle$, where

$$\begin{aligned} q(x) &= \Big(\int_{-1}^{1} \sqrt{\tfrac{1}{2}}\cos(\pi t)\,dt\Big)\sqrt{\tfrac{1}{2}} + \Big(\int_{-1}^{1} \sqrt{\tfrac{3}{2}}\,t\cos(\pi t)\,dt\Big)\sqrt{\tfrac{3}{2}}x \\ &\quad + \Big(\int_{-1}^{1} \sqrt{\tfrac{45}{8}}\big(t^2 - \tfrac{1}{3}\big)\cos(\pi t)\,dt\Big)\sqrt{\tfrac{45}{8}}\big(x^2 - \tfrac{1}{3}\big). \end{aligned}$$

A bit of calculus applied to the equation above shows that

$$q(x) = \tfrac{15}{2\pi^2}\big(1 - 3x^2\big).$$

The same procedure shows that if we want to find $q \in \mathcal{P}_5(\mathbf{R})$ such that 6.45 holds for all $p \in \mathcal{P}_5(\mathbf{R})$, then we should take

$$q(x) = \tfrac{105}{8\pi^4}\Big(\big(27 - 2\pi^2\big) + \big(24\pi^2 - 270\big)x^2 + \big(315 - 30\pi^2\big)x^4\Big).$$

Suppose V is finite-dimensional and φ a linear functional on V. Then 6.43 gives a formula for the vector v that satisfies

$$\varphi(u) = \langle u, v\rangle$$

for all $u \in V$. Specifically, we have

$$v = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n.$$

The right side of the equation above seems to depend on the orthonormal basis $e_1, \dots, e_n$ as well as on φ. However, 6.42 tells us that v is uniquely determined by φ. Thus the right side of the equation above is the same regardless of which orthonormal basis $e_1, \dots, e_n$ of V is chosen.

For two additional different proofs of the Riesz representation theorem, see 6.58 and also Exercise 13 in Section 6C.

Exercises 6B

1 Suppose $e_1, \dots, e_m$ is a list of vectors in V such that

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \dots, a_m \in \mathbf{F}$. Show that $e_1, \dots, e_m$ is an orthonormal list.

This exercise provides a converse to 6.24.

2 (a) Suppose $\theta \in \mathbf{R}$. Show that both

$$(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \quad \text{and} \quad (\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$$

are orthonormal bases of $\mathbf{R}^2$.

(b) Show that each orthonormal basis of $\mathbf{R}^2$ is of the form given by one of the two possibilities in (a).

3 Suppose $e_1, \dots, e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = \left|\langle v, e_1\rangle\right|^2 + \cdots + \left|\langle v, e_m\rangle\right|^2 \iff v \in \operatorname{span}(e_1, \dots, e_m).$$

4 Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} fg.$$

Hint: The following formulas should help.

$$(\sin x)(\cos y) = \frac{\sin(x-y) + \sin(x+y)}{2}$$

$$(\sin x)(\sin y) = \frac{\cos(x-y) - \cos(x+y)}{2}$$

$$(\cos x)(\cos y) = \frac{\cos(x-y) + \cos(x+y)}{2}$$

5 Suppose $f\colon [-\pi, \pi] \to \mathbf{R}$ is continuous. For each nonnegative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \quad \text{and} \quad b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x)\sin(kx)\,dx.$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} \left(a_k^2 + b_k^2\right) \le \int_{-\pi}^{\pi} f^2.$$

The inequality above is actually an equality for all continuous functions $f\colon [-\pi, \pi] \to \mathbf{R}$. However, proving that this inequality is an equality involves Fourier series techniques beyond the scope of this book.

6 Suppose $e_1, \dots, e_n$ is an orthonormal basis of V.

(a) Prove that if $v_1, \dots, v_n$ are vectors in V such that

$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$

for each k, then $v_1, \dots, v_n$ is a basis of V.

(b) Show that there exist $v_1, \dots, v_n \in V$ such that

$$\|e_k - v_k\| \leq \frac{1}{\sqrt{n}}$$

for each k, but $v_1, \dots, v_n$ is not linearly independent.

This exercise states in (a) that an appropriately small perturbation of an orthonormal basis is a basis. Then (b) shows that the number $1/\sqrt{n}$ on the right side of the inequality in (a) cannot be lower.

7 Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ has an upper-triangular matrix with respect to the basis $(1,0,0)$, $(1,1,1)$, $(1,1,2)$. Find an orthonormal basis of $\mathbf{R}^3$ with respect to which T has an upper-triangular matrix.

8 Make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining $\langle p, q\rangle = \int_0^1 pq$ for all $p, q \in \mathcal{P}_2(\mathbf{R})$.

(a) Apply the Gram–Schmidt procedure to the basis $1, x, x^2$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

(b) The differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to the basis $1, x, x^2$, which is not an orthonormal basis. Find the matrix of the differentiation operator on $\mathcal{P}_2(\mathbf{R})$ with respect to the orthonormal basis produced in (a) and verify that this matrix is upper triangular, as expected from the proof of 6.37.

9 Suppose $e_1, \dots, e_m$ is the result of applying the Gram–Schmidt procedure to a linearly independent list $v_1, \dots, v_m$ in V. Prove that $\langle v_k, e_k\rangle > 0$ for each $k = 1, \dots, m$.

10 Suppose $v_1, \dots, v_m$ is a linearly independent list in V. Explain why the orthonormal list produced by the formulas of the Gram–Schmidt procedure (6.32) is the only orthonormal list $e_1, \dots, e_m$ in V such that $\langle v_k, e_k\rangle > 0$ and $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$ for each $k = 1, \dots, m$.

The result in this exercise is used in the proof of 7.58.

11 Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that $p\left(\frac{1}{2}\right) = \int_0^1 pq$ for every $p \in \mathcal{P}_2(\mathbf{R})$.

12 Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x)\cos(\pi x)\,dx = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

13 Show that a list $v_1, \dots, v_m$ of vectors in V is linearly dependent if and only if the Gram–Schmidt formula in 6.32 produces $f_k = 0$ for some $k \in \{1, \dots, m\}$.

This exercise gives an alternative to Gaussian elimination techniques for determining whether a list of vectors in an inner product space is linearly dependent.

14 Suppose V is a real inner product space and $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \dots, e_m$ of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for all $k \in \{1, \dots, m\}$.

15 Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle u, v \rangle_1 = 0$ if and only if $\langle u, v \rangle_2 = 0$. Prove that there is a positive number c such that $\langle u, v \rangle_1 = c\langle u, v \rangle_2$ for every $u, v \in V$.

This exercise shows that if two inner products have the same pairs of orthogonal vectors, then each of the inner products is a scalar multiple of the other inner product.

16 Suppose V is finite-dimensional. Suppose $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$ are inner products on V with corresponding norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Prove that there exists a positive number c such that $\|v\|_1 \leq c\|v\|_2$ for every $v \in V$.

17 Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if T is an operator on V such that 1 is the only eigenvalue of T and $\|Tv\| \leq \|v\|$ for all $v \in V$, then T is the identity operator.

18 Suppose $u_1, \dots, u_m$ is a linearly independent list in V. Show that there exists $v \in V$ such that $\langle u_k, v \rangle = 1$ for all $k \in \{1, \dots, m\}$.

19 Suppose $v_1, \dots, v_n$ is a basis of V. Prove that there exists a basis $u_1, \dots, u_n$ of V such that

$$\langle v_j, u_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

20 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that

$$ST = TS$$

for all $S, T \in \mathcal{E}$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

This exercise strengthens Exercise 9(b) in Section 5E (in the context of inner product spaces) by asserting that the basis in that exercise can be chosen to be orthonormal.

21 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, and all eigenvalues of T have absolute value less than 1. Let $\epsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \leq \epsilon\|v\|$ for every $v \in V$.

22 Suppose $C[-1,1]$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g\rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1,1]$. Let φ be the linear functional on $C[-1,1]$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C[-1,1]$ such that

$$\varphi(f) = \langle f, g\rangle$$

for every $f \in C[-1,1]$.

This exercise shows that the Riesz representation theorem (6.42) does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.

23 For all $u, v \in V$, define $d(u,v) = \|u - v\|$.

(a) Show that d is a metric on V.

(b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).

(c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

This exercise requires familiarity with metric spaces.

orthogonality at the Supreme Court

Law professor Richard Friedman presenting a case before the U.S. Supreme Court in 2010:

Mr. Friedman: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging—
Chief Justice Roberts: I'm sorry. Entirely what?
Mr. Friedman: Orthogonal. Right angle. Unrelated. Irrelevant.
Chief Justice Roberts: Oh.
Justice Scalia: What was that adjective? I liked that.
Mr. Friedman: Orthogonal.
Chief Justice Roberts: Orthogonal.
Mr. Friedman: Right, right.
Justice Scalia: Orthogonal, ooh. (Laughter.)
Justice Kennedy: I knew this case presented us a problem. (Laughter.)