Exercises

30. Sketch the set of points $(x_1, x_2) = \boldsymbol{x}^T$ in $\mathbb{R}^2$ such that

(a) $\|\boldsymbol{x}\|_2 = 1$ (b) $\|\boldsymbol{x}\|_1 = 1$ (c) $\|\boldsymbol{x}\|_\infty = 1$

Orthonormal Sets

Orthonormal Sets

- Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ be nonzero vectors in an inner product space V. If $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$ whenever $i \neq j$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ is said to be an *orthogonal set* of vectors

- If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ are linearly independent

- Proof

Suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are mutually orthogonal nonzero vectors and

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \tag{1}$$

If $1 \leq j \leq n$, then, taking the inner product of $\mathbf{v}_j$ with both sides of equation (1), we see that

$$\begin{aligned} c_1\langle \mathbf{v}_j, \mathbf{v}_1\rangle + c_2\langle \mathbf{v}_j, \mathbf{v}_2\rangle + \cdots + c_n\langle \mathbf{v}_j, \mathbf{v}_n\rangle &= 0 \\ c_j\|\mathbf{v}_j\|^2 &= 0 \end{aligned}$$

and hence all the scalars $c_1, c_2, \ldots, c_n$ must be 0. ■

Orthonormal Sets

- An *orthonormal set* of vectors is an orthogonal set of unit vectors
- The set $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ will be orthonormal if and only if

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

- Given any orthogonal set of nonzero vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$, it is possible to form an orthonormal set by defining

$$\boldsymbol{u}_i = \left(\frac{1}{\|\boldsymbol{v}_i\|}\right) \boldsymbol{v}_i$$

Orthonormal Sets

Ex 2. Form an orthonormal set

$$\boldsymbol{v}_1 = (1, 1, 1)^T \qquad \boldsymbol{v}_2 = (2, 1, -3)^T \qquad \boldsymbol{v}_3 = (4, -5, 1)^T$$

- Solution

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{v}_1\|}\right)\mathbf{v}_1 = \frac{1}{\sqrt{3}}(1, 1, 1)^T$$

$$\mathbf{u}_2 = \left(\frac{1}{\|\mathbf{v}_2\|}\right)\mathbf{v}_2 = \frac{1}{\sqrt{14}}(2, 1, -3)^T$$

$$\mathbf{u}_3 = \left(\frac{1}{\|\mathbf{v}_3\|}\right)\mathbf{v}_3 = \frac{1}{\sqrt{42}}(4, -5, 1)^T$$

Orthonormal Sets

- If $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k\}$ is an orthonormal set in an inner product space V, then B is a basis for the subspace $S = Span(\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k)$.

- B is an *orthonormal basis* for S

- Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ be an orthonormal basis for an inner product space V. If $\boldsymbol{v} = \sum_{i=1}^{n} c_i \boldsymbol{u}_i$, then $c_i = \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle$

- Proof

$$\langle \mathbf{v}, \mathbf{u}_i \rangle = \langle \sum_{j=1}^{n} c_j \mathbf{u}_j, \mathbf{u}_i \rangle = \sum_{j=1}^{n} c_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = \sum_{j=1}^{n} c_j \delta_{ji} = c_i$$

Orthonormal Sets

- Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ be an orthonormal basis for an inner product space V. If $\boldsymbol{u} = \sum_{i=1}^{n} a_i \boldsymbol{u}_i$ and $\boldsymbol{v} = \sum_{i=1}^{n} b_i \boldsymbol{u}_i$, then

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} a_i b_i$$

- Proof

By Theorem 5.5.2,

$$\langle \mathbf{v}, \mathbf{u}_i \rangle = b_i \qquad i = 1, \ldots, n$$

Therefore,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \sum_{i=1}^{n} a_i \mathbf{u}_i, \mathbf{v} \right\rangle = \sum_{i=1}^{n} a_i \langle \mathbf{u}_i, \mathbf{v} \rangle = \sum_{i=1}^{n} a_i \langle \mathbf{v}, \mathbf{u}_i \rangle = \sum_{i=1}^{n} a_i b_i$$

Parseval's Formula

- If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ is an orthonormal basis for an inner product space V and $\boldsymbol{v} = \sum_{i=1}^{n} c_i \boldsymbol{u}_i$, then

$$\|\boldsymbol{v}\|^2 = \sum_{i=1}^{n} c_i^2$$

- Proof

If $\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u}_i$, then, by Corollary 5.5.3,

$$\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle = \sum_{i=1}^{n} c_i^2$$

Parseval's Formula

Ex 4. The vectors $\boldsymbol{u}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$ and $\boldsymbol{u}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T$ form an orthonormal basis for $\mathbb{R}^2$

- Solution

The vectors

$$\mathbf{u}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T \quad \text{and} \quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T$$

form an orthonormal basis for $\mathbb{R}^2$. If $\mathbf{x} \in \mathbb{R}^2$, then

$$\mathbf{x}^T\mathbf{u}_1 = \frac{x_1 + x_2}{\sqrt{2}} \quad \text{and} \quad \mathbf{x}^T\mathbf{u}_2 = \frac{x_1 - x_2}{\sqrt{2}}$$

It follows from Theorem 5.5.2 that

$$\mathbf{x} = \frac{x_1 + x_2}{\sqrt{2}}\mathbf{u}_1 + \frac{x_1 - x_2}{\sqrt{2}}\mathbf{u}_2$$

and it follows from Corollary 5.5.4 that

$$\|\mathbf{x}\|^2 = \left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 + \left(\frac{x_1 - x_2}{\sqrt{2}}\right)^2 = x_1^2 + x_2^2$$

Orthogonal Matrices

- An $n \times n$ matrix Q is said to be an *orthogonal matrix* if the column vectors of Q form an orthonormal set in $\mathbb{R}^n$
- An $n \times n$ matrix Q is orthogonal if and only of $Q^T Q = I$
- Proof

It follows from the definition that an $n \times n$ matrix Q is orthogonal if and only if its column vectors satisfy

$$\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij}$$

However, $\mathbf{q}_i^T \mathbf{q}_j$ is the (i, j) entry of the matrix $Q^T Q$. Thus Q is orthogonal if and only if $Q^T Q = I$. ■

- If Q is an orthogonal matrix, then Q is invertible and $Q^{-1} = Q^T$

Orthogonal Matrices

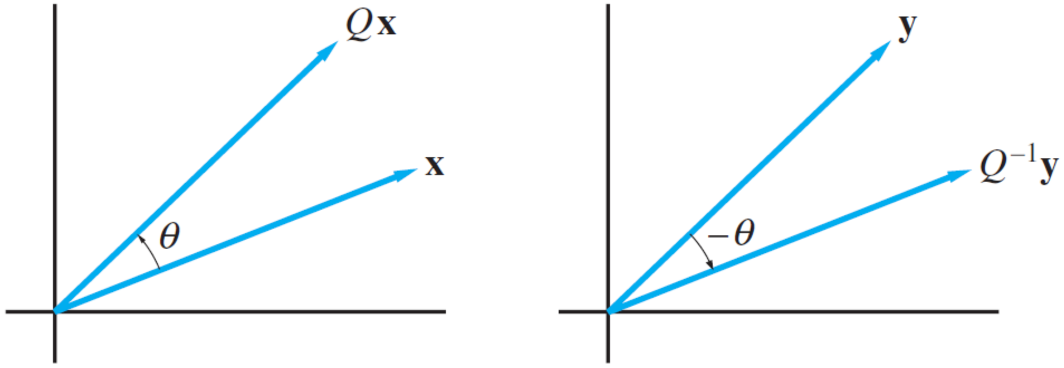
Ex 6. For any fixed θ, the matrix

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is orthogonal and

$$Q^{-1} = Q^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

- The matrix Q can be thought of as a linear transformation from $\mathbb{R}^2$ onto $\mathbb{R}^2$ that has the effect of rotating each vector by an angle θ while leaving the length of the vector unchanged
- Q^{-1} is a rotation by the angle $-\theta$



(a) (b)

Orthogonal Matrices

- In general, inner products are preserved under multiplication by an orthogonal matrix

$$\langle Q\boldsymbol{x}, Q\boldsymbol{y}\rangle = (Q\boldsymbol{y})^T Q\boldsymbol{x} = \boldsymbol{y}^T Q^T Q\boldsymbol{x} = \boldsymbol{y}^T \boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$$

- If $\boldsymbol{x} = \boldsymbol{y}$, then $\|Q\boldsymbol{x}\|^2 = \|\boldsymbol{x}\|^2$ and hence $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$

- Multiplication by an orthogonal matrix preserves the lengths of vectors

- **Properties of Orthogonal Matrices**
  - If Q is an $n \times n$ orthogonal matrix, then

    a. the column vectors of Q forms an orthonormal basis for $\mathbb{R}^n$
    b. $Q^T Q = I$
    c. $Q^T = Q^{-1}$
    d. $\langle Q\boldsymbol{x}, Q\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$
    e. $\|Q\boldsymbol{x}\|^2 = \|\boldsymbol{x}\|^2$

Permutation Matrices

- A *permutation matrix* is a matrix formed from the identity matrix by reordering its columns
  - Permutation matrices are orthogonal matrices

- If P is the permutation matrix formed by reordering the columns of I in the order $(k_1, \cdots, k_n)$, then $P = \left(\boldsymbol{e}_{k_1}, \cdots, \boldsymbol{e}_{k_n}\right)$

- If A is an $m \times n$ matrix, then

$$AP = \left(A\boldsymbol{e}_{k_1}, \cdots, A\boldsymbol{e}_{k_n}\right) = \left(\boldsymbol{a}_{k_1}, \cdots, \boldsymbol{a}_{k_n}\right)$$

- Postmultiplication of A by P reorders the columns of A in the order $(k_1, \cdots, k_n)$

Permutation Matrices

- Since $P = \left(\boldsymbol{e}_{k_1}, \cdots, \boldsymbol{e}_{k_n}\right)$ is orthogonal, it follows that

$$P^{-1} = P^T = \begin{bmatrix} \boldsymbol{e}_{k_1}^T \\ \vdots \\ \boldsymbol{e}_{k_n}^T \end{bmatrix}$$

- The k_1 column of P^T will be $\boldsymbol{e}_1$, the k_2 column will be $\boldsymbol{e}_2$, and so on. Thus, P^T is a permutation matrix

- The matrix P^T can be formed directly from I by reordering its rows in the order $(k_1, \cdots, k_n)$

- In general, a permutation matrix can be formed from I by reordering either its rows or its columns

- In general, if P is an $n \times n$ permutation matrix
  - premultiplication of an $n \times r$ matrix B by P reorders the rows of B
  - postmultiplication of an $m \times n$ matrix A by P reorders the columns of A

Orthonormal Sets and Least Squares

- If the column vectors of A form an orthonormal set of vectors in $\mathbb{R}^m$, then $A^TA = I$ and the solution to the least squares problem $A\boldsymbol{x} = \boldsymbol{b}$ is

$$\widehat{\boldsymbol{x}} = A^T\boldsymbol{b}$$

- Proof

The (i, j) entry of A^TA is formed from the ith row of A^T and the jth column of A. Thus, the (i, j) entry is actually the scalar product of the ith and jth columns of A. Since the column vectors of A are orthonormal, it follows that

$$A^TA = (\delta_{ij}) = I$$

Consequently, the normal equations simplify to

$$\widehat{\boldsymbol{x}} = (A^TA)^{-1}A^T\boldsymbol{b} = A^T\boldsymbol{b} \qquad \blacksquare$$

Orthonormal Sets and Least Squares

- If we have an orthonormal basis for $R(A)$, the projection $\boldsymbol{p} = A\hat{\boldsymbol{x}}$ can be determined in terms of the basis elements

- Let S be a subspace of an inner product space V and let $\boldsymbol{x} \in V$. Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ be an orthonormal basis for S. If $\boldsymbol{p} = \sum_{i=1}^{n} c_i \boldsymbol{u}_i$ where $c_i = \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle$ for each i, then $\boldsymbol{p} - \boldsymbol{x} \in S^{\perp}$

- Proof

We will show first that $(\mathbf{p} - \mathbf{x}) \perp \mathbf{u}_i$ for each i:

$$\begin{aligned}
\langle \mathbf{u}_i, \mathbf{p} - \mathbf{x} \rangle &= \langle \mathbf{u}_i, \mathbf{p} \rangle - \langle \mathbf{u}_i, \mathbf{x} \rangle \\
&= \langle \mathbf{x}_i, \sum_{j=1}^{n} c_j \mathbf{u}_j \rangle - c_i \\
&= \sum_{j=1}^{n} c_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle - c_i \\
&= 0
\end{aligned}$$

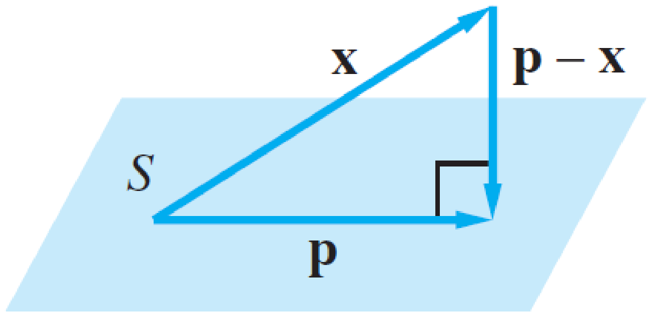
So $\mathbf{p} - \mathbf{x}$ is orthogonal to all the $\mathbf{u}_i$'s. If $\mathbf{y} \in S$, then

$$\mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$$

and hence

$$\langle \mathbf{p} - \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{p} - \mathbf{x}, \sum_{i=1}^{n} \alpha_i \mathbf{u}_i \rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{p} - \mathbf{x}, \mathbf{u}_i \rangle = 0$$

If $\mathbf{x} \in S$, the preceding result is trivial, since, by Theorem 5.5.2, $\mathbf{p} - \mathbf{x} = \mathbf{0}$. If $\mathbf{x} \notin S$, then $\mathbf{p}$ is the element in S closest to $\mathbf{x}$.



Orthonormal Sets and Least Squares

- $\boldsymbol{p}$ is the element of S that is closest to $\boldsymbol{x}$; that is, $\|\boldsymbol{y}-\boldsymbol{x}\| > \|\boldsymbol{p}-\boldsymbol{x}\|$ for any $\boldsymbol{y} \neq \boldsymbol{p}$ in S

- Proof

If $\mathbf{y} \in S$ and $\mathbf{y} \neq \mathbf{p}$, then

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|(\mathbf{y}-\mathbf{p}) + (\mathbf{p}-\mathbf{x})\|^2$$

Since $\mathbf{y}-\mathbf{p} \in S$, it follows from Theorem 5.5.7 and the Pythagorean law that

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

Therefore, $\|\mathbf{y}-\mathbf{x}\| > \|\mathbf{p}-\mathbf{x}\|$.

The vector $\mathbf{p}$ defined by (3) and (4) is said to be the *projection of* $\mathbf{x}$ *onto* S.

Orthonormal Sets and Least Squares

- Let S be a nonzero subspace of $\mathbb{R}^m$ and let $\boldsymbol{b} \in \mathbb{R}^m$. If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ is an orthonormal basis for S and $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]$, then the projection $\boldsymbol{p}$ of $\boldsymbol{b}$ onto S is given by

$$\boldsymbol{p} = UU^T\boldsymbol{b}$$

- Proof

It follows from Theorem 5.5.7 that the projection $\mathbf{p}$ of $\mathbf{b}$ onto S is given by

$$\mathbf{p} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = U\mathbf{c}$$

where

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{b} \\ \mathbf{u}_2^T\mathbf{b} \\ \vdots \\ \mathbf{u}_k^T\mathbf{b} \end{pmatrix} = U^T\mathbf{b}$$

Therefore,

$$\mathbf{p} = UU^T\mathbf{b} \qquad \blacksquare$$

Orthonormal Sets and Least Squares

- If P is a projection matrix corresponding to a subspace of S of $\mathbb{R}^m$, then, for any $\boldsymbol{b} \in \mathbb{R}^m$, the projection $\boldsymbol{p}$ of $\boldsymbol{b}$ onto S is unique

- If Q is also a projection matrix corresponding to S, then

$$Q\boldsymbol{b} = \boldsymbol{p} = P\boldsymbol{b}$$

$$\boldsymbol{q}_j = Q\boldsymbol{e}_j = P\boldsymbol{e}_j = \boldsymbol{p}_j \qquad \text{for } j = 1, \cdots, m$$

  and hence $Q = P$

- The projection matrix for a subspace S of $\mathbb{R}^m$ is unique

Orthonormal Sets and Least Squares

Ex 7. Let S be the set of all vectors in $\mathbb{R}^3$ of the form $(x, y, 0)^T$. Find the vector $\boldsymbol{p}$ in S that is closest to $\boldsymbol{w} = (5, 3, 4)^T$

- Solution

Let $\mathbf{u}_1 = (1, 0, 0)^T$ and $\mathbf{u}_2 = (0, 1, 0)^T$. Clearly, $\mathbf{u}_1$ and $\mathbf{u}_2$ form an orthonormal basis for S. Now,

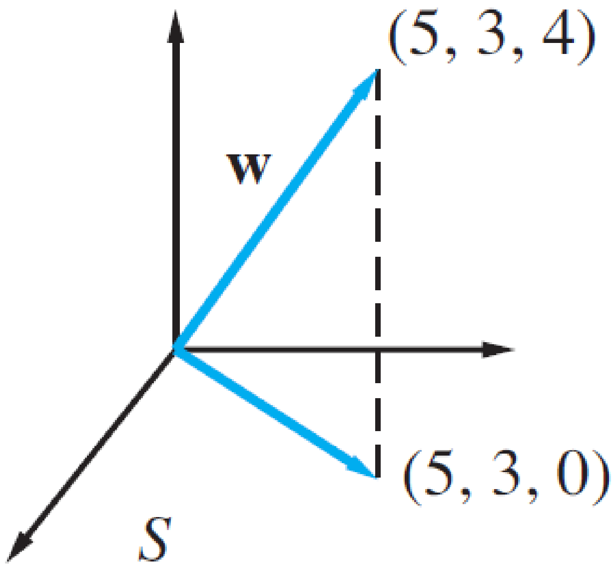
$$c_1 = \mathbf{w}^T\mathbf{u}_1 = 5$$
$$c_2 = \mathbf{w}^T\mathbf{u}_2 = 3$$

The vector $\mathbf{p}$ turns out to be exactly what we would expect:

$$\mathbf{p} = 5\mathbf{u}_1 + 3\mathbf{u}_2 = (5, 3, 0)^T$$

Alternatively, $\mathbf{p}$ could have been calculated using the projection matrix UU^T:

$$\mathbf{p} = UU^T\mathbf{w} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$$ ■



Exercises

12. If Q is an $n \times n$ orthogonal matrix and $\boldsymbol{x}$ and $\boldsymbol{y}$ are nonzero vectors in $\mathbb{R}^n$, then how does the angle θ_2 between $Q\boldsymbol{x}$ and $Q\boldsymbol{y}$ compare with the angle θ_1 between $\boldsymbol{x}$ and $\boldsymbol{y}$?