1 Section 6.B Exercises

Exercises with solutions from Section 6.B of [LADR].

Exercise 6.B.1. (a) Suppose $\theta \in \mathbf{R}$. Show that $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$ and $(\cos\theta, \sin\theta)$, $(\sin\theta, -\cos\theta)$ 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 of part (a).

Solution. (a) Observe that

$$\langle(\cos\theta, \sin\theta), (\cos\theta, \sin\theta)\rangle = \langle(-\sin\theta, \cos\theta), (-\sin\theta, \cos\theta)\rangle = \cos^2\theta + \sin^2\theta = 1,$$

$$\langle(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\rangle = \cos\theta\sin\theta - \cos\theta\sin\theta = 0.$$

Thus $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$ is an orthonormal basis of $\mathbf{R}^2$, and we can similarly show that $(\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$ is an orthonormal basis of $\mathbf{R}^2$.
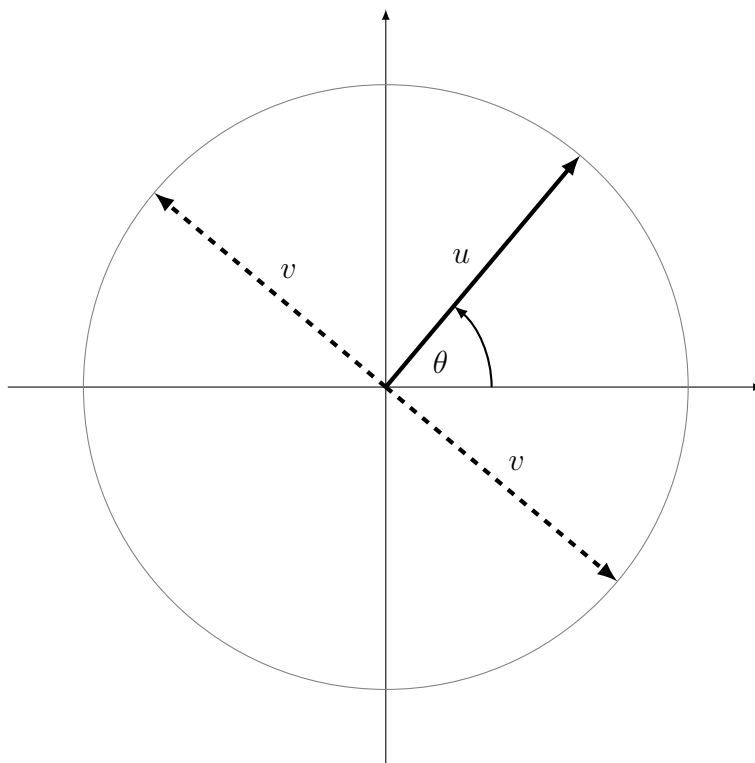

Figure 1: Orthonormal basis of $\mathbf{R}^2$

(b) Suppose u, v is an orthonormal basis of $\mathbf{R}^2$. Let θ be the angle that u makes with the positive x-axis, as shown in Figure 1. Note that $\|u\| = 1$, so that u lies on the circle of radius 1 centered at the origin in $\mathbf{R}^2$. It follows that $u = (\cos\theta, \sin\theta)$.

Since u and v are orthogonal, plane geometry tells us that v either makes an angle of $\theta + \frac{\pi}{2}$ or $\theta - \frac{\pi}{2}$ with the positive x-axis, and since $\|v\| = 1$ we know that v also lies on the circle of radius 1 centered at the origin in $\mathbf{R}^2$. It follows that

$$v = \left(\cos\left(\theta + \tfrac{\pi}{2}\right), \sin\left(\theta + \tfrac{\pi}{2}\right)\right) = (-\sin\theta, \cos\theta)$$

$$\text{or} \quad v = \left(\cos\left(\theta - \tfrac{\pi}{2}\right), \sin\left(\theta - \tfrac{\pi}{2}\right)\right) = (\sin\theta, -\cos\theta).$$

Exercise 6.B.2. Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. Let $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$$

if and only if $v \in \text{span}(e_1, \ldots, e_m)$.

Solution. Suppose $v \in \text{span}(e_1, \ldots e_m)$. Note that $e_1, \ldots, e_m$ is linearly independent (6.26) and hence is a basis of $\text{span}(e_1, \ldots, e_m)$. It follows from 6.30 that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2.$$

Now suppose that $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$, let $v' = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m$, and note that $\|v'\|^2 = \|v\|^2$ by 6.30. Observe that

$$\langle v', v\rangle = \left\langle \sum_{j=1}^{m}\langle v, e_j\rangle e_j, v\right\rangle = \sum_{j=1}^{m}\langle\langle v, e_j\rangle e_j, v\rangle = \sum_{j=1}^{m}|\langle v, e_j\rangle|^2 = \|v\|^2.$$

It follows that

$$\|v' - v\|^2 = \langle v' - v, v' - v\rangle = \langle v', v'\rangle + \langle v, v\rangle - 2\text{Re}\langle v', v\rangle = 2\|v\|^2 - 2\|v\|^2 = 0,$$

from which it follows that $v = v' \in \text{span}(e_1, \ldots, e_m)$.

Exercise 6.B.3. 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 $\mathbf{R}^3$ (use the usual inner product on $\mathbf{R}^3$) with respect to which T has an upper-triangular matrix.

Solution. Let $u_1 = (1, 0, 0)$, $u_2 = (1, 1, 1)$, and $u_3 = (1, 1, 2)$. According to the proof of 6.37, if we apply the Gram-Schmidt procedure to the basis u_1, u_2, u_3, we will obtain an orthonormal basis of $\mathbf{R}^3$ with respect to which T has an upper-triangular matrix.

Since $\|u_1\| = 1$, we can take $e_1 = u_1$. We calculate

$$e_2 = \frac{u_2 - \langle u_2, e_1 \rangle e_1}{\|u_2 - \langle u_2, e_1 \rangle e_1\|} = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

$$\text{and} \quad e_3 = \frac{u_3 - \langle u_3, e_2 \rangle e_2 - \langle u_3, e_1 \rangle e_1}{\|u_3 - \langle u_3, e_2 \rangle e_2 - \langle u_3, e_1 \rangle e_1\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right).$$

Thus e_1, e_2, e_3 is the desired orthonormal basis.

Exercise 6.B.4. Suppose n is a positive integer. Prove that

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

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

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) \; \mathrm{d}x.$$

[*The orthonormal list above is often used for modeling periodic phenomena such as tides.*]

Solution. We calculate, for $1 \le j \le n$,

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}} \right\rangle = \int_{-\pi}^{\pi} \left(\frac{1}{\sqrt{2\pi}} \right)^2 \; \mathrm{d}x = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}x = 1;$$

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\cos jx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \left(\frac{\cos jx}{\sqrt{\pi}} \right)^2 \; \mathrm{d}x = \frac{2}{j\pi} \int_{0}^{j\pi} \cos^2 y \; \mathrm{d}y = \frac{1}{j\pi} [y + \sin y \cos y]_{y=0}^{y=j\pi} = 1;$$

$$\left\langle \frac{\sin jx}{\sqrt{\pi}}, \frac{\sin jx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \left(\frac{\sin jx}{\sqrt{\pi}} \right)^2 \; \mathrm{d}x = \frac{2}{j\pi} \int_{0}^{j\pi} \sin^2 y \; \mathrm{d}y = \frac{1}{j\pi} [y - \sin y \cos y]_{y=0}^{y=j\pi} = 1;$$

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\cos jx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} \frac{\cos jx}{\sqrt{\pi}} \; \mathrm{d}x = \frac{\sqrt{2}}{j\pi} \int_{0}^{j\pi} \cos y \; \mathrm{d}y = \frac{\sqrt{2}}{j\pi} [\sin y]_{y=0}^{y=j\pi} = 0;$$

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\sin jx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} \frac{\sin jx}{\sqrt{\pi}} \; \mathrm{d}x = 0;$$

where we have used that $\sin jx$ is an odd function for the last equality. For $1 \le j, k \le n$ such

that $j \neq k$, we have

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\cos kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{\cos jx}{\sqrt{\pi}} \frac{\cos kx}{\sqrt{\pi}} \, \mathrm{d}x = \frac{1}{\pi} \int_0^{\pi} \cos((j-k)x) + \cos((j+k)x) \, \mathrm{d}x$$

$$= \frac{1}{\pi} \left[\frac{\sin((j-k)x)}{j-k} \right]_{x=0}^{x=\pi} + \frac{1}{\pi} \left[\frac{\sin((j+k)x)}{j+k} \right]_{x=0}^{x=\pi} = 0;$$

$$\left\langle \frac{\sin jx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{\sin jx}{\sqrt{\pi}} \frac{\sin kx}{\sqrt{\pi}} \, \mathrm{d}x = \frac{1}{\pi} \int_0^{\pi} \cos((j-k)x) - \cos((j+k)x) \, \mathrm{d}x$$

$$= \frac{1}{\pi} \left[\frac{\sin((j-k)x)}{j-k} \right]_{x=0}^{x=\pi} - \frac{1}{\pi} \left[\frac{\sin((j+k)x)}{j+k} \right]_{x=0}^{x=\pi} = 0.$$

Finally, for any $1 \leq j, k \leq n$, we have

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{\cos jx}{\sqrt{\pi}} \frac{\sin kx}{\sqrt{\pi}} \, \mathrm{d}x = 0,$$

where we have used that $\cos jx \sin kx$ is an odd function for the last equality.

Exercise 6.B.5. On $\mathcal{P}_2(\mathbf{R})$, consider the inner product given by

$$\langle p, q \rangle = \int_0^1 p(x)q(x) \, \mathrm{d}x.$$

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

Solution. Since $\int_0^1 \mathrm{d}x = 1$, we can take $e_1 = 1$. We calculate

$$\langle x, e_1 \rangle = \int_0^1 x \, \mathrm{d}x = \frac{1}{2} \quad \text{and} \quad \|x - \langle x, e_1 \rangle e_1\| = \left(\int_0^1 \left(x - \tfrac{1}{2}\right)^2 \, \mathrm{d}x \right)^{1/2} = \frac{1}{2\sqrt{3}}.$$

and thus

$$e_2 = \frac{x - \langle x, e_1 \rangle e_1}{\|x - \langle x, e_1 \rangle e_1\|} = 2\sqrt{3}\left(x - \tfrac{1}{2}\right).$$

Similarly,

$$\langle x^2, e_1 \rangle = \int_0^1 x^2 \, \mathrm{d}x = \frac{1}{3}, \quad \langle x^2, e_2 \rangle = 2\sqrt{3} \int_0^1 x^2 \left(x - \tfrac{1}{2}\right) \, \mathrm{d}x = \frac{1}{2\sqrt{3}},$$

$$\text{and} \quad \|x^2 - \langle x^2, e_2 \rangle e_2 - \langle x^2, e_1 \rangle e_1\| = \left(\int_0^1 \left(x^2 - x + \tfrac{1}{6}\right)^2 \, \mathrm{d}x \right)^{1/2} = \frac{1}{6\sqrt{5}}.$$

This gives us

$$e_3 = \frac{x^2 - \langle x^2, e_2 \rangle e_2 - \langle x^2, e_1 \rangle e_1}{\| x^2 - \langle x^2, e_2 \rangle e_2 - \langle x^2, e_1 \rangle e_1 \|} = 6\sqrt{5}\left(x^2 - x + \tfrac{1}{6}\right).$$

Thus

$$e_1 = 1, \quad e_2 = 2\sqrt{3}\left(x - \tfrac{1}{2}\right), \quad e_3 = 6\sqrt{5}\left(x^2 - x + \tfrac{1}{6}\right)$$

is an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ with respect to the given inner product.

Exercise 6.B.6. Find an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ (with inner product as in Exercise 5) such that the differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to this basis.

Solution. Observe that the differentiation operator has the upper-triangular matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

with respect to the basis $1, x, x^2$. As the proof of 6.37 shows, if we apply the Gram-Schmidt procedure to the basis $1, x, x^2$ we will obtain an orthonormal basis with respect to which the differentiation operator has an upper-triangular matrix. This orthonormal basis is precisely the one obtained in Exercise 6.B.5.

Exercise 6.B.7. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$p\left(\tfrac{1}{2}\right) = \int_0^1 p(x)q(x)\ \mathrm{d}x$$

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

Solution. Define the linear functional $\varphi : \mathcal{P}_2(\mathbf{R}) \to \mathbf{R}$ by $\varphi(p) = p\left(\tfrac{1}{2}\right)$. As 6.43 (in the proof of the Riesz Representation Theorem, 6.42) shows, if we take

$$q(x) = \varphi(e_1)e_1(x) + \varphi(e_2)e_2(x) + \varphi(e_3)e_3(x) = -15x^2 + 15x - \tfrac{3}{2},$$

where e_1, e_2, e_3 is the orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ found in Exercise 6.B.5, then

$$\varphi(p) = p\left(\tfrac{1}{2}\right) = \langle p, q \rangle = \int_0^1 p(x)q(x)\ \mathrm{d}x$$

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

Exercise 6.B.8. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x)(\cos \pi x) \, \mathrm{d}x = \int_0^1 p(x)q(x) \, \mathrm{d}x$$

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

Solution. Define the linear functional $\varphi : \mathcal{P}_2(\mathbf{R}) \to \mathbf{R}$ by $\varphi(p) = \int_0^1 p(x)(\cos \pi x) \, \mathrm{d}x$. As 6.43 (in the proof of the Riesz Representation Theorem, 6.42) shows, if we take

$$q(x) = \varphi(e_1)e_1(x) + \varphi(e_2)e_2(x) + \varphi(e_3)e_3(x) = -\frac{24}{\pi^2}\left(x - \tfrac{1}{2}\right),$$

where e_1, e_2, e_3 is the orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ found in Exercise 6.B.5, then

$$\varphi(p) = \int_0^1 p(x)(\cos \pi x) \, \mathrm{d}x = \langle p, q \rangle = \int_0^1 p(x)q(x) \, \mathrm{d}x$$

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

Exercise 6.B.9. What happens if the Gram-Schmidt Procedure is applied to a list of vectors that is not linearly independent?

Solution. Division by zero will occur. This is clear if the list of vectors contains the zero vector, but the problem arises even if this is not the case. For example, if v is non-zero, then consider applying the Gram-Schmidt procedure to the linearly dependent list $v, 2v$. We first set $e_1 = v/\|v\|$, and then run into a problem:

$$2v - \langle 2v, e_1 \rangle e_1 = 2v - 2\frac{\langle v, v \rangle}{\|v\|^2}v = 2v - 2v = 0.$$

Exercise 6.B.10. Suppose V is a real inner product space and $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \ldots, e_m$ of vectors in V such that

$$\mathrm{span}(v_1, \ldots v_j) = \mathrm{span}(e_1, \ldots, e_j)$$

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

Solution. Apply the Gram-Schmidt procedure (6.31) to the list $v_1, \ldots, v_m$ to obtain an orthonormal list of vectors $e_1, \ldots, e_m$ satisfying

$$\mathrm{span}(v_1, \ldots v_j) = \mathrm{span}(e_1, \ldots, e_j)$$

for all $j \in \{1, \ldots, m\}$. Let $\{-1, 1\}^m$ be the collection of functions $f : \{1, \ldots, m\} \to \{-1, 1\}$. For $f \in \{-1, 1\}^m$, it is straightforward to verify that the list $f(1)e_1, \ldots, f(m)e_m$ is orthonormal and satisfies

$$\mathrm{span}(v_1, \ldots v_j) = \mathrm{span}(f(1)e_1, \ldots, f(j)e_j)$$

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

Suppose that $u_1, \ldots, u_m$ is an orthonormal list of vectors such that

$$\mathrm{span}(v_1, \ldots v_j) = \mathrm{span}(e_1, \ldots, e_j) = \mathrm{span}(u_1, \ldots, u_j)$$

for all $j \in \{1, \ldots, m\}$. In particular, we have $\mathrm{span}(e_1) = \mathrm{span}(u_1)$, from which it follows that $u_1 = \lambda e_1$. Since $\|u_1\| = \|e_1\| = 1$, we see that $\lambda = \pm 1$; set $f(1) = \lambda$. Since $\mathrm{span}(e_1, e_2) = \mathrm{span}(u_1, u_2)$, we must have $u_2 \in \mathrm{span}(e_1, e_2)$. Since e_1, e_2 is an orthonormal basis of $\mathrm{span}(e_1, e_2)$, 6.30 implies that $u_2 = \langle u_2, e_1 \rangle e_1 + \langle u_2, e_2 \rangle e_2$. Note that the orthonormality of the list u_1, u_2 implies that

$$0 = \langle u_1, u_2 \rangle = \langle f(1)e_1, u_2 \rangle = f(1)\langle e_1, u_2 \rangle,$$

which implies that $\langle e_1, u_2 \rangle = 0$ since $f(1) = \pm 1$. Thus $u_2 = \lambda e_2$ for some $\lambda \in \mathbf{R}$; since $\|u_2\| = \|e_2\| = 1$, we again see that $\lambda = \pm 1$ and we can set $f(2) = \lambda$.

We continue in this manner, obtaining an $f \in \{-1, 1\}^m$ such that $u_j = f(j)e_j$ for each $1 \leq j \leq m$. It follows that the orthonormal lists $u_1, \ldots, u_m$ satisfying

$$\mathrm{span}(v_1, \ldots v_j) = \mathrm{span}(e_1, \ldots, e_j) = \mathrm{span}(u_1, \ldots, u_j)$$

for all $j \in \{1, \ldots, m\}$ are precisely those of the form $f(1)e_1, \ldots, f(m)e_m$ for $f \in \{-1, 1\}^m$. Since there are 2^m such functions, and distinct elements of $\{-1, 1\}^m$ give rise to distinct orthonormal lists, we see that there are exactly 2^m orthonormal lists with the desired property.

Exercise 6.B.11. Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle v, w \rangle_1 = 0$ if and only if $\langle v, w \rangle_2 = 0$. Prove that there is a positive number c such that $\langle v, w \rangle_1 = c\langle v, w \rangle_2$ for every $v, w \in V$.

Solution. If $V = \{0\}$ then we may take any $c > 0$ we like, since the only inner product on V is the map $(0, 0) \mapsto 0$. Suppose therefore that $V \neq \{0\}$ and for each non-zero $w \in V$ define

$$c_w = \frac{\langle w, w \rangle_1}{\langle w, w \rangle_2} > 0.$$

Suppose $v, w \in V$ are non-zero. By orthogonal decomposition (6.14), we have

$$\left\langle v - \frac{\langle v, w \rangle_2}{\langle w, w \rangle_2} w, w \right\rangle_2 = 0.$$

Our assumption is that orthogonality with respect to $\langle \cdot, \cdot \rangle_2$ is equivalent to orthogonality with respect to $\langle \cdot, \cdot \rangle_1$ and thus

$$\left\langle v - \frac{\langle v, w \rangle_2}{\langle w, w \rangle_2} w, w \right\rangle_1 = 0 \quad \Longleftrightarrow \quad \langle v, w \rangle_1 - \frac{\langle v, w \rangle_2}{\langle w, w \rangle_2} \langle w, w \rangle_1 = 0 \quad \Longleftrightarrow \quad \langle v, w \rangle_1 = c_w \langle v, w \rangle_2.$$

Reversing the roles of v and w gives $\langle w, v \rangle_1 = c_v \langle w, v \rangle_2$ and combining this with conjugate symmetry gives us

$$c_w \langle v, w \rangle_2 = \langle v, w \rangle_1 = \overline{\langle w, v \rangle_1} = \overline{c_v \langle w, v \rangle_2} = c_v \langle v, w \rangle_2.$$

Let us summarize our findings:

$$\text{for all non-zero } v, w \in V, \qquad \langle v, w \rangle_1 = c_w \langle v, w \rangle_2, \tag{1}$$

$$\text{for all non-zero } v, w \in V, \qquad c_w \langle v, w \rangle_2 = c_v \langle v, w \rangle_2. \tag{2}$$

Now, given non-zero $v, w \in V$, there exists a non-zero $u \in V$ such that $\langle u, v \rangle_2 \neq 0$ and $\langle w, u \rangle_2 \neq 0$; if $\langle v, w \rangle_2 \neq 0$ then take $u = v$ and if $\langle v, w \rangle_2 = 0$ then take $u = v + w$ (this is a special case of Exercise 6.B.13). Using (2), we have

$$c_v \langle u, v \rangle_2 = c_u \langle u, v \rangle_2 \quad \text{and} \quad c_u \langle w, u \rangle_2 = c_w \langle w, u \rangle_2.$$

Since $\langle u, v \rangle_2 \neq 0$ and $\langle w, u \rangle_2 \neq 0$, these two equations imply that $c_v = c_u = c_w$. If we denote this common value by c (noting that $c > 0$), then we have shown that $c_w = c$ for all non-zero $w \in V$. It follows from (1) that $\langle v, w \rangle_1 = c \langle v, w \rangle_2$ for all non-zero $v, w \in V$; evidently this equation also holds if either $v = 0$ or $w = 0$ and thus $\langle v, w \rangle_1 = c \langle v, w \rangle_2$ for all $v, w \in V$.

Exercise 6.B.12. Suppose V is finite-dimensional and $\langle \cdot, \cdot \rangle_1, \langle \cdot, \cdot \rangle_2$ 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 $v \in V$.

Solution. While this follows from the fact that any two norms on a finite-dimensional vector space are equivalent, which we proved in the solution to Exercise 6.A.29, there is an easier proof for the special case where the two norms both arise from inner products.

By 6.34, there exists an orthonormal basis $e_1, \ldots, e_n$ of V with respect to $\langle \cdot, \cdot \rangle_2$. Let $v \in V$ be given, so that $v = a_1 e_1 + \cdots + a_n e_n$ for some scalars $a_1, \ldots, a_n$. Observe that

$$|a_1| + \cdots + |a_n| \leq n \max\{|a_1|, \ldots, |a_n|\} \leq n \sqrt{|a_1|^2 + \cdots + |a_n|^2} = n \|v\|_2, \tag{1}$$

where the last equality follows from 6.25. If we let $M = \max\{\|e_1\|_1, \ldots, \|e_n\|_1\}$, which is positive since each $e_j \neq 0$, then it follows from (1) and the triangle inequality that

$$\|v\|_1 \leq |a_1|\|e_1\|_1 + \cdots + |a_n|\|e_n\|_1 \leq M(|a_1| + \cdots + |a_n|) \leq nM\|v\|_2.$$

Thus the desired positive constant is $c = nM$.

Exercise 6.B.13. Suppose $v_1, \ldots, v_m$ is a linearly independent list in V. Show that there exists $w \in V$ such that $\langle w, v_j \rangle > 0$ for all $j \in \{1, \ldots, m\}$.

Solution 1. Let $U = \text{span}(v_1, \ldots, v_m)$ and define a linear functional $\varphi : U \to \mathbf{F}$ by

$$\varphi(v) = a_1 + \cdots + a_m,$$

where $a_1, \ldots, a_m$ are the unique scalars satisfying $v = a_1 v_1 + \cdots a_m v_m$. Since U is finite-dimensional, we may appeal to the Riesz representation theorem (6.42) to obtain a unique $w \in U$ such that $\varphi(v) = \langle v, w \rangle$ for all $v \in U$. In particular, for any $1 \leq j \leq m$, we have

$$1 = \varphi(v_j) = \langle v_j, w \rangle \quad \implies \quad \langle w, v_j \rangle > 0.$$

Solution 2. Let $U = \text{span}(v_1, \ldots, v_m)$ and note that $\dim U = m$ since $v_1, \ldots, v_m$ is a basis of U. Define a linear map $T : U \to \mathbf{F}^m$ by

$$Tu = (\langle u, v_1 \rangle, \ldots, \langle u, v_m \rangle).$$

Suppose that $u \in U$ is such that $Tu = 0$, which is the case if and only if $\langle u, v_j \rangle = 0$ for all $1 \leq j \leq m$. Since $u \in U$, we have $u = a_1 v_1 + \cdots + a_m v_m$ for some scalars $a_1, \ldots, a_m$. Observe that

$$\langle u, u \rangle = \langle u, a_1 v_1 + \cdots + a_m v_m \rangle = \overline{a_1}\langle u, v_1 \rangle + \cdots + \overline{a_m}\langle u, v_m \rangle = 0.$$

It follows that $u = 0$ and hence that T is injective. Since $\dim U = \dim \mathbf{F}^m = m$, rank-nullity (3.22) implies that T is surjective. Consequently, there exists some $w \in U$ such that

$$Tw = (\langle w, v_1 \rangle, \ldots, \langle w, v_m \rangle) = (1, \ldots, 1).$$

Exercise 6.B.14. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $v_1, \ldots, v_n$ are vectors in V such that

$$\|e_j - v_j\| < \frac{1}{\sqrt{n}}$$

for each j. Prove that $v_1, \ldots, v_n$ is a basis of V.

Solution. It will suffice to show that $v_1, \ldots, v_n$ is linearly independent, so suppose that $\sum_{j=1}^{n} a_j v_j = 0$. Then

$$\sum_{j=1}^{n} |a_j|^2 = \left\| \sum_{j=1}^{n} a_j e_j \right\|^2 \tag{6.25}$$

$$= \left\| \sum_{j=1}^{n} a_j (e_j - v_j) \right\|^2 \qquad \left(\sum_{j=1}^{n} a_j v_j = 0 \right)$$

$$\leq \left(\sum_{j=1}^{n} |a_j| \|e_j - v_j\| \right)^2 \qquad \text{(triangle inequality)}$$

$$\leq \left(\sum_{j=1}^{n} |a_j|^2 \right) \left(\sum_{j=1}^{n} \|e_j - v_j\|^2 \right). \qquad \text{(Cauchy-Schwarz inequality)}$$

By assumption we have $\sum_{j=1}^{n} \|e_j - v_j\|^2 < 1$; it follows that $\sum_{j=1}^{n} |a_j|^2 = 0$, which is the case if and only if each $a_j = 0$.

Exercise 6.B.15. Suppose $C_{\mathbf{R}}([-1, 1])$ is the vector space of continuous real-valued functions on the interval $[-1, 1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, \mathrm{d}x$$

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

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

for every $f \in C_{\mathbf{R}}([-1, 1])$.

[*The exercise above shows that the Riesz Representation Theorem (6.42) does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.*]

Solution. Suppose such a g exists and define $h \in C_{\mathbf{R}}([-1, 1])$ by $h(x) = x^2 g(x)$. Then

$$0 = h(0) = \int_{-1}^{1} [x g(x)]^2 \, \mathrm{d}x,$$

which is the case if and only if $x g(x) = 0$ for all $x \in [-1, 1]$ (as the integrand is a non-negative and continuous function). This implies that $g(x) = 0$ for all $x \in [-1, 1]$, except possibly at $x = 0$; but g is continuous, so in fact g must satisfy $g(0) = 0$ also. This implies that

$$f(0) = \int_{-1}^{1} f(x) g(x) \, \mathrm{d}x = 0$$

for all $f \in C_{\mathbf{R}}([-1, 1])$, which is clearly not true. We may conclude that no such g exists.

Exercise 6.B.16. Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, all the eigenvalues of T have absolute value less than 1, and $\epsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \leq \epsilon \|v\|$ for every $v \in V$.

Solution. We will first translate the problem into a statement involving column vectors in $\mathbf{C}^{n,1}$ and matrices in $\mathbf{C}^{n,n}$, where $n = \dim V$. For the avoidance of confusion, we shall use the following notation.

- $\langle \cdot, \cdot \rangle_V : V \times V \to \mathbf{C}$. This is the given inner product on V.

- $\| \cdot \|_V : V \to [0, \infty)$. This is the norm on V arising from $\langle \cdot, \cdot \rangle_V$, i.e.

$$\|v\|_V = \sqrt{\langle v, v \rangle}.$$

- $\langle \cdot, \cdot \rangle : \mathbf{C}^{n,1} \times \mathbf{C}^{n,1} \to \mathbf{C}$. This is the Euclidean inner product on $\mathbf{C}^{n,1}$, i.e.

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \overline{y_n},$$

  where $x = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^{\mathrm{t}}$ and $y = \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix}^{\mathrm{t}}$.

- $\| \cdot \| : \mathbf{C}^{n,1} \to [0, \infty)$. This is the Euclidean norm on $\mathbf{C}^{n,1}$ arising from $\langle \cdot, \cdot \rangle$, i.e.

$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2},$$

  where $x = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^{\mathrm{t}}$.

For this exercise, we will use the following useful lemmas about matrices.

Lemma 1. If $A \in \mathbf{C}^{n,n}$, then there exists a non-negative constant $C \in \mathbf{R}$, depending only on A and n, such that

$$\|Ax\| \leq C \|x\| \text{ for all } x \in \mathbf{C}^{n,1}.$$

Proof. Suppose that

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n} \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{so that} \quad Ax = \begin{pmatrix} \sum_{k=1}^{n} A_{1,k} x_k \\ \sum_{k=1}^{n} A_{2,k} x_k \\ \vdots \\ \sum_{k=1}^{n} A_{n,k} x_k \end{pmatrix}.$$

Then

$$\|Ax\|^2 = \sum_{j=1}^{n} \left| \sum_{k=1}^{n} A_{j,k} x_k \right|^2$$

$$\leq \sum_{j=1}^{n} \left(\sum_{k=1}^{n} |A_{j,k}||x_k| \right)^2 \qquad \text{(triangle inequality)}$$

$$\leq \sum_{j=1}^{n} \left(\sum_{k=1}^{n} |A_{j,k}|^2 \right) \left(\sum_{k=1}^{n} |x_k|^2 \right) \qquad \text{(Cauchy-Schwarz inequality)}$$

$$= \left(\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{j,k}|^2 \right) \|x\|^2.$$

It follows that $\|Ax\| \leq C\|x\|$, where

$$C := \left(\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{j,k}|^2 \right)^{1/2}$$

is real, non-negative, and depends only on A and n. $\qquad\square$

In what follows, by a strictly upper-triangular matrix we mean a matrix that is upper-triangular and whose diagonal entries are also zero.

Lemma 2. Suppose that $A \in \mathbf{C}^{n,n}$ is diagonal and $B \in \mathbf{C}^{n,n}$ is strictly upper-triangular. Then AB and BA are both strictly upper-triangular.

Proof. To show that AB is strictly upper-triangular, we need to show that $(AB)_{i,k} = 0$ for $i \geq k$. Indeed,

$$(AB)_{i,k} = \sum_{j=1}^{n} A_{i,j} B_{j,k} = A_{i,i} B_{i,k} = 0;$$

the second equality follows since A is diagonal and the last equality follows since B is strictly upper-triangular. The proof that BA is strictly upper-triangular is similar. $\qquad\square$

Lemma 3. Suppose that $A \in \mathbf{C}^{n,n}$ is strictly upper-triangular. Then $A^n = 0$.

Proof. Let us think of A as a linear map $\mathbf{C}^{n,1} \to \mathbf{C}^{n,1}$; our goal is to show that A^n is the zero map.

Let $e_1, \ldots, e_n$ be the standard basis of $\mathbf{C}^{n,1}$. Define $U_0 = \{0\}$ and $U_j = \text{span}(e_1, \ldots, e_j)$ for $j = 1, \ldots, n$; note that $U_n = \mathbf{C}^{n,1}$. The fact that A is upper-triangular means that $A(U_j) \subseteq U_{j-1}$

for each j and so
$$A(U_n) \subseteq U_{n-1} \implies A^2(U_n) \subseteq A(U_{n-1}) \subseteq U_{n-2}.$$
Continuing in this manner, we see that $A^n(U_n) \subseteq U_0 = \{0\}$, i.e. $A^n(\mathbf{C}^{n,1}) = \{0\}$. It follows that A^n is the zero map. $\square$

Let us return to the exercise. Schur's theorem (6.38) implies that there is an orthonormal (with respect to $\langle \cdot, \cdot \rangle_V$) basis $e_1, \ldots, e_n$ of V such that the matrix $U := \mathcal{M}(T, (e_1, \ldots, e_n)) \in \mathbf{C}^{n,n}$ is upper-triangular. Given a vector $v = x_1 e_1 + \cdots + x_n e_n \in V$, let $x = \mathcal{M}(v) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^{\mathrm{t}} \in \mathbf{C}^{n,1}$ (recall the matrix of a vector from 3.62). Observe that

$$\|v\|_V = \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2} = \|x\|,$$

where the first equality follows from 6.25. Furthermore, for any positive integer m we have $U^m = \mathcal{M}(T^m)$ by 3.43; combining this with 3.65 we see that $\mathcal{M}(T^m v) = U^m x$. Consequently, the inequality $\|T^m v\|_V \le \epsilon \|v\|_V$ is equivalent to $\|U^m x\| \le \epsilon \|x\|$. Given this, to prove that there exists a positive integer m such that

$$\|T^m v\|_V \le \epsilon \|v\|_V \text{ for all } v \in V,$$

it will suffice to show that there exists a positive integer m such that

$$\|U^m x\| \le \epsilon \|x\| \text{ for all } x \in \mathbf{C}^{n,1}.$$

Let D be the diagonal matrix whose diagonal entries are exactly those of U, and let N be the strictly upper-triangular matrix whose entries above the diagonal are exactly those of U, so that $U = D + N$. Let C be the non-negative real constant obtained from Lemma 1 for the matrix N and let
$$\rho = \max\{|D_{1,1}|, \ldots, |D_{n,n}|\} = \max\{|U_{1,1}|, \ldots, |U_{n,n}|\};$$
note that $0 \le \rho < 1$ since the diagonal elements of D and U are precisely the eigenvalues of T. Note further that
$$\|Dx\| = \left(|D_{1,1}|^2 |x_1|^2 + \cdots + |D_{n,n}|^2 |x_n|^2 \right)^{1/2} \le \rho \|x\| \tag{1}$$
for any $x \in \mathbf{C}^{n,1}$.

For a positive integer m, let H_m be the collection of all matrix products of length m formed from the matrices D and N. For example,

$$H_1 = \{D, N\}, \quad H_2 = \{DD, DN, ND, NN\},$$
$$H_3 = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}.$$

It can then be shown by induction that

$$(D + N)^m = \sum_{A \in H_m} A.$$

Observe that if $m \geq n, A \in H_m$, and A contains n or more N's, then by Lemma 2 the matrix product A must contain at least n consecutive strictly upper-triangular matrices; it follows from Lemma 3 that $A = 0$. Given this, we are only interested in matrix products $A \in H_m$ which contain $n - 1$ or less N's; let H'_m be the collection of all such matrix products, so that

$$(D + N)^m = \sum_{A \in H'_m} A.$$

Suppose $A \in H'_m$ is a product of k N's and $m - k$ D's in some order, where $0 \leq k \leq n - 1$; note that there are $\binom{m}{k}$ such products (think of m D's in a row; we have to replace k of them with N's). If $x \in \mathbf{C}^{n,1}$, then it follows from Lemma 1 and inequality (1) that

$$\|Ax\| \leq \rho^{m-k} C^k \|x\|.$$

This is best demonstrated by example:

$$\|DNDx\| \leq \rho\|NDx\| \leq \rho C\|Dx\| \leq \rho^2 C\|x\|.$$

It now follows that for any $x \in \mathbf{C}^{n,1}$ we have

$$\|U^m x\| = \|(D + N)^m x\| = \left\| \sum_{A \in H'_m} Ax \right\| \leq \sum_{A \in H'_m} \|Ax\| \leq \sum_{k=0}^{n-1} \binom{m}{k} \rho^{m-k} C^k \|x\|.$$

For $0 \leq k \leq n - 1$, we have

$$\binom{m}{k} = \frac{m!}{k!(m-k)!} \leq \frac{m!}{(m-k)!} = m(m-1)\cdots(m-k+1) \leq m^k \leq m^{n-1}.$$

Thus, letting $\mu = \max\{1, C, C^2, \ldots, C^{n-1}\}$, we have

$$\|U^m x\| \leq \mu n m^{n-1} \rho^{m-n+1} \|x\|$$

for any $x \in \mathbf{C}^{n,1}$. Since $0 \leq \rho < 1$, L'Hôpital's rule shows that $\lim_{m \to \infty} \mu n m^{n-1} \rho^{m-n+1} = 0$. Consequently, we may choose a positive integer m so that

$$\|U^m x\| \leq \epsilon \|x\|$$

for every $x \in \mathbf{C}^{n,1}$.

Exercise 6.B.17. For $u \in V$, let Φu denote the linear functional on V defined by

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

for $v \in V$.

(a) Show that if $\mathbf{F} = \mathbf{R}$, then Φ is a linear map from V to V'. (Recall from Section 3.F that $V' = \mathcal{L}(V, \mathbf{F})$ and that V' is called the dual space of V.)

(b) Show that if $\mathbf{F} = \mathbf{C}$ and $V \neq \{0\}$, then Φ is not a linear map.

(c) Show that Φ is injective.

(d) Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. Use parts (a) and (c) and a dimension-counting argument (but without using 6.42) to show that Φ is an isomorphism from V onto V'.

[Part (d) gives an alternative proof of the Riesz Representation Theorem (6.42) when $\mathbf{F} = \mathbf{R}$. Part (d) also gives a natural isomorphism (meaning that it does not depend on a choice of basis) from a finite-dimensional real inner product space onto its dual space.]

Solution. (a) This follows from 6.7 (d) and (e), noting that $\overline{\lambda} = \lambda$ if $\lambda \in \mathbf{R}$.

(b) Since $V \neq \{0\}$, there is some non-zero $u \in V$. Observe that

$$(\Phi i u)(u) = \langle u, iu \rangle = -i \langle u, u \rangle \quad \text{and} \quad (i\Phi u)(u) = i \langle u, u \rangle.$$

These are not equal since $\langle u, u \rangle \neq 0$. It follows that $\Phi i u \neq i \Phi u$ and hence that Φ is not a linear map.

(c) Despite part (b), we can still use the equivalence

$$\Phi \text{ is injective} \quad \Longleftrightarrow \quad \text{null } \Phi = \{0\}$$

for either $\mathbf{F} = \mathbf{R}$ or $\mathbf{F} = \mathbf{C}$, since the proof of this uses only the additivity of Φ, which follows from 6.7 (d). Suppose therefore that $u \in V$ is such that $\Phi u = 0$, which is to say that $\langle v, u \rangle = 0$ for all $v \in V$. In particular we have $\langle u, u \rangle = 0$, which is the case if and only if $u = 0$. It follows that null $\Phi = \{0\}$ and hence that Φ is injective.

(d) Parts (a) and (c) imply that Φ is an injective linear map from V to V'. Since $\dim V = \dim V'$ (3.95), the rank-nullity theorem (3.22) implies that Φ must be surjective. Thus Φ is an isomorphism.

[LADR] Axler, S. (2015) *Linear Algebra Done Right.* 3rd edition.