5.6 UNITARY AND ORTHOGONAL MATRICES

The purpose of this section is to examine square matrices whose columns (or rows) are orthonormal. The standard inner product and the euclidean 2-norm are the only ones used in this section, so distinguishing subscripts are omitted.

Unitary and Orthogonal Matrices

- A ***unitary matrix*** is defined to be a *complex* matrix $\mathbf{U}_{n\times n}$ whose columns (or rows) constitute an orthonormal basis for $\mathcal{C}^n$.
- An ***orthogonal matrix*** is defined to be a *real* matrix $\mathbf{P}_{n\times n}$ whose columns (or rows) constitute an orthonormal basis for $\Re^n$.

Unitary and orthogonal matrices have some nice features, one of which is the fact that they are easy to invert. To see why, notice that the columns of $\mathbf{U}_{n\times n} = \big(\mathbf{u}_1 \,|\, \mathbf{u}_2 \,|\, \cdots \,|\mathbf{u}_n\big)$ are an orthonormal set if and only if

$$[\mathbf{U}^*\mathbf{U}]_{ij} = \mathbf{u}_i^*\mathbf{u}_j = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{when } i \neq j, \end{cases} \Longleftrightarrow \mathbf{U}^*\mathbf{U} = \mathbf{I} \Longleftrightarrow \mathbf{U}^{-1} = \mathbf{U}^*.$$

Notice that because $\mathbf{U}^*\mathbf{U} = \mathbf{I} \Longleftrightarrow \mathbf{U}\mathbf{U}^* = \mathbf{I}$, the columns of $\mathbf{U}$ are orthonormal if and only if the rows of $\mathbf{U}$ are orthonormal, and this is why the definitions of unitary and orthogonal matrices can be stated either in terms of orthonormal columns or orthonormal rows.

Another nice feature is that multiplication by a unitary matrix doesn't change the length of a vector. Only the direction can be altered because

$$\|\mathbf{U}\mathbf{x}\|^2 = \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{x} = \mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2 \quad \forall\ \mathbf{x} \in \mathcal{C}^n. \tag{5.6.1}$$

Conversely, if (5.6.1) holds, then $\mathbf{U}$ must be unitary. To see this, set $\mathbf{x} = \mathbf{e}_i$ in (5.6.1) to observe $\mathbf{u}_i^*\mathbf{u}_i = 1$ for each i, and then set $\mathbf{x} = \mathbf{e}_j + \mathbf{e}_k$ for $j \neq k$ to obtain $0 = \mathbf{u}_j^*\mathbf{u}_k + \mathbf{u}_k^*\mathbf{u}_j = 2\,\mathrm{Re}\,(\mathbf{u}_j^*\mathbf{u}_k)$. By setting $\mathbf{x} = \mathbf{e}_j + \mathrm{i}\mathbf{e}_k$ in (5.6.1) it also follows that $0 = 2\,\mathrm{Im}\,(\mathbf{u}_j^*\mathbf{u}_k)$, so $\mathbf{u}_j^*\mathbf{u}_k = 0$ for each $j \neq k$, and thus (5.6.1) guarantees that $\mathbf{U}$ is unitary.

In the case of orthogonal matrices, everything is real so that $(\star)^*$ can be replaced by $(\star)^T$. Below is a summary of these observations.

Characterizations

- The following statements are equivalent to saying that a complex matrix $\mathbf{U}_{n\times n}$ is unitary.
  - ▷ $\mathbf{U}$ has orthonormal columns.
  - ▷ $\mathbf{U}$ has orthonormal rows.
  - ▷ $\mathbf{U}^{-1} = \mathbf{U}^*$.
  - ▷ $\|\mathbf{Ux}\|_2 = \|\mathbf{x}\|_2$ for every $\mathbf{x} \in \mathcal{C}^{n\times 1}$.
- The following statements are equivalent to saying that a real matrix $\mathbf{P}_{n\times n}$ is orthogonal.
  - ▷ $\mathbf{P}$ has orthonormal columns.
  - ▷ $\mathbf{P}$ has orthonormal rows.
  - ▷ $\mathbf{P}^{-1} = \mathbf{P}^T$.
  - ▷ $\|\mathbf{Px}\|_2 = \|\mathbf{x}\|_2$ for every $\mathbf{x} \in \Re^{n\times 1}$.

Example 5.6.1

- The identity matrix $\mathbf{I}$ is an orthogonal matrix.
- All permutation matrices (products of elementary interchange matrices) are orthogonal—recall Exercise 3.9.4.
- The matrix
$$\mathbf{P} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{pmatrix}$$
is an orthogonal matrix because $\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = \mathbf{I}$ or, equivalently, because the columns (and rows) constitute an orthonormal set.
- The matrix $\mathbf{U} = \frac{1}{2}\begin{pmatrix} 1+\mathrm{i} & -1+\mathrm{i} \\ 1+\mathrm{i} & 1-\mathrm{i} \end{pmatrix}$ is unitary because $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I}$ or, equivalently, because the columns (and rows) are an orthonormal set.
- An orthogonal matrix can be considered to be unitary, but a unitary matrix is generally not orthogonal.

In general, a linear operator $\mathbf{T}$ on a vector space $\mathcal{V}$ with the property that $\|\mathbf{Tx}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathcal{V}$ is called an ***isometry*** on $\mathcal{V}$. The isometries on $\Re^n$ are precisely the orthogonal matrices, and the isometries on $\mathcal{C}^n$ are the unitary matrices. The term "isometry" has an advantage in that it can be used to treat the real and complex cases simultaneously, but for clarity we will often revert back to the more cumbersome "orthogonal" and "unitary" terminology.

The geometrical concepts of projection, reflection, and rotation are among the most fundamental of all linear transformations in $\Re^2$ and $\Re^3$ (see Example 4.7.1 for three simple examples), so pursuing these ideas in higher dimensions is only natural. The reflector and rotator given in Example 4.7.1 are isometries (because they preserve length), but the projector is not. We are about to see that the same is true in more general settings.

Elementary Orthogonal Projectors

For a vector $\mathbf{u} \in \mathcal{C}^{n\times 1}$ such that $\|\mathbf{u}\| = 1$, a matrix of the form

$$\mathbf{Q} = \mathbf{I} - \mathbf{u}\mathbf{u}^* \tag{5.6.2}$$

is called an ***elementary orthogonal projector.*** More general projectors are discussed on pp. 386 and 429.

To understand the nature of elementary projectors consider the situation in $\Re^3$. Suppose that $\|\mathbf{u}_{3\times 1}\| = 1$, and let $\mathbf{u}^\perp$ denote the space (the plane through the origin) consisting of all vectors that are perpendicular to $\mathbf{u}$—we call $\mathbf{u}^\perp$ the ***orthogonal complement*** of $\mathbf{u}$ (a more general definition appears on p. 403). The matrix $\mathbf{Q} = \mathbf{I} - \mathbf{u}\mathbf{u}^T$ is the orthogonal projector onto $\mathbf{u}^\perp$ in the sense that $\mathbf{Q}$ maps each $\mathbf{x} \in \Re^{3\times 1}$ to its orthogonal projection in $\mathbf{u}^\perp$ as shown in Figure 5.6.1.

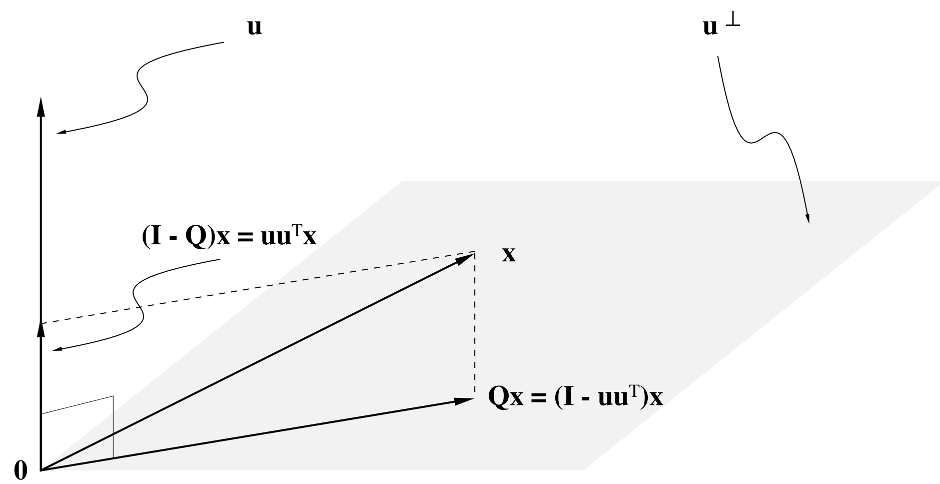


FIGURE 5.6.1

To see this, observe that each $\mathbf{x}$ can be resolved into two components

$$\mathbf{x} = (\mathbf{I} - \mathbf{Q})\mathbf{x} + \mathbf{Q}\mathbf{x}, \quad \text{where} \quad (\mathbf{I} - \mathbf{Q})\mathbf{x} \perp \mathbf{Q}\mathbf{x}.$$

The vector $(\mathbf{I} - \mathbf{Q})\mathbf{x} = \mathbf{u}(\mathbf{u}^T\mathbf{x})$ is on the line determined by $\mathbf{u}$, and $\mathbf{Q}\mathbf{x}$ is in the plane $\mathbf{u}^\perp$ because $\mathbf{u}^T\mathbf{Q}\mathbf{x} = \mathbf{0}$.

The situation is exactly as depicted in Figure 5.6.1. Notice that $(\mathbf{I}-\mathbf{Q})\mathbf{x} = \mathbf{u}\mathbf{u}^T\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the line determined by $\mathbf{u}$ and $\left\|\mathbf{u}\mathbf{u}^T\mathbf{x}\right\| = |\mathbf{u}^T\mathbf{x}|$. This provides a nice interpretation of the magnitude of the standard inner product. Below is a summary.

Geometry of Elementary Projectors

For vectors $\mathbf{u}, \mathbf{x} \in \mathcal{C}^{n\times 1}$ such that $\|\mathbf{u}\| = 1$,

- $(\mathbf{I} - \mathbf{u}\mathbf{u}^*)\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the orthogonal complement $\mathbf{u}^\perp$, the space of all vectors orthogonal to $\mathbf{u}$; (5.6.3)
- $\mathbf{u}\mathbf{u}^*\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the one-dimensional space $span\{\mathbf{u}\}$; (5.6.4)
- $|\mathbf{u}^*\mathbf{x}|$ represents the length of the orthogonal projection of $\mathbf{x}$ onto the one-dimensional space $span\{\mathbf{u}\}$. (5.6.5)

In passing, note that elementary projectors are never isometries—they can't be because they are not unitary matrices in the complex case and not orthogonal matrices in the real case. Furthermore, isometries are nonsingular but elementary projectors are singular.

Example 5.6.2

Problem: Determine the orthogonal projection of $\mathbf{x}$ onto $span\{\mathbf{u}\}$, and then find the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}^\perp$ for $\mathbf{x} = \begin{pmatrix} 2\\ 0\\ 1\end{pmatrix}$ and $\mathbf{u} = \begin{pmatrix} 2\\ -1\\ 3\end{pmatrix}$.

Solution: We cannot apply (5.6.3) and (5.6.4) directly because $\|\mathbf{u}\| \neq 1$, but this is not a problem because

$$\left\|\frac{\mathbf{u}}{\|\mathbf{u}\|}\right\| = 1, \qquad span\{\mathbf{u}\} = span\left\{\frac{\mathbf{u}}{\|\mathbf{u}\|}\right\}, \quad \text{and} \quad \mathbf{u}^\perp = \left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right)^\perp.$$

Consequently, the orthogonal projection of $\mathbf{x}$ onto $span\{\mathbf{u}\}$ is given by

$$\left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right)\left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\right)^T\mathbf{x} = \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\mathbf{x} = \frac{1}{2}\begin{pmatrix} 2\\ -1\\ 3\end{pmatrix},$$

and the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}^\perp$ is

$$\left(\mathbf{I} - \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{x} = \mathbf{x} - \frac{\mathbf{u}\mathbf{u}^T\mathbf{x}}{\mathbf{u}^T\mathbf{u}} = \frac{1}{2}\begin{pmatrix} 2\\ 1\\ -1\end{pmatrix}.$$

There is nothing special about the numbers in this example. For every nonzero vector $\mathbf{u} \in \mathcal{C}^{n\times 1}$, the orthogonal projectors onto $span\{\mathbf{u}\}$ and $\mathbf{u}^\perp$ are

$$\mathbf{P_u} = \frac{\mathbf{uu}^*}{\mathbf{u}^*\mathbf{u}} \qquad \text{and} \qquad \mathbf{P}_{\mathbf{u}^\perp} = \mathbf{I} - \frac{\mathbf{uu}^*}{\mathbf{u}^*\mathbf{u}}. \tag{5.6.6}$$

Elementary Reflectors

For $\mathbf{u}_{n\times 1} \neq \mathbf{0}$, the ***elementary reflector*** about $\mathbf{u}^\perp$ is defined to be

$$\mathbf{R} = \mathbf{I} - 2\frac{\mathbf{uu}^*}{\mathbf{u}^*\mathbf{u}} \tag{5.6.7}$$

or, equivalently,

$$\mathbf{R} = \mathbf{I} - 2\mathbf{uu}^* \quad \text{when} \quad \|\mathbf{u}\| = 1. \tag{5.6.8}$$

Elementary reflectors are also called ***Householder transformations,***[46] and they are analogous to the simple reflector given in Example 4.7.1. To understand why, suppose $\mathbf{u} \in \Re^{3\times 1}$ and $\|\mathbf{u}\| = 1$ so that $\mathbf{Q} = \mathbf{I} - \mathbf{uu}^T$ is the orthogonal projector onto the plane $\mathbf{u}^\perp$. For each $\mathbf{x} \in \Re^{3\times 1}$, $\mathbf{Qx}$ is the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}^\perp$ as shown in Figure 5.6.1. To locate $\mathbf{Rx} = (\mathbf{I} - 2\mathbf{uu}^T)\mathbf{x}$, notice that $\mathbf{Q}(\mathbf{Rx}) = \mathbf{Qx}$. In other words, $\mathbf{Qx}$ is simultaneously the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}^\perp$ as well as the orthogonal projection of $\mathbf{Rx}$ onto $\mathbf{u}^\perp$. This together with $\|\mathbf{x} - \mathbf{Qx}\| = |\mathbf{u}^T\mathbf{x}| = \|\mathbf{Qx} - \mathbf{Rx}\|$ implies that $\mathbf{Rx}$ *is the reflection of* $\mathbf{x}$ *about the plane* $\mathbf{u}^\perp$, exactly as depicted in Figure 5.6.2. (Reflections about more general subspaces are examined in Exercise 5.13.21.)

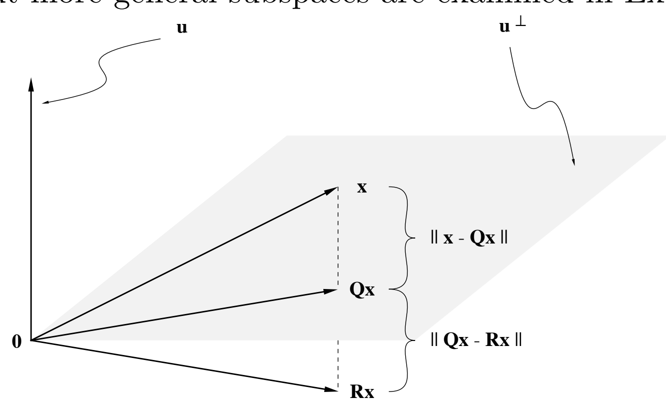


FIGURE 5.6.2

[46] Alston Scott Householder (1904–1993) was one of the first people to appreciate and promote the use of elementary reflectors for numerical applications. Although his 1937 Ph.D. dissertation at University of Chicago concerned the calculus of variations, Householder's passion was mathematical biology, and this was the thrust of his career until it was derailed by the war effort in 1944. Householder joined the Mathematics Division of Oak Ridge National Laboratory in 1946 and became its director in 1948. He stayed at Oak Ridge for the remainder of his career, and he became a leading figure in numerical analysis and matrix computations. Like his counterpart J. Wallace Givens (p. 333) at the Argonne National Laboratory, Householder was one of the early presidents of SIAM.

Properties of Elementary Reflectors

- All elementary reflectors $\mathbf{R}$ are unitary, hermitian, and involutory ($\mathbf{R}^2 = \mathbf{I}$). That is,
$$\mathbf{R} = \mathbf{R}^* = \mathbf{R}^{-1}. \tag{5.6.9}$$
- If $\mathbf{x}_{n\times 1}$ is a vector whose first entry is $x_1 \neq 0$, and if
$$\mathbf{u} = \mathbf{x} \pm \mu \|\mathbf{x}\| \mathbf{e}_1, \quad \text{where} \quad \mu = \begin{cases} 1 & \text{if } x_1 \text{ is real,} \\ x_1/|x_1| & \text{if } x_1 \text{ is not real,} \end{cases} \tag{5.6.10}$$
is used to build the elementary reflector $\mathbf{R}$ in (5.6.7), then
$$\mathbf{R}\mathbf{x} = \mp\mu \|\mathbf{x}\| \mathbf{e}_1. \tag{5.6.11}$$
In other words, this $\mathbf{R}$ "reflects" $\mathbf{x}$ onto the first coordinate axis.
Computational Note: To avoid cancellation when using floating-point arithmetic for real matrices, set $\mathbf{u} = \mathbf{x} + \operatorname{sign}(x_1) \|\mathbf{x}\| \mathbf{e}_1$.

Proof of (5.6.9). It is clear that $\mathbf{R} = \mathbf{R}^*$, and the fact that $\mathbf{R} = \mathbf{R}^{-1}$ is established simply by verifying that $\mathbf{R}^2 = \mathbf{I}$.

Proof of (5.6.10). Observe that $\mathbf{R} = \mathbf{I} - 2\hat{\mathbf{u}}\hat{\mathbf{u}}^*$, where $\hat{\mathbf{u}} = \mathbf{u}/\|\mathbf{u}\|$.

Proof of (5.6.11). Write $\mathbf{R}\mathbf{x} = \mathbf{x} - 2\mathbf{u}\mathbf{u}^*\mathbf{x}/\mathbf{u}^*\mathbf{u} = \mathbf{x} - (2\mathbf{u}^*\mathbf{x}/\mathbf{u}^*\mathbf{u})\mathbf{u}$ and verify that $2\mathbf{u}^*\mathbf{x} = \mathbf{u}^*\mathbf{u}$ to conclude $\mathbf{R}\mathbf{x} = \mathbf{x} - \mathbf{u} = \mp\mu \|\mathbf{x}\| \mathbf{e}_1$. ∎

Example 5.6.3

Problem: Given $\mathbf{x} \in \mathcal{C}^{n\times 1}$ such that $\|\mathbf{x}\| = 1$, construct an orthonormal basis for $\mathcal{C}^n$ that contains $\mathbf{x}$.

Solution: An efficient solution is to build a unitary matrix that contains $\mathbf{x}$ as its first column. Set $\mathbf{u} = \mathbf{x} \pm \mu\mathbf{e}_1$ in $\mathbf{R} = \mathbf{I} - 2(\mathbf{u}\mathbf{u}^*/\mathbf{u}^*\mathbf{u})$ and notice that (5.6.11) guarantees $\mathbf{R}\mathbf{x} = \mp\mu\mathbf{e}_1$, so multiplication on the left by $\mathbf{R}$ (remembering that $\mathbf{R}^2 = \mathbf{I}$) produces $\mathbf{x} = \mp\mu\mathbf{R}\mathbf{e}_1 = [\mp\mu\mathbf{R}]_{*1}$. Since $|\mp\mu| = 1$, $\mathbf{U} = \mp\mu\mathbf{R}$ is a unitary matrix with $\mathbf{U}_{*1} = \mathbf{x}$, so the columns of $\mathbf{U}$ provide the desired orthonormal basis. For example, to construct an orthonormal basis for $\Re^4$ that includes $\mathbf{x} = (1/3)\begin{pmatrix} -1 & 2 & 0 - 2 \end{pmatrix}^T$, set

$$\mathbf{u} = \mathbf{x} - \mathbf{e}_1 = \frac{1}{3}\begin{pmatrix} -4 \\ 2 \\ 0 \\ -2 \end{pmatrix} \text{ and compute } \mathbf{R} = \mathbf{I} - 2\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} = \frac{1}{3}\begin{pmatrix} -1 & 2 & 0 & -2 \\ 2 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ -2 & 1 & 0 & 2 \end{pmatrix}.$$

The columns of $\mathbf{R}$ do the job.

Now consider rotation, and begin with a basic problem in $\Re^2$. If a nonzero vector $\mathbf{u} = (u_1, u_2)$ is rotated counterclockwise through an angle θ to produce $\mathbf{v} = (v_1, v_2)$, how are the coordinates of $\mathbf{v}$ related to the coordinates of $\mathbf{u}$? To answer this question, refer to Figure 5.6.3, and use the fact that $\|\mathbf{u}\| = \nu = \|\mathbf{v}\|$ (rotation is an isometry) together with some elementary trigonometry to obtain

$$\begin{aligned} v_1 &= \nu\cos(\phi+\theta) = \nu(\cos\theta\cos\phi - \sin\theta\sin\phi), \\ v_2 &= \nu\sin(\phi+\theta) = \nu(\sin\theta\cos\phi + \cos\theta\sin\phi). \end{aligned} \tag{5.6.12}$$

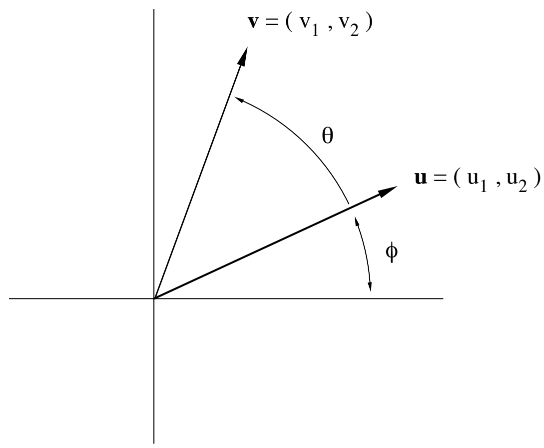


Figure 5.6.3

Substituting $\cos\phi = u_1/\nu$ and $\sin\phi = u_2/\nu$ into (5.6.12) yields

$$\begin{aligned} v_1 &= (\cos\theta)u_1 - (\sin\theta)u_2, \\ v_2 &= (\sin\theta)u_1 + (\cos\theta)u_2, \end{aligned} \quad \text{or} \quad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{5.6.13}$$

In other words, $\mathbf{v} = \mathbf{P}\mathbf{u}$, where $\mathbf{P}$ is the ***rotator*** (rotation operator)

$$\mathbf{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{5.6.14}$$

Notice that $\mathbf{P}$ is an orthogonal matrix because $\mathbf{P}^T\mathbf{P} = \mathbf{I}$. This means that if $\mathbf{v} = \mathbf{P}\mathbf{u}$, then $\mathbf{u} = \mathbf{P}^T\mathbf{v}$, and hence $\mathbf{P}^T$ is also a rotator, but in the opposite direction of that associated with $\mathbf{P}$. That is, $\mathbf{P}^T$ is the rotator associated with the angle $-\theta$. This is confirmed by the fact that if θ is replaced by $-\theta$ in (5.6.14), then $\mathbf{P}^T$ is produced.

Rotating vectors in $\Re^3$ around any one of the coordinate axes is similar. For example, consider rotation around the z-axis. Suppose that $\mathbf{v} = (v_1, v_2, v_3)$ is obtained by rotating $\mathbf{u} = (u_1, u_2, u_3)$ counterclockwise[47] through an angle θ around the z-axis. Just as before, the goal is to determine the relationship between the coordinates of $\mathbf{u}$ and $\mathbf{v}$. Since we are rotating around the z-axis,

[47] This is from the perspective of looking down the z-axis onto the xy-plane.

it is evident (see Figure 5.6.4) that the third coordinates are unaffected—i.e., $v_3 = u_3$. To see how the xy-coordinates of $\mathbf{u}$ and $\mathbf{v}$ are related, consider the orthogonal projections

$$\mathbf{u}_p = (u_1,\, u_2,\, 0) \quad \text{and} \quad \mathbf{v}_p = (v_1,\, v_2,\, 0)$$

of $\mathbf{u}$ and $\mathbf{v}$ onto the xy-plane.

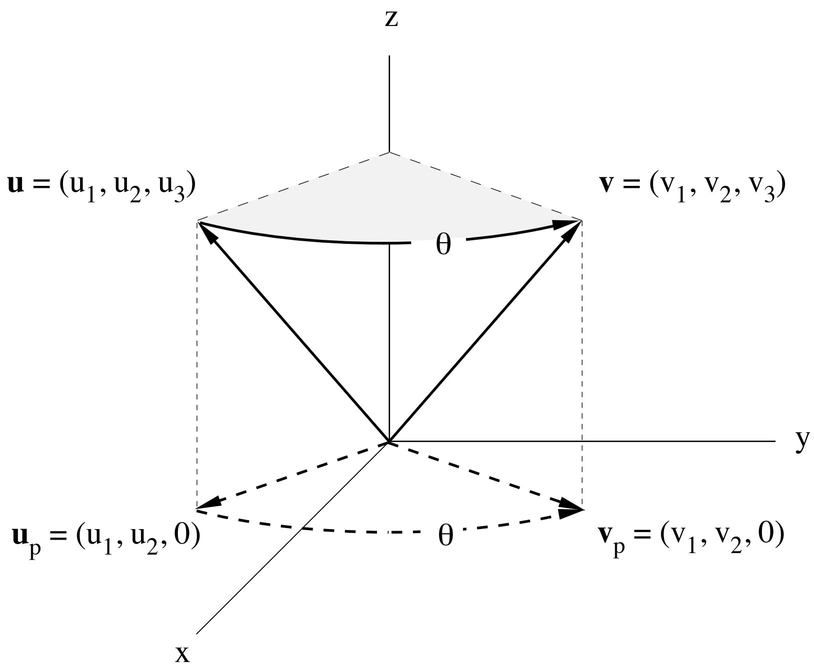


Figure 5.6.4

It's apparent from Figure 5.6.4 that the problem has been reduced to rotation in the xy-plane, and we already know how to do this. Combining (5.6.13) with the fact that $v_3 = u_3$ produces the equation

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

so

$$\mathbf{P}_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is the matrix that rotates vectors in $\Re^3$ counterclockwise around the z-axis through an angle θ. It is easy to verify that $\mathbf{P}_z$ is an orthogonal matrix and that $\mathbf{P}_z^{-1} = \mathbf{P}_z^T$ rotates vectors *clockwise* around the z-axis.

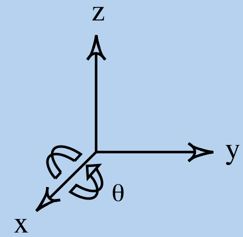
By using similar techniques, it is possible to derive orthogonal matrices that rotate vectors around the x-axis or around the y-axis. Below is a summary of these rotations in $\Re^3$.

Rotations in $\mathbf{R}^3$

A vector $\mathbf{u} \in \Re^3$ can be rotated counterclockwise through an angle θ around a coordinate axis by means of a multiplication $\mathbf{P}_\star\mathbf{u}$ in which $\mathbf{P}_\star$ is an appropriate orthogonal matrix as described below.

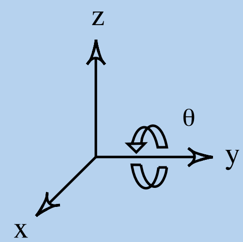
Rotation around the x-Axis

$$\mathbf{P}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$



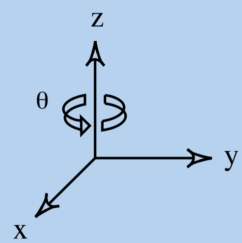
Rotation around the y-Axis

$$\mathbf{P}_y = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$



Rotation around the z-Axis

$$\mathbf{P}_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Note: The minus sign appears above the diagonal in $\mathbf{P}_x$ and $\mathbf{P}_z$, but below the diagonal in $\mathbf{P}_y$. This is not a mistake—it's due to the orientation of the positive x-axis with respect to the yz-plane.

Example 5.6.4

3-D Rotational Coordinates. Suppose that three counterclockwise rotations are performed on the three-dimensional solid shown in Figure 5.6.5. First rotate the solid in View (a) 90° around the x-axis to obtain the orientation shown in View (b). Then rotate View (b) 45° around the y-axis to produce View (c) and, finally, rotate View (c) 60° around the z-axis to end up with View (d). You can follow the process by watching how the notch, the vertex $\mathbf{v}$, and the lighter shaded face move.

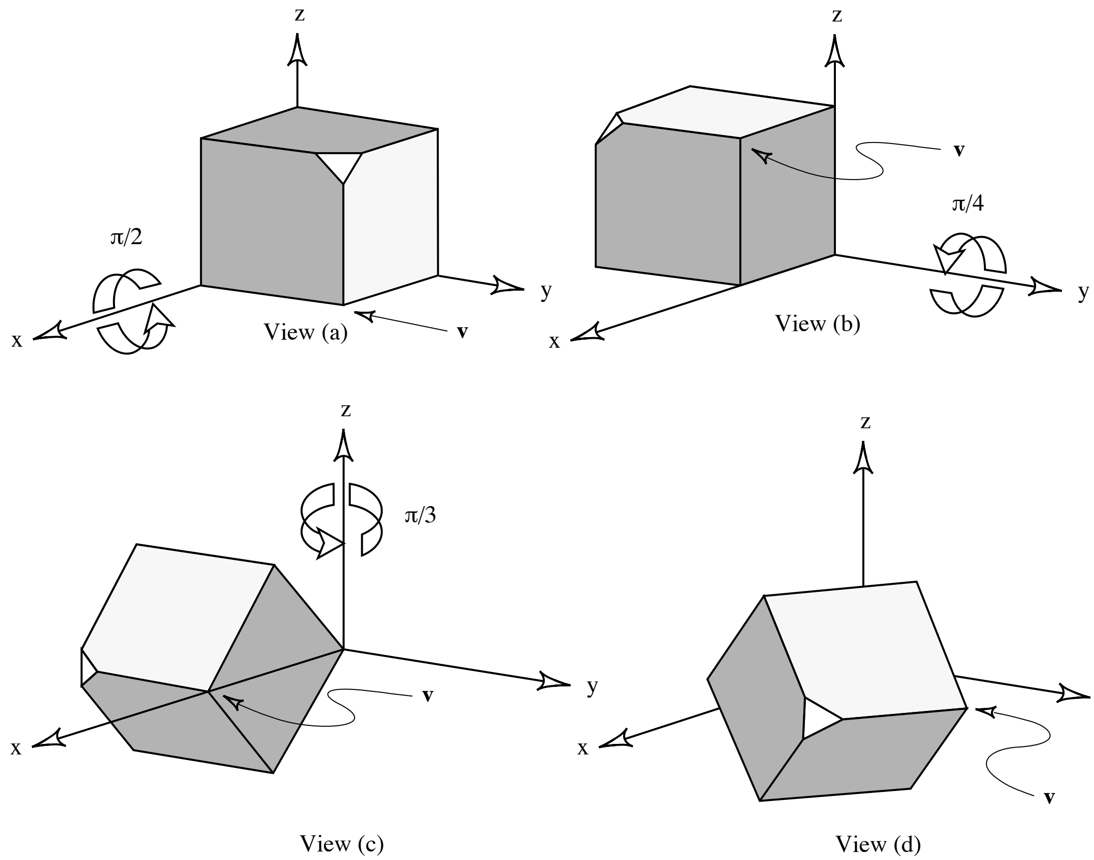


FIGURE 5.6.5

Problem: If the coordinates of each vertex in View (a) are specified, what are the coordinates of each vertex in View (d)?

Solution: If $\mathbf{P}_x$ is the rotator that maps points in View (a) to corresponding points in View (b), and if $\mathbf{P}_y$ and $\mathbf{P}_z$ are the respective rotators carrying View (b) to View (c) and View (c) to View (d), then

$$\mathbf{P}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{P}_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad \mathbf{P}_z = \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so

$$\mathbf{P} = \mathbf{P}_z\mathbf{P}_y\mathbf{P}_x = \frac{1}{2\sqrt{2}}\begin{pmatrix} 1 & 1 & \sqrt{6} \\ \sqrt{3} & \sqrt{3} & -\sqrt{2} \\ -2 & 2 & 0 \end{pmatrix} \tag{5.6.15}$$

is the orthogonal matrix that maps points in View (a) to their corresponding images in View (d). For example, focus on the vertex labeled $\mathbf{v}$ in View (a), and let $\mathbf{v}_a$, $\mathbf{v}_b$, $\mathbf{v}_c$, and $\mathbf{v}_d$ denote its respective coordinates in Views (a), (b), (c), and (d). If $\mathbf{v}_a = (\,1 \quad 1 \quad 0\,)^T$, then $\mathbf{v}_b = \mathbf{P}_x\mathbf{v}_a = (\,1 \quad 0 \quad 1\,)^T$,

$$\mathbf{v}_c = \mathbf{P}_y\mathbf{v}_b = \mathbf{P}_y\mathbf{P}_x\mathbf{v}_a = \begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{v}_d = \mathbf{P}_z\mathbf{v}_c = \mathbf{P}_z\mathbf{P}_y\mathbf{P}_x\mathbf{v}_a = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{6}/2 \\ 0 \end{pmatrix}.$$

More generally, if the coordinates of each of the ten vertices in View (a) are placed as columns in a ***vertex matrix,***

$$\mathbf{V}_a = \begin{array}{c} \begin{array}{cccc} \mathbf{v}_1 & \mathbf{v}_2 & & \mathbf{v}_{10} \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} x_1 & x_2 & \cdots & x_{10} \\ y_1 & y_2 & \cdots & y_{10} \\ z_1 & z_2 & \cdots & z_{10} \end{pmatrix} \end{array}, \quad \text{then} \quad \mathbf{V}_d = \mathbf{P}_z\mathbf{P}_y\mathbf{P}_x\mathbf{V}_a = \begin{array}{c} \begin{array}{cccc} \hat{\mathbf{v}}_1 & \hat{\mathbf{v}}_2 & & \hat{\mathbf{v}}_{10} \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_{10} \\ \hat{y}_1 & \hat{y}_2 & \cdots & \hat{y}_{10} \\ \hat{z}_1 & \hat{z}_2 & \cdots & \hat{z}_{10} \end{pmatrix} \end{array}$$

is the vertex matrix for the orientation shown in View (d). The polytope in View (d) is drawn by identifying pairs of vertices $(\mathbf{v}_i, \mathbf{v}_j)$ in $\mathbf{V}_a$ that have an edge between them, and by drawing an edge between the corresponding vertices $(\hat{\mathbf{v}}_i, \hat{\mathbf{v}}_j)$ in $\mathbf{V}_d$.

Example 5.6.5

3-D Computer Graphics. Consider the problem of displaying and manipulating views of a three-dimensional solid on a two-dimensional computer display monitor. One simple technique is to use a *wire-frame representation* of the solid consisting of a mesh of points (vertices) on the solid's surface connected by straight line segments (edges). Once these vertices and edges have been defined, the resulting polytope can be oriented in any desired manner as described in Example 5.6.4, so all that remains are the following problems.

Problem: How should the vertices and edges of a three-dimensional polytope be plotted on a two-dimensional computer monitor?

Solution: Assume that the screen represents the yz-plane, and suppose the x-axis is orthogonal to the screen so that it points toward the viewer's eye as shown in Figure 5.6.6.

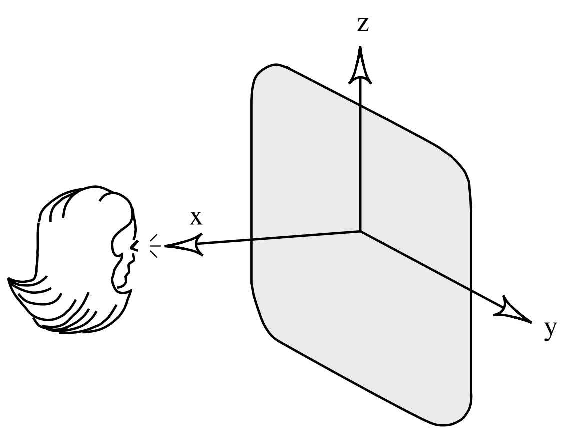


Figure 5.6.6

A solid in the xyz-coordinate system appears to the viewer as the orthogonal projection of the solid onto the yz-plane, and the projection of a polytope is easy to draw. Just set the x-coordinate of each vertex to 0 (i.e., ignore the x-coordinates), plot the (y, z)-coordinates on the yz-plane (the screen), and draw edges between appropriate vertices. For example, suppose that the vertices of the polytope in Figure 5.6.5 are numbered as indicated below in Figure 5.6.7,

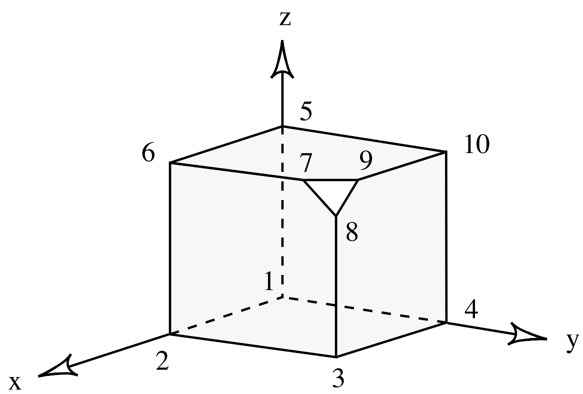


FIGURE 5.6.7

and suppose that the associated vertex matrix is

$$\mathbf{V} = \begin{array}{c} \\ x \\ y \\ z \end{array}\!\!\begin{array}{c} \begin{array}{cccccccccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5 & \mathbf{v}_6 & \mathbf{v}_7 & \mathbf{v}_8 & \mathbf{v}_9 & \mathbf{v}_{10} \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & .8 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & .8 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & .8 & 1 & 1 \end{pmatrix} \end{array}.$$

There are 15 edges, and they can be recorded in an ***edge matrix***

$$\mathbf{E} = \begin{array}{c} \begin{array}{ccccccccccccccc} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 & \mathbf{e}_5 & \mathbf{e}_6 & \mathbf{e}_7 & \mathbf{e}_8 & \mathbf{e}_9 & \mathbf{e}_{10} & \mathbf{e}_{11} & \mathbf{e}_{12} & \mathbf{e}_{13} & \mathbf{e}_{14} & \mathbf{e}_{15} \end{array} \\ \begin{pmatrix} 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 7 & 8 & 9 & 10 \\ 2 & 3 & 4 & 1 & 5 & 6 & 8 & 10 & 6 & 7 & 8 & 9 & 9 & 10 & 5 \end{pmatrix} \end{array}$$

in which the k^{th} column represents an edge between the indicated pair of vertices. To display the image of the polytope in Figure 5.6.7 on a monitor, (i) drop the first row from $\mathbf{V}$, (ii) plot the remaining yz-coordinates on the screen, (iii) draw edges between appropriate vertices as dictated by the information in the edge matrix $\mathbf{E}$. To display the image of the polytope after it has been rotated counterclockwise around the x-, y-, and z-axes by 90°, 45°, and 60°, respectively, use the orthogonal matrix $\mathbf{P} = \mathbf{P}_z\mathbf{P}_y\mathbf{P}_x$ determined in (5.6.15) and compute the product

$$\mathbf{PV} = \begin{pmatrix} 0 & .354 & .707 & .354 & .866 & 1.22 & 1.5 & 1.4 & 1.5 & 1.22 \\ 0 & .612 & 1.22 & .612 & -.5 & .112 & .602 & .825 & .602 & .112 \\ 0 & -.707 & 0 & .707 & 0 & -.707 & -.141 & 0 & .141 & .707 \end{pmatrix}.$$

Now proceed as before—(i) ignore the first row of $\mathbf{PV}$, (ii) plot the points in the second and third row of $\mathbf{PV}$ as yz-coordinates on the monitor, (iii) draw edges between appropriate vertices as indicated by the edge matrix $\mathbf{E}$.

Problem: In addition to rotation, how can a polytope (or its image on a computer monitor) be translated?

Solution: Translation of a polytope to a different point in space is accomplished by adding a constant to each of its coordinates. For example, to translate the polytope shown in Figure 5.6.7 to the location where vertex 1 is at $\mathbf{p}^T = (x_0, y_0, z_0)$ instead of at the origin, just add $\mathbf{p}$ to every point. In particular, if $\mathbf{e}$ is the column of 1's, the translated vertex matrix is

$$\mathbf{V}_{trans} = \mathbf{V}_{orig} + \begin{pmatrix} x_0 & x_0 & \cdots & x_0 \\ y_0 & y_0 & \cdots & y_0 \\ z_0 & z_0 & \cdots & z_0 \end{pmatrix} = \mathbf{V}_{orig} + \mathbf{p}\mathbf{e}^T \quad \text{(a rank-1 update).}$$

Of course, the edge matrix is not affected by translation.

Problem: How can a polytope (or its image on a computer monitor) be scaled?

Solution: Simply multiply every coordinate by the desired scaling factor. For example, to scale an image by a factor α, form the scaled vertex matrix

$$\mathbf{V}_{scaled} = \alpha \mathbf{V}_{orig},$$

and then connect the scaled vertices with appropriate edges as dictated by the edge matrix $\mathbf{E}$.

Problem: How can the faces of a polytope that are hidden from the viewer's perspective be detected so that they can be omitted from the drawing on the screen?

Solution: A complete discussion of this tricky problem would carry us too far astray, but one clever solution relying on the cross product of vectors in $\Re^3$ is presented in Exercise 5.6.21 for the case of *convex* polytopes.

Rotations in higher dimensions are straightforward generalizations of rotations in $\Re^3$. Recall from p. 328 that rotation around any particular axis in $\Re^3$ amounts to rotation in the complementary plane, and the associated 3×3 rotator is constructed by embedding a 2×2 rotator in the appropriate position in a 3×3 identity matrix. For example, rotation around the y-axis is rotation in the xz-plane, and the corresponding rotator is produced by embedding

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

in the "xz-position" of $\mathbf{I}_{3\times 3}$ to form

$$\mathbf{P}_y = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}.$$

These observations directly extend to higher dimensions.

Plane Rotations

Orthogonal matrices of the form

$$
\begin{array}{cc}
 & \begin{array}{ccccccccc} & & \text{col } i & & & \text{col } j & & & \\ & & \downarrow & & & \downarrow & & & \end{array} \\
\mathbf{P}_{ij} = & \left(\begin{array}{ccccccccc}
1 &&&&&&&& \\
& \ddots &&&&&&& \\
&& c &&& s &&& \\
&&& 1 &&&&& \\
&&&& \ddots &&&& \\
&& -s &&& c &&& \\
&&&&&& 1 && \\
&&&&&&& \ddots & \\
&&&&&&&& 1
\end{array}\right)
\begin{array}{l} \\ \\ \longleftarrow \text{row } i \\ \\ \\ \longleftarrow \text{row } j \\ \\ \\ \\ \end{array}
\end{array}
$$

in which $c^2 + s^2 = 1$ are called ***plane rotation matrices*** because they perform a rotation in the (i, j)-plane of $\Re^n$. The entries c and s are meant to suggest cosine and sine, respectively, but designating a rotation angle θ as is done in $\Re^2$ and $\Re^3$ is not useful in higher dimensions.

Plane rotations matrices $\mathbf{P}_{ij}$ are also called ***Givens***[48] ***rotations***. Applying $\mathbf{P}_{ij}$ to $\mathbf{0} \neq \mathbf{x} \in \Re^n$ rotates the (i, j)-coordinates of $\mathbf{x}$ in the sense that

$$
\mathbf{P}_{ij}\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ cx_i + sx_j \\ \vdots \\ -sx_i + cx_j \\ \vdots \\ x_n \end{pmatrix}
\begin{array}{l} \\ \\ \longleftarrow i \\ \\ \longleftarrow j \\ \\ \\ \end{array}.
$$

If x_i and x_j are not both zero, and if we set

$$
c = \frac{x_i}{\sqrt{x_i^2 + x_j^2}} \quad \text{and} \quad s = \frac{x_j}{\sqrt{x_i^2 + x_j^2}}, \tag{5.6.16}
$$

[48] J. Wallace Givens, Jr. (1910–1993) pioneered the use of plane rotations in the early days of automatic matrix computations. Givens graduated from Lynchburg College in 1928, and he completed his Ph.D. at Princeton University in 1936. After spending three years at the Institute for Advanced Study in Princeton as an assistant of O. Veblen, Givens accepted an appointment at Cornell University but later moved to Northwestern University. In addition to his academic career, Givens was the Director of the Applied Mathematics Division at Argonne National Laboratory and, like his counterpart A. S. Householder (p. 324) at Oak Ridge National Laboratory, Givens served as an early president of SIAM.

then

$$\mathbf{P}_{ij}\mathbf{x} = \begin{array}{c} \begin{pmatrix} x_1 \\ \vdots \\ \sqrt{x_i^2 + x_j^2} \\ \vdots \\ 0 \\ \vdots \\ x_n \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \longleftarrow i \\ \\ \longleftarrow j \\ \\ \\ \end{array}.$$

This means that we can selectively annihilate any component—the j^{th} in this case—by a rotation in the (i, j)-plane without affecting any entry except x_i and x_j. Consequently, plane rotations can be applied to annihilate *all* components below any particular "pivot." For example, to annihilate all entries below the first position in $\mathbf{x}$, apply a sequence of plane rotations as follows:

$$\mathbf{P}_{12}\mathbf{x} = \begin{pmatrix} \sqrt{x_1^2+x_2^2} \\ 0 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{P}_{13}\mathbf{P}_{12}\mathbf{x} = \begin{pmatrix} \sqrt{x_1^2+x_2^2+x_3^2} \\ 0 \\ 0 \\ x_4 \\ \vdots \\ x_n \end{pmatrix}, \quad \ldots, \quad \mathbf{P}_{1n}\cdots\mathbf{P}_{13}\mathbf{P}_{12}\mathbf{x} = \begin{pmatrix} \|\mathbf{x}\| \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The product of plane rotations is generally not another plane rotation, but such a product is always an orthogonal matrix, and hence it is an isometry. If we are willing to interpret "rotation in $\Re^n$" as a sequence of plane rotations, then we can say that it is always possible to "rotate" each nonzero vector onto the first coordinate axis. Recall from (5.6.11) that we can also do this with a reflection. More generally, the following statement is true.

Rotations in $\Re^n$

Every nonzero vector $\mathbf{x} \in \Re^n$ can be rotated to the i^{th} coordinate axis by a sequence of $n - 1$ plane rotations. In other words, there is an orthogonal matrix $\mathbf{P}$ such that

$$\mathbf{P}\mathbf{x} = \|\mathbf{x}\|\, \mathbf{e}_i, \tag{5.6.17}$$

where $\mathbf{P}$ has the form

$$\mathbf{P} = \mathbf{P}_{in} \cdots \mathbf{P}_{i,i+1}\mathbf{P}_{i,i-1} \cdots \mathbf{P}_{i1}.$$

Example 5.6.6

Problem: If $\mathbf{x} \in \Re^n$ is a vector such that $\|\mathbf{x}\| = 1$, explain how to use plane rotations to construct an orthonormal basis for $\Re^n$ that contains $\mathbf{x}$.

Solution: This is almost the same problem as that posed in Example 5.6.3, and, as explained there, the goal is to construct an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}_{*1} = \mathbf{x}$. But this time we need to use plane rotations rather than an elementary reflector. Equation (5.6.17) asserts that we can build an orthogonal matrix from a sequence of plane rotations $\mathbf{P} = \mathbf{P}_{1n} \cdots \mathbf{P}_{13}\mathbf{P}_{12}$ such that $\mathbf{Px} = \mathbf{e}_1$. Thus $\mathbf{x} = \mathbf{P}^T\mathbf{e}_1 = \mathbf{P}^T_{*1}$, and hence the columns of $\mathbf{Q} = \mathbf{P}^T$ serve the purpose. For example, to extend

$$\mathbf{x} = \frac{1}{3}\begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \end{pmatrix}$$

to an orthonormal basis for $\Re^4$, sequentially annihilate the second and fourth components of $\mathbf{x}$ by using (5.6.16) to construct the following plane rotations:

$$\mathbf{P}_{12}\mathbf{x} = \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} & 0 & 0 \\ -2/\sqrt{5} & -1/\sqrt{5} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \frac{1}{3}\begin{pmatrix} -1 \\ 2 \\ 0 \\ -2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} \sqrt{5} \\ 0 \\ 0 \\ -2 \end{pmatrix},$$

$$\mathbf{P}_{14}\Big(\mathbf{P}_{12}\mathbf{x}\Big) = \begin{pmatrix} \sqrt{5}/3 & 0 & 0 & -2/3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & \sqrt{5}/3 \end{pmatrix} \frac{1}{3}\begin{pmatrix} \sqrt{5} \\ 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore, the columns of

$$\mathbf{Q} = (\mathbf{P}_{14}\mathbf{P}_{12})^T = \mathbf{P}^T_{12}\mathbf{P}^T_{14} = \begin{pmatrix} -1/3 & -2/\sqrt{5} & 0 & -2/3\sqrt{5} \\ 2/3 & -1/\sqrt{5} & 0 & 4/3\sqrt{5} \\ 0 & 0 & 1 & 0 \\ -2/3 & 0 & 0 & \sqrt{5}/3 \end{pmatrix}$$

are an orthonormal set containing the specified vector $\mathbf{x}$.

Exercises for section 5.6

5.6.1. Determine which of the following matrices are isometries.

(a) $\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}$. (b) $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$. (d) $\begin{pmatrix} e^{i\theta_1} & 0 & \cdots & 0 \\ 0 & e^{i\theta_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{i\theta_n} \end{pmatrix}$.

5.6.2. Is $\begin{pmatrix} \dfrac{1+i}{\sqrt{3}} & \dfrac{1+i}{\sqrt{6}} \\ \dfrac{i}{\sqrt{3}} & \dfrac{-2\,i}{\sqrt{6}} \end{pmatrix}$ a unitary matrix?

5.6.3. (a) How many 3×3 matrices are both diagonal and orthogonal?
(b) How many $n \times n$ matrices are both diagonal and orthogonal?
(c) How many $n \times n$ matrices are both diagonal and unitary?

5.6.4. (a) Under what conditions on the real numbers α and β will

$$\mathbf{P} = \begin{pmatrix} \alpha+\beta & \beta-\alpha \\ \alpha-\beta & \beta+\alpha \end{pmatrix}$$

be an orthogonal matrix?

(b) Under what conditions on the real numbers α and β will

$$\mathbf{U} = \begin{pmatrix} 0 & \alpha & 0 & \mathrm{i}\beta \\ \alpha & 0 & \mathrm{i}\beta & 0 \\ 0 & \mathrm{i}\beta & 0 & \alpha \\ \mathrm{i}\beta & 0 & \alpha & 0 \end{pmatrix}$$

be a unitary matrix?

5.6.5. Let $\mathbf{U}$ and $\mathbf{V}$ be two $n \times n$ unitary (orthogonal) matrices.
(a) Explain why the product $\mathbf{UV}$ must be unitary (orthogonal).
(b) Explain why the sum $\mathbf{U}+\mathbf{V}$ need not be unitary (orthogonal).
(c) Explain why $\begin{pmatrix} \mathbf{U}_{n\times n} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{m\times m} \end{pmatrix}$ must be unitary (orthogonal).

5.6.6. Cayley Transformation. Prove, as Cayley did in 1846, that if $\mathbf{A}$ is skew hermitian (or real skew symmetric), then

$$\mathbf{U} = (\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A})^{-1} = (\mathbf{I}+\mathbf{A})^{-1}(\mathbf{I}-\mathbf{A})$$

is unitary (orthogonal) by first showing that $(\mathbf{I}+\mathbf{A})^{-1}$ exists for skew-hermitian matrices, and $(\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A})^{-1} = (\mathbf{I}+\mathbf{A})^{-1}(\mathbf{I}-\mathbf{A})$ (recall Exercise 3.7.6). **Note:** There is a more direct approach, but it requires the diagonalization theorem for normal matrices—see Exercise 7.5.5.

5.6.7. Suppose that $\mathbf{R}$ and $\mathbf{S}$ are elementary reflectors.
(a) Is $\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}$ an elementary reflector?
(b) Is $\begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{pmatrix}$ an elementary reflector?

5.6.8. (a) Explain why the standard inner product is invariant under a unitary transformation. That is, if $\mathbf{U}$ is any unitary matrix, and if $\mathbf{u} = \mathbf{Ux}$ and $\mathbf{v} = \mathbf{Uy}$, then

$$\mathbf{u}^*\mathbf{v} = \mathbf{x}^*\mathbf{y}.$$

(b) Given any two vectors $\mathbf{x}, \mathbf{y} \in \Re^n$, explain why the angle between them is invariant under an orthogonal transformation. That is, if $\mathbf{u} = \mathbf{Px}$ and $\mathbf{v} = \mathbf{Py}$, where $\mathbf{P}$ is an orthogonal matrix, then

$$\cos\theta_{\mathbf{u},\mathbf{v}} = \cos\theta_{\mathbf{x},\mathbf{y}}.$$

5.6.9. Let $\mathbf{U}_{m\times r}$ be a matrix with orthonormal columns, and let $\mathbf{V}_{k\times n}$ be a matrix with orthonormal rows. For an arbitrary $\mathbf{A} \in \mathcal{C}^{r\times k}$, solve the following problems using the matrix 2-norm (p. 281) and the Frobenius matrix norm (p. 279).

(a) Determine the values of $\|\mathbf{U}\|_2$, $\|\mathbf{V}\|_2$, $\|\mathbf{U}\|_F$, and $\|\mathbf{V}\|_F$.
(b) Show that $\|\mathbf{UAV}\|_2 = \|\mathbf{A}\|_2$. (**Hint:** Start with $\|\mathbf{UA}\|_2$.)
(c) Show that $\|\mathbf{UAV}\|_F = \|\mathbf{A}\|_F$.

Note: In particular, these properties are valid when $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices. Because of parts (b) and (c), the 2-norm and the F-norm are said to be ***unitarily invariant norms.***

5.6.10. Let $\mathbf{u} = \begin{pmatrix} -2 \\ 1 \\ 3 \\ -1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 \\ 4 \\ 0 \\ -1 \end{pmatrix}$.

(a) Determine the orthogonal projection of $\mathbf{u}$ onto $span\{\mathbf{v}\}$.
(b) Determine the orthogonal projection of $\mathbf{v}$ onto $span\{\mathbf{u}\}$.
(c) Determine the orthogonal projection of $\mathbf{u}$ onto $\mathbf{v}^\perp$.
(d) Determine the orthogonal projection of $\mathbf{v}$ onto $\mathbf{u}^\perp$.

5.6.11. Consider elementary orthogonal projectors $\mathbf{Q} = \mathbf{I} - \mathbf{uu}^*$.

(a) Prove that $\mathbf{Q}$ is singular.
(b) Now prove that if $\mathbf{Q}$ is $n \times n$, then $rank(\mathbf{Q}) = n - 1$.
Hint: Recall Exercise 4.4.10.

5.6.12. For vectors $\mathbf{u}, \mathbf{x} \in \mathcal{C}^n$ such that $\|\mathbf{u}\| = 1$, let $\mathbf{p}$ be the orthogonal projection of $\mathbf{x}$ onto $span\{\mathbf{u}\}$. Explain why $\|\mathbf{p}\| \le \|\mathbf{x}\|$ with equality holding if and only if $\mathbf{x}$ is a scalar multiple of $\mathbf{u}$.

5.6.13. Let $\mathbf{x} = (1/3)\begin{pmatrix} 1 \\ -2 \\ -2 \end{pmatrix}$.

(a) Determine an elementary reflector $\mathbf{R}$ such that $\mathbf{Rx}$ lies on the x-axis.

(b) Verify by direct computation that your reflector $\mathbf{R}$ is symmetric, orthogonal, and involutory.

(c) Extend $\mathbf{x}$ to an orthonormal basis for $\Re^3$ by using an elementary reflector.

5.6.14. Let $\mathbf{R} = \mathbf{I} - 2\mathbf{uu}^*$, where $\|\mathbf{u}_{n\times 1}\| = 1$. If $\mathbf{x}$ is a *fixed point* for $\mathbf{R}$ in the sense that $\mathbf{Rx} = \mathbf{x}$, and if $n > 1$, prove that $\mathbf{x}$ must be orthogonal to $\mathbf{u}$, and then sketch a picture of this situation in $\Re^3$.

5.6.15. Let $\mathbf{x}, \mathbf{y} \in \Re^{n\times 1}$ be vectors such that $\|\mathbf{x}\| = \|\mathbf{y}\|$ but $\mathbf{x} \neq \mathbf{y}$. Explain how to construct an elementary reflector $\mathbf{R}$ such that $\mathbf{Rx} = \mathbf{y}$.
Hint: The vector $\mathbf{u}$ that defines $\mathbf{R}$ can be determined visually in $\Re^3$ by considering Figure 5.6.2.

5.6.16. Let $\mathbf{x}_{n\times 1}$ be a vector such that $\|\mathbf{x}\| = 1$, and partition $\mathbf{x}$ as

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \tilde{\mathbf{x}} \end{pmatrix}, \qquad \text{where } \tilde{\mathbf{x}} \text{ is } n-1\times 1.$$

(a) If the entries of $\mathbf{x}$ are real, and if $x_1 \neq 1$, show that

$$\mathbf{P} = \begin{pmatrix} x_1 & \tilde{\mathbf{x}}^T \\ \tilde{\mathbf{x}} & \mathbf{I} - \alpha\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T \end{pmatrix}, \qquad \text{where} \quad \alpha = \frac{1}{1-x_1}$$

is an orthogonal matrix.

(b) Suppose that the entries of $\mathbf{x}$ are complex. If $|x_1| \neq 1$, and if μ is the number defined in (5.6.10), show that the matrix

$$\mathbf{U} = \begin{pmatrix} x_1 & \mu^2\tilde{\mathbf{x}}^* \\ \tilde{\mathbf{x}} & \mu(\mathbf{I} - \alpha\tilde{\mathbf{x}}\tilde{\mathbf{x}}^*) \end{pmatrix}, \qquad \text{where} \quad \alpha = \frac{1}{1-|x_1|}$$

is unitary. **Note:** These results provide an easy way to extend a given vector to an orthonormal basis for the entire space $\Re^n$ or $\mathcal{C}^n$.

5.6.17. Perform the following sequence of rotations in $\Re^3$ beginning with

$$\mathbf{v}_0 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

1. Rotate $\mathbf{v}_0$ counterclockwise 45° around the x-axis to produce $\mathbf{v}_1$.
2. Rotate $\mathbf{v}_1$ clockwise 90° around the y-axis to produce $\mathbf{v}_2$.
3. Rotate $\mathbf{v}_2$ counterclockwise 30° around the z-axis to produce $\mathbf{v}_3$.

Determine the coordinates of $\mathbf{v}_3$ as well as an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{v}_0 = \mathbf{v}_3$.

5.6.18. Does it matter in what order rotations in $\Re^3$ are performed? For example, suppose that a vector $\mathbf{v} \in \Re^3$ is first rotated counterclockwise around the x-axis through an angle θ, and then that vector is rotated counterclockwise around the y-axis through an angle ϕ. Is the result the same as first rotating $\mathbf{v}$ counterclockwise around the y-axis through an angle ϕ followed by a rotation counterclockwise around the x-axis through an angle θ?

5.6.19. For each nonzero vector $\mathbf{u} \in \mathcal{C}^n$, prove that $\dim \mathbf{u}^\perp = n - 1$.

5.6.20. A matrix satisfying $\mathbf{A}^2 = \mathbf{I}$ is said to be an ***involution*** or an ***involutory matrix***, and a matrix $\mathbf{P}$ satisfying $\mathbf{P}^2 = \mathbf{P}$ is called a ***projector*** or is said to be an ***idempotent matrix***—properties of such matrices are developed on p. 386. Show that there is a one-to-one correspondence between the set of involutions and the set of projectors in $\mathcal{C}^{n\times n}$. **Hint:** Consider the relationship between the projectors in (5.6.6) and the reflectors (which are involutions) in (5.6.7) on p. 324.

5.6.21. When using a computer to generate and display a three-dimensional convex polytope such as the one in Example 5.6.4, it is desirable to not draw those faces that should be hidden from the perspective of a viewer positioned as shown in Figure 5.6.6. The operation of ***cross product*** in $\Re^3$ (usually introduced in elementary calculus courses) can be used to decide which faces are visible and which are not. Recall that if

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \text{ and } \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad \text{then} \quad \mathbf{u} \times \mathbf{v} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix},$$

and $\mathbf{u} \times \mathbf{v}$ is a vector orthogonal to both $\mathbf{u}$ and $\mathbf{v}$. The direction of $\mathbf{u} \times \mathbf{v}$ is determined from the so-called right-hand rule as illustrated in Figure 5.6.8.

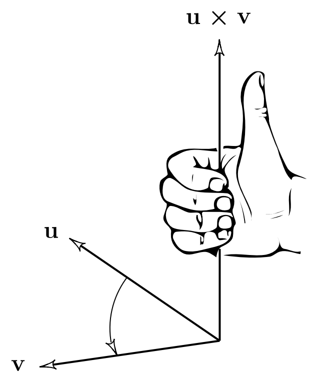


FIGURE 5.6.8

Assume the origin is interior to the polytope, and consider a particular face and three vertices $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$ on the face that are positioned as shown in Figure 5.6.9. The vector $\mathbf{n} = (\mathbf{p}_1 - \mathbf{p}_0) \times (\mathbf{p}_2 - \mathbf{p}_1)$ is orthogonal to the face, and it points in the outward direction.

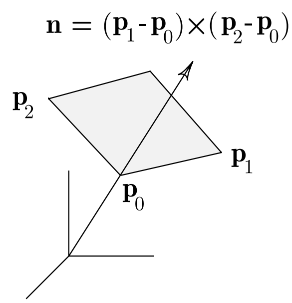


FIGURE 5.6.9

Explain why the outside of the face is visible from the perspective indicated in Figure 5.6.6 if and only if the first component of the outward normal vector $\mathbf{n}$ is positive. In other words, the face is drawn if and only if $n_1 > 0$.

Thus the modified Gram–Schmidt algorithm produces

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 10^{-3} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

which is as close to being an orthonormal set as one could reasonably hope to obtain by using 3-digit arithmetic.

5.5.10. Yes. In both cases r_{ij} is the (i, j)-entry in the upper-triangular matrix R in the QR factorization.

5.5.11. $p_0(x) = 1/\sqrt{2}, \quad p_1(x) = \sqrt{3/2}\, x, \quad p_2(x) = \sqrt{5/8}\,(3x^2 - 1)$

Solutions for exercises in section 5. 6

5.6.1. (a), (c), and (d).

5.6.2. Yes, because $\mathbf{U}^*\mathbf{U} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

5.6.3. (a) Eight: $\mathbf{D} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}$ (b) 2^n : $\mathbf{D} = \begin{pmatrix} \pm 1 & 0 & \cdots & 0 \\ 0 & \pm 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm 1 \end{pmatrix}$

(c) There are infinitely many because each diagonal entry can be any point on the unit circle in the complex plane—these matrices have the form given in part (d) of Exercise 5.6.1.

5.6.4. (a) When $\alpha^2 + \beta^2 = 1/2$. (b) When $\alpha^2 + \beta^2 = 1$.

5.6.5. (a) $(\mathbf{UV})^*(\mathbf{UV}) = \mathbf{V}^*\mathbf{U}^*\mathbf{UV} = \mathbf{V}^*\mathbf{V} = \mathbf{I}$.

(b) Consider $\mathbf{I} + (-\mathbf{I}) = \mathbf{0}$.

(c)

$$\begin{aligned} \begin{pmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{pmatrix}^* \begin{pmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{pmatrix} &= \begin{pmatrix} \mathbf{U}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^* \end{pmatrix} \begin{pmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{U}^*\mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^*\mathbf{V} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \end{aligned}$$

5.6.6. Recall from (3.7.8) or (4.2.10) that $(\mathbf{I}+\mathbf{A})^{-1}$ exists if and only if $N(\mathbf{I} + \mathbf{A}) = \mathbf{0}$, and write $\mathbf{x} \in N(\mathbf{I} + \mathbf{A}) \implies \mathbf{x} = -\mathbf{Ax} \implies \mathbf{x}^*\mathbf{x} = -\mathbf{x}^*\mathbf{Ax}$. But taking the conjugate transpose of both sides yields $\mathbf{x}^*\mathbf{x} = -\mathbf{x}^*\mathbf{A}^*\mathbf{x} = \mathbf{x}^*\mathbf{Ax}$, so $\mathbf{x}^*\mathbf{x} = 0$, and thus $\mathbf{x} = \mathbf{0}$. Replacing $\mathbf{A}$ by $-\mathbf{A}$ in Exercise 3.7.6 gives $\mathbf{A}(\mathbf{I} + \mathbf{A})^{-1} = (\mathbf{I} + \mathbf{A})^{-1}\mathbf{A}$, so

$$\begin{aligned} (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1} &= (\mathbf{I} + \mathbf{A})^{-1} - \mathbf{A}(\mathbf{I} + \mathbf{A})^{-1} \\ &= (\mathbf{I} + \mathbf{A})^{-1} - (\mathbf{I} + \mathbf{A})^{-1}\mathbf{A} = (\mathbf{I} + \mathbf{A})^{-1}(\mathbf{I} - \mathbf{A}). \end{aligned}$$

These results together with the fact that $\mathbf{A}$ is skew hermitian produce

$$\begin{aligned}\mathbf{U}^*\mathbf{U} &= (\mathbf{I}+\mathbf{A})^{-1^*}(\mathbf{I}-\mathbf{A})^*(\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A})^{-1} \\ &= (\mathbf{I}+\mathbf{A})^{*^{-1}}(\mathbf{I}-\mathbf{A})^*(\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A})^{-1} \\ &= (\mathbf{I}-\mathbf{A})^{-1}(\mathbf{I}+\mathbf{A})(\mathbf{I}-\mathbf{A})(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I}.\end{aligned}$$

5.6.7. (a) Yes—because if $\mathbf{R}=\mathbf{I}-2\mathbf{u}\mathbf{u}^*$, where $\|\mathbf{u}\|=1$, then

$$\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{pmatrix} = \mathbf{I}-2\begin{pmatrix}\mathbf{0}\\ \mathbf{u}\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathbf{u}^*\end{pmatrix} \quad \text{and} \quad \left\|\begin{pmatrix}\mathbf{0}\\ \mathbf{u}\end{pmatrix}\right\| = 1.$$

(b) No—Suppose $\mathbf{R}=\mathbf{I}-2\mathbf{u}\mathbf{u}^*$ and $\mathbf{S}=\mathbf{I}-2\mathbf{v}\mathbf{v}^*$, where $\|\mathbf{u}\|=1$ and $\|\mathbf{v}\|=1$ so that

$$\begin{pmatrix}\mathbf{R} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}\end{pmatrix} = \mathbf{I}-2\begin{pmatrix}\mathbf{u}\mathbf{u}^* & \mathbf{0}\\ \mathbf{0} & \mathbf{v}\mathbf{v}^*\end{pmatrix}.$$

If we could find a vector $\mathbf{w}$ such that $\|\mathbf{w}\|=1$ and

$$\begin{pmatrix}\mathbf{R} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}\end{pmatrix} = \mathbf{I}-2\mathbf{w}\mathbf{w}^*, \quad \text{then} \quad \mathbf{w}\mathbf{w}^* = \begin{pmatrix}\mathbf{u}\mathbf{u}^* & \mathbf{0}\\ \mathbf{0} & \mathbf{v}\mathbf{v}^*\end{pmatrix}.$$

But this is impossible because (recall Example 3.9.3)

$$rank\,(\mathbf{w}\mathbf{w}^*) = 1 \quad \text{and} \quad rank\begin{pmatrix}\mathbf{u}\mathbf{u}^* & \mathbf{0}\\ \mathbf{0} & \mathbf{v}\mathbf{v}^*\end{pmatrix} = 2.$$

5.6.8. (a) $\mathbf{u}^*\mathbf{v} = (\mathbf{U}\mathbf{x})^*\mathbf{U}\mathbf{y} = \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y} = \mathbf{x}^*\mathbf{y}$

(b) The fact that $\mathbf{P}$ is an isometry means $\|\mathbf{u}\|=\|\mathbf{x}\|$ and $\|\mathbf{v}\|=\|\mathbf{y}\|$. Use this together with part (a) and the definition of cosine given in (5.4.1) to obtain

$$\cos\theta_{\mathbf{u},\mathbf{v}} = \frac{\mathbf{u}^T\mathbf{v}}{\|\mathbf{u}\|\;\|\mathbf{v}\|} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\;\|\mathbf{y}\|} = \cos\theta_{\mathbf{x},\mathbf{y}}.$$

5.6.9. (a) Since $\mathbf{U}_{m\times r}$ has orthonormal columns, we have $\mathbf{U}^*\mathbf{U}=\mathbf{I}_r$ so that

$$\|\mathbf{U}\|_2^2 = \max_{\|\mathbf{x}\|_2=1}\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{x} = \max_{\|\mathbf{x}\|_2=1}\mathbf{x}^*\mathbf{x} = 1.$$

This together with $\|\mathbf{A}\|_2=\|\mathbf{A}^*\|_2$—recall (5.2.10)—implies $\|\mathbf{V}\|_2=1$. For the Frobenius norm we have

$$\|\mathbf{U}\|_F = [trace\,(\mathbf{U}^*\mathbf{U})]^{1/2} = [trace\,(\mathbf{I})]^{1/2} = \sqrt{r}.$$

$trace\,(\mathbf{AB}) = trace\,(\mathbf{BA})$ (Example 3.6.5) and $\mathbf{V}\mathbf{V}^*=\mathbf{I}_k \implies \|\mathbf{V}\|_F=\sqrt{k}$.

(b) First show that $\|\mathbf{UA}\|_2 = \|\mathbf{A}\|_2$ by writing

$$\begin{aligned}\|\mathbf{UA}\|_2^2 &= \max_{\|\mathbf{x}\|_2=1} \|\mathbf{UAx}\|_2^2 = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^*\mathbf{A}^*\mathbf{U}^*\mathbf{UAx} = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^*\mathbf{A}^*\mathbf{Ax}\\ &= \max_{\|\mathbf{x}\|_2=1} \|\mathbf{Ax}\|_2^2 = \|\mathbf{A}\|_2^2.\end{aligned}$$

Now use this together with $\|\mathbf{A}\|_2 = \|\mathbf{A}^*\|_2$ to observe that

$$\|\mathbf{AV}\|_2 = \|\mathbf{V}^*\mathbf{A}^*\|_2 = \|\mathbf{A}^*\|_2 = \|\mathbf{A}\|_2.$$

Therefore, $\|\mathbf{UAV}\|_2 = \|\mathbf{U}(\mathbf{AV})\|_2 = \|\mathbf{AV}\|_2 = \|\mathbf{A}\|_2.$

(c) Use $\mathit{trace}\,(\mathbf{AB}) = \mathit{trace}\,(\mathbf{BA})$ with $\mathbf{U}^*\mathbf{U} = \mathbf{I}_r$ and $\mathbf{VV}^* = \mathbf{I}_k$ to write

$$\begin{aligned}\|\mathbf{UAV}\|_F^2 &= \mathit{trace}\left((\mathbf{UAV})^*\mathbf{UAV}\right) = \mathit{trace}\,(\mathbf{V}^*\mathbf{A}^*\mathbf{U}^*\mathbf{UAV})\\ &= \mathit{trace}\,(\mathbf{V}^*\mathbf{A}^*\mathbf{AV}) = \mathit{trace}\,(\mathbf{A}^*\mathbf{AVV}^*)\\ &= \mathit{trace}\,(\mathbf{A}^*\mathbf{A}) = \|\mathbf{A}\|_F^2.\end{aligned}$$

5.6.10. Use (5.6.6) to compute the following quantities.

(a) $$\frac{\mathbf{vv}^T}{\mathbf{v}^T\mathbf{v}}\mathbf{u} = \left(\frac{\mathbf{v}^T\mathbf{u}}{\mathbf{v}^T\mathbf{v}}\right)\mathbf{v} = \frac{1}{6}\mathbf{v} = \frac{1}{6}\begin{pmatrix}1\\4\\0\\-1\end{pmatrix}$$

(b) $$\frac{\mathbf{uu}^T}{\mathbf{u}^T\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u}^T\mathbf{v}}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{u} = \frac{1}{5}\mathbf{u} = \frac{1}{5}\begin{pmatrix}-2\\1\\3\\-1\end{pmatrix}$$

(c) $$\left(\mathbf{I} - \frac{\mathbf{vv}^T}{\mathbf{v}^T\mathbf{v}}\right)\mathbf{u} = \mathbf{u} - \left(\frac{\mathbf{v}^T\mathbf{u}}{\mathbf{v}^T\mathbf{v}}\right)\mathbf{v} = \mathbf{u} - \frac{1}{6}\mathbf{v} = \frac{1}{6}\begin{pmatrix}-13\\2\\18\\-5\end{pmatrix}$$

(d) $$\left(\mathbf{I} - \frac{\mathbf{uu}^T}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{v} = \mathbf{v} - \left(\frac{\mathbf{u}^T\mathbf{v}}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{u} = \mathbf{v} - \frac{1}{5}\mathbf{u} = \frac{1}{5}\begin{pmatrix}7\\19\\-3\\-4\end{pmatrix}$$

5.6.11. (a) $N(\mathbf{Q}) \neq \{\mathbf{0}\}$ because $\mathbf{Qu} = \mathbf{0}$ and $\|\mathbf{u}\| = 1 \implies \mathbf{u} \neq \mathbf{0}$, so $\mathbf{Q}$ must be singular by (4.2.10).

(b) The result of Exercise 4.4.10 insures that $n-1 \le \mathit{rank}\,(\mathbf{Q})$, and the result of part (a) says $\mathit{rank}\,(\mathbf{Q}) \le n-1$, and therefore $\mathit{rank}\,(\mathbf{Q}) = n-1$.

5.6.12. Use (5.6.5) in conjunction with the CBS inequality given in (5.1.3) to write

$$\|\mathbf{p}\| = |\mathbf{u}^*\mathbf{x}| \le \|\mathbf{u}\|\,\|\mathbf{x}\| = \|\mathbf{x}\|.$$

The fact that equality holds if and only if $\mathbf{x}$ is a scalar multiple of $\mathbf{u}$ follows from the result of Exercise 5.1.9.

5.6.13. (a) Set $\mathbf{u} = \mathbf{x} - \|\mathbf{x}\|\,\mathbf{e}_1 = -2/3\begin{pmatrix}1\\1\\1\end{pmatrix}$, and compute

$$\mathbf{R} = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} = \frac{1}{3}\begin{pmatrix}1 & -2 & -2\\ -2 & 1 & -2\\ -2 & -2 & 1\end{pmatrix}.$$

(You could also use $\mathbf{u} = \mathbf{x} + \|\mathbf{x}\|\,\mathbf{e}_1$.)

(b) Verify that $\mathbf{R} = \mathbf{R}^T$, $\mathbf{R}^T\mathbf{R} = \mathbf{I}$, and $\mathbf{R}^2 = \mathbf{I}$.

(c) The columns of the reflector $\mathbf{R}$ computed in part (a) do the job.

5.6.14. $\mathbf{R}\mathbf{x} = \mathbf{x} \implies 2\mathbf{u}\mathbf{u}^*\mathbf{x} = \mathbf{0} \implies \mathbf{u}^*\mathbf{x} = 0$ because $\mathbf{u} \neq \mathbf{0}$.

5.6.15. If $\mathbf{R}\mathbf{x} = \mathbf{y}$ in Figure 5.6.2, then the line segment between $\mathbf{x} - \mathbf{y}$ is parallel to the line determined by $\mathbf{u}$, so $\mathbf{x} - \mathbf{y}$ itself must be a scalar multiple of $\mathbf{u}$. If $\mathbf{x} - \mathbf{y} = \alpha\mathbf{u}$, then

$$\mathbf{u} = \frac{\mathbf{x} - \mathbf{y}}{\alpha} = \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|}.$$

It is straightforward to verify that this choice of $\mathbf{u}$ produces the desired reflector.

5.6.16. You can verify by direct multiplication that $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ and $\mathbf{U}^*\mathbf{U} = \mathbf{I}$, but you can also recognize that $\mathbf{P}$ and $\mathbf{U}$ are elementary reflectors that come from Example 5.6.3 in the sense that

$$\mathbf{P} = \mathbf{I} - 2\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}, \quad \text{where} \quad \mathbf{u} = \mathbf{x} - \mathbf{e}_1 = \begin{pmatrix} x_1 - 1\\ \tilde{\mathbf{x}}\end{pmatrix}$$

and

$$\mathbf{U} = \mu\left(\mathbf{I} - 2\frac{\mathbf{u}\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}}\right), \quad \text{where} \quad \mathbf{u} = \mathbf{x} - \mu\mathbf{e}_1 = \begin{pmatrix} x_1 - \mu\\ \tilde{\mathbf{x}}\end{pmatrix}.$$

5.6.17. The final result is

$$\mathbf{v}_3 = \begin{pmatrix} -\sqrt{2}/2\\ \sqrt{6}/2\\ 1\end{pmatrix}$$

and

$$\mathbf{Q} = \mathbf{P}_z(\pi/6)\mathbf{P}_y(-\pi/2)\mathbf{P}_x(\pi/4) = \frac{1}{4}\begin{pmatrix} 0 & -\sqrt{6}-\sqrt{2} & -\sqrt{6}+\sqrt{2}\\ 0 & \sqrt{6}-\sqrt{2} & -\sqrt{6}-\sqrt{2}\\ 4 & 0 & 0\end{pmatrix}.$$

5.6.18. It matters because the rotation matrices given on p. 328 generally do not commute with each other (this is easily verified by direct multiplication). For example, this means that it is generally the case that

$$\mathbf{P}_y(\phi)\mathbf{P}_x(\theta)\mathbf{v} \neq \mathbf{P}_x(\theta)\mathbf{P}_y(\phi)\mathbf{v}.$$

5.6.19. As pointed out in Example 5.6.2, $\mathbf{u}^\perp = (\mathbf{u}/\|\mathbf{u}\|)^\perp$, so we can assume without any loss of generality that $\mathbf{u}$ has unit norm. We also know that any vector of unit norm can be extended to an orthonormal basis for $\mathcal{C}^n$—Examples 5.6.3 and 5.6.6 provide two possible ways to accomplish this. Let $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ be such an orthonormal basis for $\mathcal{C}^n$.

Claim: $span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\} = \mathbf{u}^\perp$.

Proof. $\mathbf{x} \in span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\} \implies \mathbf{x} = \sum_i \alpha_i \mathbf{v}_i \implies \mathbf{u}^*\mathbf{x} = \sum_i \alpha_i \mathbf{u}^*\mathbf{v}_i = 0 \implies \mathbf{x} \in \mathbf{u}^\perp$, and thus $span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\} \subseteq \mathbf{u}^\perp$. To establish the reverse inclusion, write $\mathbf{x} = \alpha_0\mathbf{u} + \sum_i \alpha_i \mathbf{v}_i$, and then note that $\mathbf{x} \perp \mathbf{u} \implies 0 = \mathbf{u}^*\mathbf{x} = \alpha_0 \implies \mathbf{x} \in span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$, and hence $\implies \mathbf{u}^\perp \subseteq span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$.

Consequently, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ is a basis for $\mathbf{u}^\perp$ because it is a spanning set that is linearly independent—recall (4.3.14)—and thus $\dim \mathbf{u}^\perp = n-1$.

5.6.20. The relationship between the matrices in (5.6.6) and (5.6.7) on p. 324 suggests that if $\mathbf{P}$ is a projector, then $\mathbf{A} = \mathbf{I} - 2\mathbf{P}$ is an involution—and indeed this is true because $\mathbf{A}^2 = (\mathbf{I} - 2\mathbf{P})^2 = \mathbf{I} - 4\mathbf{P} + 4\mathbf{P}^2 = \mathbf{I}$. Similarly, if $\mathbf{A}$ is an involution, then $\mathbf{P} = (\mathbf{I} - \mathbf{A})/2$ is easily verified to be a projector. Thus each projector uniquely defines an involution, and vice versa.

5.6.21. The outside of the face is visible from the perspective indicated in Figure 5.6.6 if and only if the angle θ between $\mathbf{n}$ and the positive x-axis is between $-90°$ and $+90°$. This is equivalent to saying that the cosine between $\mathbf{n}$ and $\mathbf{e}_1$ is positive, so the desired conclusion follows from the fact that

$$\cos\theta > 0 \iff \frac{\mathbf{n}^T\mathbf{e}_1}{\|\mathbf{n}\|\,\|\mathbf{e}_1\|} > 0 \iff \mathbf{n}^T\mathbf{e}_1 > 0 \iff n_1 > 0.$$

Solutions for exercises in section 5. 7

5.7.1. (a) Householder reduction produces

$$\mathbf{R}_2\mathbf{R}_1\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3/5 & 4/5 \\ 0 & 4/5 & 3/5 \end{pmatrix} \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ -2/3 & 1/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 & 19 & -34 \\ -2 & -5 & 20 \\ 2 & 8 & 37 \end{pmatrix}$$
$$= \begin{pmatrix} 3 & 15 & 0 \\ 0 & 15 & -30 \\ 0 & 0 & 45 \end{pmatrix} = \mathbf{R},$$

so

$$\mathbf{Q} = (\mathbf{R}_2\mathbf{R}_1)^T = \begin{pmatrix} 1/3 & 14/15 & -2/15 \\ -2/3 & 1/3 & 2/3 \\ 2/3 & -2/15 & 11/15 \end{pmatrix}.$$