Theory of Matrices

Chapter 3. Basic Properties of Matrices

3.7. Orthogonality—Proofs of Theorems

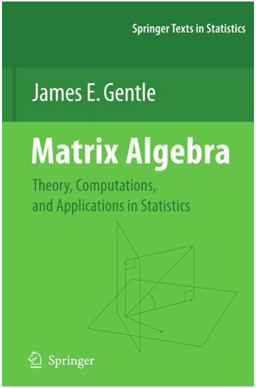

Table of contents

Theorem 3.7.1

Theorem 3.7.1. Let Q be an $n \times m$ matrix. For $n \leq m$, Q is orthogonal if and only if $QQ^T = I_n$. For $n \geq m$, Q is orthogonal if and only if $Q^T Q = I_m$. A square matrix Q is orthogonal if and only if $QQ^T = Q^T Q = I$ (so a square matrix Q is orthogonal if and only if it is invertible and $Q^{-1} = Q^T$).

Proof. First, suppose $n \leq m$. If Q is orthogonal then the row rank of Q equals the column rank of Q by Theorem 3.3.2 and so Q must have n orthonormal rows (since it cannot have more orthonormal [and hence linearly independent] columns than rows). Let the rows of Q be $r_1, r_2, \ldots, r_n$. Then the columns of Q^T are $r_1^T, r_2^T, \ldots, r_n^T$.

Theorem 3.7.1

Theorem 3.7.1. Let Q be an $n \times m$ matrix. For $n \leq m$, Q is orthogonal if and only if $QQ^T = I_n$. For $n \geq m$, Q is orthogonal if and only if $Q^TQ = I_m$. A square matrix Q is orthogonal if and only if $QQ^T = Q^TQ = I$ (so a square matrix Q is orthogonal if and only if it is invertible and $Q^{-1} = Q^T$).

Proof. First, suppose $n \leq m$. If Q is orthogonal then the row rank of Q equals the column rank of Q by Theorem 3.3.2 and so Q must have n orthonormal rows (since it cannot have more orthonormal [and hence linearly independent] columns than rows). Let the rows of Q be $r_1, r_2, \ldots, r_n$. Then the columns of Q^T are $r_1^T, r_2^T, \ldots, r_n^T$. So the (i, j) entry of QQ^T is $r_i r_j^T = \langle r_i, r_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ since the rows of Q are orthonormal. Since QQ^T is $n \times n$, then $QQ^T = I_n$.

Theorem 3.7.1

Theorem 3.7.1. Let Q be an $n \times m$ matrix. For $n \leq m$, Q is orthogonal if and only if $QQ^T = I_n$. For $n \geq m$, Q is orthogonal if and only if $Q^TQ = I_m$. A square matrix Q is orthogonal if and only if $QQ^T = Q^TQ = I$ (so a square matrix Q is orthogonal if and only if it is invertible and $Q^{-1} = Q^T$).

Proof. First, suppose $n \leq m$. If Q is orthogonal then the row rank of Q equals the column rank of Q by Theorem 3.3.2 and so Q must have n orthonormal rows (since it cannot have more orthonormal [and hence linearly independent] columns than rows). Let the rows of Q be $r_1, r_2, \ldots, r_n$. Then the columns of Q^T are $r_1^T, r_2^T, \ldots, r_n^T$. So the (i,j) entry of QQ^T is $r_i r_j^T = \langle r_i, r_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ since the rows of Q are orthonormal. Since QQ^T is $n \times n$, then $QQ^T = I_n$. Conversely, if $QQ^T = I_n$ then the (i,j) entry of QQ^T is $\langle r_i, r_j \rangle$ as given above and so the rows of Q form an orthonormal set and Q is orthogonal.

Theorem 3.7.1

Theorem 3.7.1. Let Q be an $n \times m$ matrix. For $n \leq m$, Q is orthogonal if and only if $QQ^T = I_n$. For $n \geq m$, Q is orthogonal if and only if $Q^TQ = I_m$. A square matrix Q is orthogonal if and only if $QQ^T = Q^TQ = I$ (so a square matrix Q is orthogonal if and only if it is invertible and $Q^{-1} = Q^T$).

Proof. First, suppose $n \leq m$. If Q is orthogonal then the row rank of Q equals the column rank of Q by Theorem 3.3.2 and so Q must have n orthonormal rows (since it cannot have more orthonormal [and hence linearly independent] columns than rows). Let the rows of Q be $r_1, r_2, \ldots, r_n$. Then the columns of Q^T are $r_1^T, r_2^T, \ldots, r_n^T$. So the (i,j) entry of QQ^T is $r_i r_j^T = \langle r_i, r_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ since the rows of Q are orthonormal. Since QQ^T is $n \times n$, then $QQ^T = I_n$. Conversely, if $QQ^T = I_n$ then the (i,j) entry of QQ^T is $\langle r_i, r_j \rangle$ as given above and so the rows of Q form an orthonormal set and Q is orthogonal.

Theorem 3.7.1 (continued)

Proof (continued). Second, suppose $n \geq m$. If Q is orthogonal then, similar to the case $n \leq m$, it must be that Q has m orthonormal columns. Let the columns of Q be $c_1, c_2, \ldots, c_m$. Then the rows of Q^T are $c_1^T, c_2^T, \ldots, c_m^T$. So the (i,j) entry of Q^TQ is $c_i^Tc_j = \langle c_i, c_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ since the columns of Q are orthonormal. Since Q^TQ is $m \times m$ then $Q^TQ = I_m$. Conversely, if $Q^TQ = I_m$ then the (i,j) entry of Q^TQ is $\langle c_i, c_j \rangle$ as given above and so the columns of Q form an orthonormal set and Q is orthogonal.

Theorem 3.7.1 (continued)

Proof (continued). Second, suppose $n \geq m$. If Q is orthogonal then, similar to the case $n \leq m$, it must be that Q has m orthonormal columns. Let the columns of Q be $c_1, c_2, \ldots, c_m$. Then the rows of Q^T are $c_1^T, c_2^T, \ldots, c_m^T$. So the (i, j) entry of $Q^T Q$ is $c_i^T c_j = \langle c_i, c_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ since the columns of Q are orthonormal. Since $Q^T Q$ is $m \times m$ then $Q^T Q = I_m$. Conversely, if $Q^T Q = I_m$ then the (i, j) entry of $Q^T Q$ is $\langle c_i, c_j \rangle$ as given above and so the columns of Q form an orthonormal set and Q is orthogonal.

If Q is $n \times n$, then combining the first two cases we have that Q is orthogonal if and only if $QQ^T = I = Q^T Q$. □

Theorem 3.7.1 (continued)

Proof (continued). Second, suppose $n \geq m$. If Q is orthogonal then, similar to the case $n \leq m$, it must be that Q has m orthonormal columns. Let the columns of Q be $c_1, c_2, \ldots, c_m$. Then the rows of Q^T are $c_1^T, c_2^T, \ldots, c_m^T$. So the (i, j) entry of $Q^T Q$ is
$c_i^T c_j = \langle c_i, c_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ since the columns of Q are orthonormal.
Since $Q^T Q$ is $m \times m$ then $Q^T Q = I_m$. Conversely, if $Q^T Q = I_m$ then the (i, j) entry of $Q^T Q$ is $\langle c_i, c_j \rangle$ as given above and so the columns of Q form an orthonormal set and Q is orthogonal.

If Q is $n \times n$, then combining the first two cases we have that Q is orthogonal if and only if $QQ^T = I = Q^T Q$. □

Corollary 3.7.2

Corollary 3.7.2. For Q a square orthogonal matrix, we have $\det(Q) = \pm 1$. For Q an $n \times m$ orthogonal matrix Q with $n \geq m$, we have $\langle Q, Q \rangle = m$.

Proof. By Theorem 3.7.1 and Theorem 3.2.4, $\det(QQ^T) = \det(I)$ or $\det(Q)\det(Q^T) = 1$. By Theorem 3.1.A, $\det(Q^T) = \det(Q)$, so $\det(Q)^2 = 1$ and $\det(Q) = \pm 1$.

Corollary 3.7.2

Corollary 3.7.2. For Q a square orthogonal matrix, we have $\det(Q) = \pm 1$. For Q an $n \times m$ orthogonal matrix Q with $n \geq m$, we have $\langle Q, Q \rangle = m$.

Proof. By Theorem 3.7.1 and Theorem 3.2.4, $\det(QQ^T) = \det(I)$ or $\det(Q)\det(Q^T) = 1$. By Theorem 3.1.A, $\det(Q^T) = \det(Q)$, so $\det(Q)^2 = 1$ and $\det(Q) = \pm 1$.

Let the columns of $n \times m$ orthogonal Q be $c_1, c_2, \ldots, c_m$. Then

$$\langle Q, Q \rangle = \sum_{j=1}^{m} c_j^T c_j = \sum_{j=1}^{m} \langle c_j, c_j \rangle = \sum_{j=1}^{m} \|c_j\|^2 = m$$

since the columns of Q form an orthonormal set of vectors. □

Corollary 3.7.2

Corollary 3.7.2. For Q a square orthogonal matrix, we have $\det(Q) = \pm 1$. For Q an $n \times m$ orthogonal matrix Q with $n \geq m$, we have $\langle Q, Q \rangle = m$.

Proof. By Theorem 3.7.1 and Theorem 3.2.4, $\det(QQ^T) = \det(I)$ or $\det(Q)\det(Q^T) = 1$. By Theorem 3.1.A, $\det(Q^T) = \det(Q)$, so $\det(Q)^2 = 1$ and $\det(Q) = \pm 1$.

Let the columns of $n \times m$ orthogonal Q be $c_1, c_2, \ldots, c_m$. Then

$$\langle Q, Q \rangle = \sum_{j=1}^{m} c_j^T c_j = \sum_{j=1}^{m} \langle c_j, c_j \rangle = \sum_{j=1}^{m} \|c_j\|^2 = m$$

since the columns of Q form an orthonormal set of vectors. □

Theorem 3.7.3

Theorem 3.7.3. Every permutation matrix is orthogonal.

Proof. First, form the elementary permutation matrix E_{pq} from I_n by interchanging rows p and q of I_n, $I_n \stackrel{R_p \leftrightarrow R_q}{\longrightarrow} E_{pq} = [e_{ij}]$. So we have $e_{ij} = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii} = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{pj} = 0$ for $j \neq q$ and $e_{pq} = 1$. For $i = q$ we have $e_{qj} = 0$ for $j \neq p$ and $e_{qp} = 1$.

Theorem 3.7.3

Theorem 3.7.3. Every permutation matrix is orthogonal.

Proof. First, form the elementary permutation matrix E_{pq} from I_n by interchanging rows p and q of I_n, $I_n \overset{R_p \leftrightarrow R_q}{\longrightarrow} E_{pq} = [e_{ij}]$. So we have $e_{ij} = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii} = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{pj} = 0$ for $j \neq q$ and $e_{pq} = 1$. For $i = q$ we have $e_{qj} = 0$ for $j \neq p$ and $e_{qp} = 1$. So in $E_{pq}^T = [e_{ij}^t]$ we have $e_{ij}^t = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii}^t = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{jp}^t = e_{pj} = 0$ for $j \neq q$ and $e_{qp}^t = e_{pq} = 1$. For $i = q$ we have $e_{jq}^t = e_{qj} = 0$ for $j \neq p$ and $e_{pq}^t = e_{qp} = 1$.

Theorem 3.7.3

Theorem 3.7.3. Every permutation matrix is orthogonal.

Proof. First, form the elementary permutation matrix E_{pq} from I_n by interchanging rows p and q of I_n, $I_n \overset{R_p \leftrightarrow R_q}{\longrightarrow} E_{pq} = [e_{ij}]$. So we have $e_{ij} = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii} = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{pj} = 0$ for $j \neq q$ and $e_{pq} = 1$. For $i = q$ we have $e_{qj} = 0$ for $j \neq p$ and $e_{qp} = 1$. So in $E_{pq}^T = [e_{ij}^t]$ we have $e_{ij}^t = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii}^t = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{jp}^t = e_{pj} = 0$ for $j \neq q$ and $e_{qp}^t = e_{pq} = 1$. For $i = q$ we have $e_{jq}^t = e_{qj} = 0$ for $j \neq p$ and $e_{pq}^t = e_{qp} = 1$. So the (i, j) entry of $E_{pq}E_{pq}^T$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ is

$$\begin{aligned} \sum_{k=1}^{n} e_{ik} e_{kj}^t &= e_{ii} e_{ij}^t \text{ since } e_{ik} = 0 \text{ for } k \neq i \text{ here} \\ &= \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j, \end{cases} \end{aligned}$$

Theorem 3.7.3

Theorem 3.7.3. Every permutation matrix is orthogonal.

Proof. First, form the elementary permutation matrix E_{pq} from I_n by interchanging rows p and q of I_n, $I_n \overset{R_p \leftrightarrow R_q}{\longrightarrow} E_{pq} = [e_{ij}]$. So we have $e_{ij} = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii} = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{pj} = 0$ for $j \neq q$ and $e_{pq} = 1$. For $i = q$ we have $e_{qj} = 0$ for $j \neq p$ and $e_{qp} = 1$. So in $E_{pq}^T = [e_{ij}^t]$ we have $e_{ij}^t = 0$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ and $i \neq j$, and $e_{ii}^t = 1$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$. For $i = p$ we have $e_{jp}^t = e_{pj} = 0$ for $j \neq q$ and $e_{qp}^t = e_{pq} = 1$. For $i = q$ we have $e_{jq}^t = e_{qj} = 0$ for $j \neq p$ and $e_{pq}^t = e_{qp} = 1$. So the (i, j) entry of $E_{pq}E_{pq}^T$ for $i \in \{1, 2, \ldots, n\} \setminus \{p, q\}$ is

$$\begin{aligned} \sum_{k=1}^{n} e_{ik} e_{kj}^t &= e_{ii} e_{ij}^t \text{ since } e_{ik} = 0 \text{ for } k \neq i \text{ here} \\ &= \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j, \end{cases} \end{aligned}$$

Theorem 3.7.3 (continued 1)

Proof (continued). ...for $i = p$ the (i, j) entry (or the (p, j) entry) is

$$\begin{aligned} \sum_{k=1}^{n} e_{ik} e^t_{kj} = \sum_{k=1}^{n} e_{pk} e^t_{kj} &= e_{pq} e^t_{qj} \text{ since } e_{pk} = 0 \text{ for } k \neq q \\ &= e_{pq} e_{jq} = \begin{cases} 1 \text{ if } j = p \\ 0 \text{ if } j \neq p, \end{cases} \end{aligned}$$

and for $j = q$ the (i, j) entry (or the (i, q) entry) is

$$\begin{aligned} \sum_{k=1}^{n} e_{ik} e^t_{kj} = \sum_{k=1}^{n} e_{ik} e^t_{kq} &= e_{ip} e^t_{pq} \text{ since } e^t_{kq} = 0 \text{ for } k \neq p \\ &= e_{ip} e_{qp} = \begin{cases} 1 \text{ if } i = q \\ 0 \text{ if } i \neq q. \end{cases} \end{aligned}$$

That is, the (i, j) entry of $E_{pq} E^T_{pq}$ is 0 if $i \neq j$ and 1 if $i = j$; that is, $E_{pq} E^T_{pq} = I_n$ and E_{pq} is orthogonal by Theorem 3.7.1.

Theorem 3.7.3 (continued 1)

Proof (continued). ...for $i = p$ the (i, j) entry (or the (p, j) entry) is

$$\begin{aligned} \sum_{k=1}^{n} e_{ik} e_{kj}^{t} = \sum_{k=1}^{n} e_{pk} e_{kj}^{t} &= e_{pq} e_{qj}^{t} \text{ since } e_{pk} = 0 \text{ for } k \neq q \\ &= e_{pq} e_{jq} = \begin{cases} 1 \text{ if } j = p \\ 0 \text{ if } j \neq p, \end{cases} \end{aligned}$$

and for $j = q$ the (i, j) entry (or the (i, q) entry) is

$$\begin{aligned} \sum_{k=1}^{n} e_{ik} e_{kj}^{t} = \sum_{k=1}^{n} e_{ik} e_{kq}^{t} &= e_{ip} e_{pq}^{t} \text{ since } e_{kq}^{t} = 0 \text{ for } k \neq p \\ &= e_{ip} e_{qp} = \begin{cases} 1 \text{ if } i = q \\ 0 \text{ if } i \neq q. \end{cases} \end{aligned}$$

That is, the (i, j) entry of $E_{pq}E_{pq}^{T}$ is 0 if $i \neq j$ and 1 if $i = j$; that is, $E_{pq}E_{pq}^{T} = I_n$ and E_{pq} is orthogonal by Theorem 3.7.1.

Theorem 3.7.3 (continued 2)

Theorem 3.7.3. Every permutation matrix is orthogonal.

Proof (continued). Second, if E is an elementary permutation matrix resulting from interchanging columns p and q in I_n, then E is, similarly, an orthogonal matrix.

So if P is a permutation matrix, that is if $P = E_1 E_2 \cdots E_\ell$ for elementary permutation matrices $E_1, E_2, \ldots, E_\ell$ (where these correspond to row or column interchanges) then

$$\begin{aligned} P^T &= (E_1 E_2 \cdots E_\ell)^T = E_\ell^T E_{\ell-1}^T \cdots E_1^T \\ &= E_\ell^{-1} E_{\ell-1}^{-1} \cdots E_1^{-1} \text{ since each } E_i \text{ is orthogonal from above,} \\ &\quad \text{and Theorem 3.7.1} \\ &= (E_1 E_2 \cdots E_\ell)^{-1} = P^{-1}, \end{aligned}$$

and so $PP^T = PP^{-1} = I$. That is, P is orthogonal by Theorem 3.7.1. □

Theorem 3.7.3 (continued 2)

Theorem 3.7.3. Every permutation matrix is orthogonal.

Proof (continued). Second, if E is an elementary permutation matrix resulting from interchanging columns p and q in I_n, then E is, similarly, an orthogonal matrix.

So if P is a permutation matrix, that is if $P = E_1 E_2 \cdots E_\ell$ for elementary permutation matrices $E_1, E_2, \ldots, E_\ell$ (where these correspond to row or column interchanges) then

$$\begin{aligned} P^T &= (E_1 E_2 \cdots E_\ell)^T = E_\ell^T E_{\ell-1}^T \cdots E_1^T \\ &= E_\ell^{-1} E_{\ell-1}^{-1} \cdots E_1^{-1} \text{ since each } E_i \text{ is orthogonal from above,} \\ &\quad \text{ and Theorem 3.7.1} \\ &= (E_1 E_2 \cdots E_\ell)^{-1} = P^{-1}, \end{aligned}$$

and so $PP^T = PP^{-1} = I$. That is, P is orthogonal by Theorem 3.7.1. □