Section 6.5 Orthogonal Matrices and Operators

Definition

A linear operator on $\mathcal{R}^n$, $T : \mathcal{R}^n \rightarrow \mathcal{R}^n$ is said to be **norm-preserving** if

$$||T(\mathbf{u})|| = ||\mathbf{u}||, \forall \mathbf{u} \in \mathcal{R}^n$$

Example: T : linear operator on $\mathcal{R}^2$ that rotates a vector by θ.

$\Rightarrow T$ is norm-preserving.

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

Example: U : linear operator on $\mathcal{R}^n$ that has an eigenvalue $\lambda \neq \pm 1$.

$\Rightarrow U$ is not norm-preserving, since for the corresponding eigenvector $\mathbf{v}$, $||U(\mathbf{v})|| = ||\lambda \mathbf{v}|| = |\lambda| \cdot ||\mathbf{v}|| \neq ||\mathbf{v}||$.

Necessary conditions for a linear operator to be norm-preserving:

Let $Q = [\ \mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \cdots \ \ \mathbf{q}_n\ ]$ be the standard matrix of the linear operator.

Then (1) $||\mathbf{q}_j|| = ||Q\mathbf{e}_j|| = ||\mathbf{e}_j|| = 1$, and

(2) $||\mathbf{q}_i + \mathbf{q}_j||^2 = ||Q\mathbf{e}_i + Q\mathbf{e}_j||^2 = ||Q(\mathbf{e}_i + \mathbf{e}_j)||^2 = ||\mathbf{e}_i + \mathbf{e}_j||^2 = 2$

$= ||\mathbf{q}_i||^2 + ||\mathbf{q}_j||^2$, i.e., $\mathbf{q}_i$ and $\mathbf{q}_j$ are orthogonal.

Definitions

1. An $n \times n$ matrix Q is called an **orthogonal matrix** (or simply **orthogonal**) if the columns of Q form an **orthonormal basis** for $\mathcal{R}^n$.
2. A linear operator T on $\mathcal{R}^n$ is called an **orthogonal operator** (or simply **orthogonal**) if its standard matrix is an **orthogonal matrix**.

Example: $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is an orthogonal matrix.

Question: What are the sufficient and necessary conditions for Q to be an **orthogonal matrix**?

Theorem 6.9

The following conditions about an $n \times n$ matrix Q are equivalent:

(a) Q is orthogonal.

(b) $Q^T Q = I_n$.

(c) Q is invertible and $Q^{-1} = Q^T$.

(d) $Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$. (i.e., Q preserves dot products.)

(e) $||Q\mathbf{u}|| = ||\mathbf{u}||$ for any $\mathbf{u}$ in $\mathcal{R}^n$. (Q preserves norms.)

Proof (b) ⇔ (c) By definition of invertible matrices

(a) ⇒ (b) with $Q = [\ \mathbf{q}_1\ \ \mathbf{q}_2\ \ \cdots\ \ \mathbf{q}_n\ ]$, $\mathbf{q}_i \cdot \mathbf{q}_i = 1 = [Q^T Q]_{ii}\ \forall i$,

and $\mathbf{q}_i \cdot \mathbf{q}_j = 0 = [Q^T Q]_{ij}\ \forall i \neq j$.

$\Rightarrow Q^T Q = I_n$

(c) ⇒ (d) $\forall \mathbf{u}, \mathbf{v} \in \mathcal{R}^n$, $Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot Q^T Q\mathbf{v} = \mathbf{u} \cdot Q^{-1} Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$.

(d) ⇒ (e) $\forall \mathbf{u} \in \mathcal{R}^n$, $\|Q\mathbf{u}\| = (Q\mathbf{u} \cdot Q\mathbf{u})^{1/2} = (\mathbf{u} \cdot \mathbf{u})^{1/2} = \|\mathbf{u}\|$.

(e) ⇒ (a) The above necessary conditions.

Corollary

(a) Q is orthogonal if and only if the rows of Q form an **orthonormal basis** of $\mathcal{R}^n$, or equivalently, $QQ^T = I_n$.
(b) Q is orthogonal if and only if Q^T is orthogonal.

Proof $Q^TQ = I_n \Leftrightarrow QQ^T = I_n \Leftrightarrow Q^T = Q^{-1} \Leftrightarrow (Q^T)^TQ^T = I_n$.

Question: If Q is an **orthogonal matrix**, what is its determinant?

Theorem 6.10

Let P and Q be $n \times n$ orthogonal matrices.

(a) $\det Q = \pm 1$.

(b) PQ is an orthogonal matrix.

(c) Q^{-1} is an orthogonal matrix.

(d) Q^T is an orthogonal matrix.

Proof (a) $QQ^T = I_n \Rightarrow 1 = \det(I_n) = \det(QQ^T) = \det(Q)\det(Q^T)$

$= \det(Q)^2. \Rightarrow \det(Q) = \pm 1.$

(b) $(PQ)^T = Q^T P^T = Q^{-1} P^{-1} = (PQ)^{-1}.$

(c) $(Q^{-1})^T = (Q^T)^{-1} = (Q^{-1})^{-1}.$

The above results re-stated in terms of linear operators:

If T is a linear operator on $\mathcal{R}^n$, then then following statements are equivalent.
(a) T is an orthogonal operator.
(b) $T(\mathbf{u}) \cdot T(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ for all $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$. (T preserves **dot products**).
(c) $||T(\mathbf{u})|| = ||\mathbf{u}||$ for all $\mathbf{u}$ in $\mathcal{R}^n$. (T preserves **norms**.)

If T and U are orthogonal operators on $\mathcal{R}^n$, then TU and T^{-1} are orthogonal operators on $\mathcal{R}^n$.

Example: Find an orthogonal operator T on $\mathcal{R}^3$ such that

$$T\left(\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix}^T\right) = [0\ 1\ 0]^T$$

Let $\mathbf{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix}^T$ and the standard matrix of T be A.

$\Rightarrow A$ is orthogonal and $\mathbf{v} = I_n\mathbf{v} = A^T A\mathbf{v} = A^T\mathbf{e}_2$.

$\Rightarrow$ The second row of A is $\mathbf{v}$.

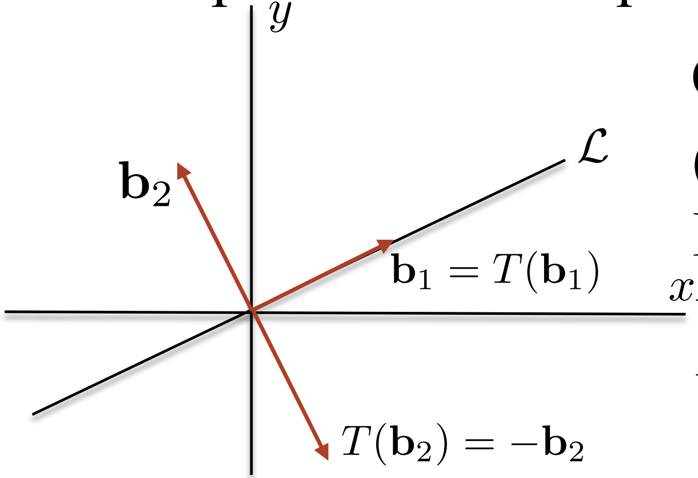
$\Rightarrow$ The other two rows of A are orthogonal to $\mathbf{v}$.

$\Rightarrow \{\mathbf{v}\}^{\perp} = \left\{\mathbf{x} : \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \mathbf{x} = 0\right\}$, which has an orthonormal basis

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

$\Rightarrow A = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$ is an acceptable matrix.

Example: reflection operator T about a line $\mathcal{L}$ passing the origin.



Question: Is T an orthogonal operator?

(An easier) Question:

Is T orthogonal if $\mathcal{L}$ is the x-axis?

$\mathbf{b}_1$ is a unit vector along $\mathcal{L}$.

$\mathbf{b}_2$ is a unit vector perpendicular to $\mathcal{L}$.

$P = [\ \mathbf{b}_1\ \ \mathbf{b}_2\ ]$ is an orthogonal matrix.

$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ is an orthonormal basis of $\mathcal{R}^2$.

$[T]_\mathcal{B}$ = diag[1 -1] is an orthogonal matrix.

Let the standard matrix of T be Q. Then $[T]_\mathcal{B} = P^{-1}QP$, or $Q = P[T]_\mathcal{B}P^{-1} \Rightarrow Q$ is an orthogonal matrix. $\Rightarrow T$ is an orthogonal operator.

Homework Set for Section 6.5

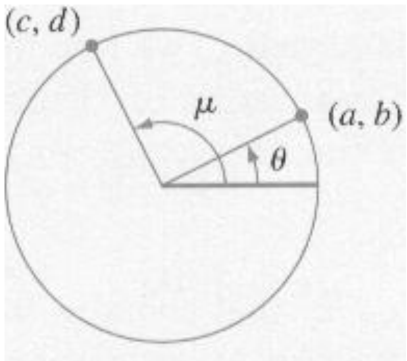
- Section 6.5: Problems 2, 4, 5, 8, 37, 38, 45, 47, 53

(*) Theorem 6.11

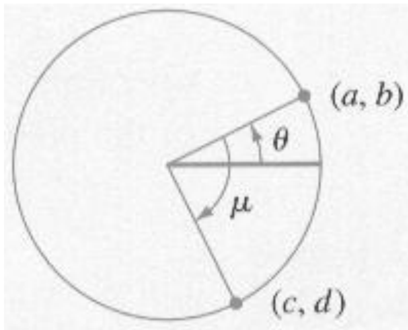
Let T be an orthogonal linear operator on $\mathcal{R}^2$ with standard matrix Q. Then the following statements are true.
(a) If $\det Q = 1$, then T is a rotation.
(b) If $\det Q = -1$, then T is a reflection.

Proof $Q = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \Rightarrow a^2 + b^2 = 1$ and $c^2 + d^2 = 1$

$\Rightarrow a = \cos\theta$, $b = \sin\theta$, $c = \cos\mu$, and $d = \sin\mu$ for some θ and μ.
$\Rightarrow \mu = \theta \pm 90^{\circ}$, since $[\,a \;\; b\,]^T$ and $[\,c \;\; d\,]^T$ are orthogonal



Case 1: $\mu = \theta + 90^{\circ}$



Case 2: $\mu = \theta - 90^{\circ}$

Case 1: $\mu = \theta + 90^\circ$.

$\Rightarrow \cos\mu = \cos(\theta + 90^\circ) = -\sin\theta$ and $\sin\mu = \sin(\theta + 90^\circ) = \cos\theta$.

$\Rightarrow Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, the rotation matrix A_θ,

and $\det Q = \cos^2\theta + \sin^2\theta = 1$.

Case 2: $\mu = \theta - 90^\circ$.

$\Rightarrow \cos\mu = \cos(\theta - 90^\circ) = \sin\theta$ and $\sin\mu = \sin(\theta - 90^\circ) = -\cos\theta$.

$\Rightarrow Q = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \Rightarrow \det Q = -\cos^2\theta - \sin^2\theta = -1$.

also, $\det(Q - tI_2) = (\cos\theta - t)(-\cos\theta - t) - \sin^2\theta$

$= t^2 - \cos^2\theta - \sin^2\theta = t^2 - 1$.

$\Rightarrow$ eigenvalues of Q are 1 and -1.

Let $\mathbf{b}_1$ and $\mathbf{b}_2$ be eigenvectors of T and $T(\mathbf{b}_1) = \mathbf{b}_1$, $T(\mathbf{b}_2) = -\mathbf{b}_2$.

$\Rightarrow \mathbf{b}_1 \cdot \mathbf{b}_2 = T(\mathbf{b}_1) \cdot T(\mathbf{b}_2) = \mathbf{b}_1 \cdot (-\mathbf{b}_2) = -\mathbf{b}_1 \cdot \mathbf{b}_2$. $\Rightarrow \mathbf{b}_1 \cdot \mathbf{b}_2 = 0$.

$\Rightarrow$ T is a reflection about a line along the direction of $\mathbf{b}_1$.

Example:

$$Q = \begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix} \Rightarrow QQ^T = I_2 \text{ and } \det Q = -1.$$

$\Rightarrow Q$ is the standard matrix of a reflection operator.

The eigenvector **x** corresponding to the eigenvalue 1 satisfies

$$(Q - I_2)\mathbf{x} = \mathbf{0}$$

$\Rightarrow \mathbf{x} = [\ 2 \quad 1\ ]^T$ is a solution, and the reflection is with respect to the line $2y = x$ or $y = (1/2)x$.

Example:

$$Q = \begin{bmatrix} -0.6 & 0.8 \\ -0.8 & -0.6 \end{bmatrix} \Rightarrow QQ^T = I_2 \text{ and } \det Q = 1.$$

$\Rightarrow Q = A_\theta$ is the standard matrix of a rotation operator, where $\cos\theta = -0.6$ and $\sin\theta = -0.8$.

$\Rightarrow \theta = 180^\circ + \cos^{-1}(0.6) \approx 233.2^\circ.$

(*) Theorem 6.12

Let T and U be orthogonal operators on $\mathcal{R}^2$. Then the following statements are true.

(a) If both T and U are reflections, then TU is a rotation.

(b) If one of T or U is a reflection and the other is a rotation, then TU is a reflection.

Proof Let P and Q be the standard matrices of T and U, respectively.

$\Rightarrow PQ$ is the standard matrices of TU, and $(PQ)^T PQ = Q^T P^T PQ = I_n$.

(a) $\det PQ = (\det P)(\det Q) = (-1)(-1) = 1$.

(b) $\det PQ = (\det P)(\det Q) = -1$.

(*) Definition

A function $F : \mathcal{R}^n \Rightarrow \mathcal{R}^n$ is a **rigid motion** if F preserves distance between vectors, i.e., $||F(\mathbf{u}) - F(\mathbf{v})|| = ||\mathbf{u} - \mathbf{v}||$ for all $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$.

Example: $F_{\mathbf{b}} : \mathcal{R}^n \rightarrow \mathcal{R}^n$ with $F_{\mathbf{b}}(\mathbf{v}) = \mathbf{v} + \mathbf{b}$, i.e., the translation by $\mathbf{b}$ is a rigid motion, and it is a linear operator if and only if $\mathbf{b} = \mathbf{0}$.

Example: Any orthogonal operator $T : \mathcal{R}^n \rightarrow \mathcal{R}^n$ is a rigid motion, since $||T(\mathbf{u}) - T(\mathbf{v})|| = ||T(\mathbf{u} - \mathbf{v})|| = ||\mathbf{u} - \mathbf{v}||$ for all $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$.

(*) Theorem 6.13

Let $T : \mathcal{R}^n \Rightarrow \mathcal{R}^n$ be a **rigid motion** such that $T(\mathbf{0}) = \mathbf{0}$.

(a) $||T(\mathbf{u})|| = ||\mathbf{u}||$ for all $\mathbf{u} \in \mathcal{R}^n$.

(b) $T(\mathbf{u}) \cdot T(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$.

(c) T is linear.

(d) T is an orthogonal operator.

Proof (a) $||T(\mathbf{u}) - T(\mathbf{0})|| = ||T(\mathbf{u}) - \mathbf{0}|| = ||\mathbf{u} - \mathbf{0}|| = ||\mathbf{u}||$ for all $\mathbf{u} \in \mathcal{R}^n$.

(b) $||T(\mathbf{u}) - T(\mathbf{v})||^2 = ||T(\mathbf{u})||^2 - 2T(\mathbf{u}) \cdot T(\mathbf{v}) + ||T(\mathbf{v})||^2 = ||\mathbf{u} - \mathbf{v}||^2$
$= ||\mathbf{u}||^2 - 2\mathbf{u} \cdot \mathbf{v} + ||\mathbf{v}||^2$.

(c) $||T(\mathbf{u}+\mathbf{v}) - T(\mathbf{u}) - T(\mathbf{v})||^2 = ||T(\mathbf{u}+\mathbf{v})||^2 + ||T(\mathbf{u})||^2 + ||T(\mathbf{v})||^2$
$- 2T(\mathbf{u}+\mathbf{v}) \cdot T(\mathbf{u}) - 2T(\mathbf{u}+\mathbf{v}) \cdot T(\mathbf{v}) + 2T(\mathbf{u}) \cdot T(\mathbf{v})$
$= ||\mathbf{u}+\mathbf{v}||^2 + ||\mathbf{u}||^2 + ||\mathbf{v}||^2 - 2(\mathbf{u}+\mathbf{v}) \cdot \mathbf{u} - 2(\mathbf{u}+\mathbf{v}) \cdot \mathbf{v} + 2\mathbf{u} \cdot \mathbf{v}$
$= \cdots = 0$, and similarly $||T(c\mathbf{u}) - cT(\mathbf{u})||^2 = 0$.

(d) implied by (a) and (c).

For any rigid motion F on $\mathcal{R}^n$, define a function $T : \mathcal{R}^n \rightarrow \mathcal{R}^n$ with $T(\mathbf{v}) = F(\mathbf{v}) - F(\mathbf{0})$. Then T is also a rigid motion and $T(\mathbf{0}) = \mathbf{0}$, so T is an orthogonal operator. Thus $F(\mathbf{v}) = F_{\mathbf{b}}T(\mathbf{v})$ with $\mathbf{b} = F(\mathbf{0})$.

(*) Definition

For $Q \in \mathcal{C}^{n\times n}$, Q is called an **orthogonal matrix** if $Q^TQ = QQ^T = I_n$, and Q is called a **unitary matrix** if $Q^HQ = QQ^H = I_n$.

Theorem 6.9'

For $Q \in \mathcal{C}^{n\times n}$, the following conditions are equivalent:

(a) Q is unitary. (i.e., $Q^HQ = I_n$)

(b) Columns of Q form an orthonormal basis for $\mathcal{C}^n$.

(c) Q is invertible and $Q^{-1} = Q^H$.

(d) $Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{C}^n$. (i.e., Q preserves dot products.)

(e) $||Q\mathbf{u}|| = ||\mathbf{u}||$ for any $\mathbf{u}$ in $\mathcal{C}^n$. (Q preserves norms.)

Proof (a) $\Leftrightarrow$ (b) with $Q = [\ \mathbf{q}_1\ \ \mathbf{q}_2\ \ \cdots\ \ \mathbf{q}_n\ ]$, $[Q^HQ]_{ii} = \mathbf{q}_i \cdot \mathbf{q}_i = 1\ \forall i$,

and $[Q^HQ]_{ij} = \mathbf{q}_j \cdot \mathbf{q}_i = 0\ \forall i \neq j$.

(a) $\Rightarrow$ (c) by the **Invertible Matrix Theorem**.

(c) $\Rightarrow$ (d) $\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}^n$, $Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot Q^HQ\mathbf{v} = \mathbf{u} \cdot Q^{-1}Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$.

(d) $\Rightarrow$ (e) $\forall \mathbf{u} \in \mathcal{R}^n$, $\|Q\mathbf{u}\| = (Q\mathbf{u} \cdot Q\mathbf{u})^{1/2} = (\mathbf{u} \cdot \mathbf{u})^{1/2} = \|\mathbf{u}\|$.

(d) $\Rightarrow$ (a) $\|\mathbf{q}_j\| = \|Q\mathbf{e}_j\| = \|\mathbf{e}_j\| = 1\ \forall i$, and

$\mathbf{q}_i \cdot \mathbf{q}_j = 0\ \forall i \neq j$ as shown below:

$$\|\mathbf{q}_i\|^2 + \|\mathbf{q}_j\|^2 + 2\mathrm{Re}\{\mathbf{q}_i \cdot \mathbf{q}_j\} = \|\mathbf{q}_i + \mathbf{q}_j\|^2$$

$$= \|Q(\mathbf{e}_i + \mathbf{e}_j)\|^2 = \|\mathbf{e}_i + \mathbf{e}_j\|^2 = 2 = \|\mathbf{q}_i\|^2 + \|\mathbf{q}_j\|^2$$

$$\Rightarrow \mathrm{Re}\{\mathbf{q}_i \cdot \mathbf{q}_j\} = 0,$$

$$\|\mathbf{q}_j\|^2 + \|\mathbf{q}_i\|^2 - 2\mathrm{Im}\{\mathbf{q}_i \cdot \mathbf{q}_j\} = \|\mathbf{q}_j + i\mathbf{q}_i\|^2$$

$$= \|Q(\mathbf{e}_j + \mathrm{i}\mathbf{e}_i)\|^2 = \|\mathbf{e}_j + i\mathbf{e}_i\|^2 = 2 = \|\mathbf{q}_j\|^2 + \|\mathbf{q}_i\|^2$$

$$\Rightarrow \mathrm{Im}\{\mathbf{q}_i \cdot \mathbf{q}_j\} = 0, \text{ where } i^2 = -1.$$

(*) Corollary

(a) $Q \in \mathcal{C}^{n\times n}$ is unitary if and only if the rows of Q form an orthonormal basis of $\mathcal{C}^n$.

(b) Q is unitary if and only if Q^H is unitary.

(*) Theorem 6.10 (Complex case)

Let P and Q be $n \times n$ unitary matrices. Then

(a) $|\det(Q)| = 1$.

(b) PQ is a unitary matrix in $\mathcal{C}^{n\times n}$.

(c) Q^{-1} is a unitary matrix in $\mathcal{C}^{n\times n}$.