Linear Algebra 2

Lecture #14
Orthogonal matrices

Irena Penev

March 9, 2023

- Some preliminaries

- Some preliminaries

Proposition 1.1

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n\times n}$. If $AB = I_n$, then A and B are both invertible and are each other's inverses.

Proof. Problem 2(b) of HW#9 from Linear Algebra 1 (winter 2022).

- Some preliminaries

Proposition 1.1

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$. If $AB = I_n$, then A and B are both invertible and are each other's inverses.

Proof. Problem 2(b) of HW#9 from Linear Algebra 1 (winter 2022).

Proposition 1.2

Let V be a finite-dimensional vector space over $\mathbb{R}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Set $n := \dim(V)$, and let $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal set in V if and only if $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis of V.

Proof.

- Some preliminaries

Proposition 1.1

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$. If $AB = I_n$, then A and B are both invertible and are each other's inverses.

Proof. Problem 2(b) of HW#9 from Linear Algebra 1 (winter 2022).

Proposition 1.2

Let V be a finite-dimensional vector space over $\mathbb{R}$, equipped with a scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$. Set $n := \dim(V)$, and let $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal set in V if and only if $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis of V.

Proof. "$\Longleftarrow$" is obvious. For "$\Longrightarrow$," note that if $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal set, then it is linearly independent, and so since $\dim(V) = n$, it is a basis of V.

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

Definition

A matrix $Q \in \mathbb{R}^{n\times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

Theorem 2.1

Let $Q \in \mathbb{R}^{n\times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

- First some examples, then a proof.

- I_n and $-I_n$ are orthogonal matrices.

- I_n and $-I_n$ are orthogonal matrices.
- Given $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, the *Householder matrix* is the $n \times n$ matrix

$$H(\mathbf{a}) \quad := \quad I_n - \tfrac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \quad = \quad I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

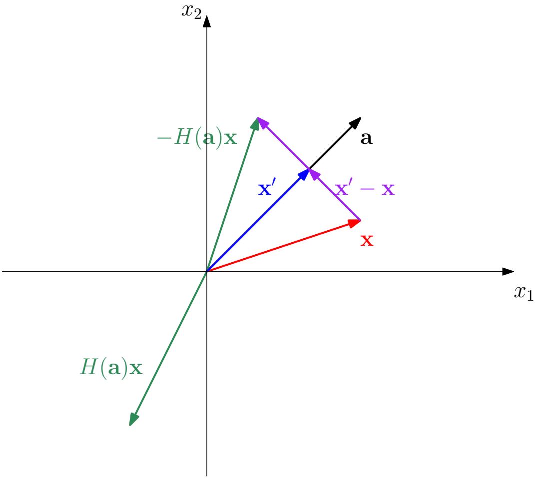
- I_n and $-I_n$ are orthogonal matrices.
- Given $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, the *Householder matrix* is the $n \times n$ matrix

$$H(\mathbf{a}) \;:=\; I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \;=\; I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

To see that $H(\mathbf{a})$ really is an orthogonal matrix, we perform the following simple calculation:

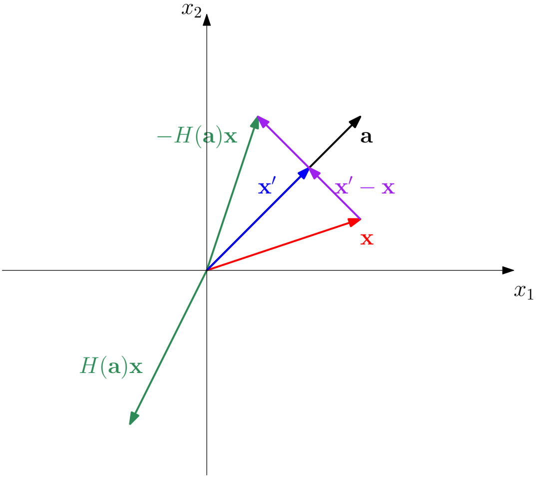
$$\begin{aligned}
H(\mathbf{a})^T\,H(\mathbf{a}) &= (I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T)^T\,(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T) \\
&= (I_n^T - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}(\mathbf{a}\mathbf{a}^T)^T)\,(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T) \\
&= (I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T)\,(I_n - \tfrac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T) \\
&= I_n - \tfrac{4}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T + \tfrac{4}{(\mathbf{a}\cdot\mathbf{a})^2}\mathbf{a}\,\underbrace{\mathbf{a}^T\mathbf{a}}_{=\mathbf{a}\cdot\mathbf{a}}\,\mathbf{a}^T \\
&= I_n - \tfrac{4}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T + \tfrac{4}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T \\
&= I_n.
\end{aligned}$$

- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.



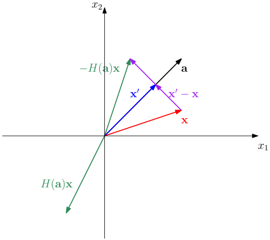
The Householder matrix $H(\mathbf{a})$ itself is the standard matrix of the linear operation that first reflects about the Span($\mathbf{a}$) line and then reflects about the origin.

- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.



The Householder matrix $H(\mathbf{a})$ itself is the standard matrix of the linear operation that first reflects about the Span($\mathbf{a}$) line and then reflects about the origin. Let's check!

- Reminder: $H(\mathbf{a}) := I_n - \frac{2}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T = I_n - \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.



  - The standard matrix of orthogonal projection onto the line Span($\mathbf{a}$) is $\mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T = \mathbf{a}(\mathbf{a}\cdot\mathbf{a})^{-1}\mathbf{a}^T = \frac{1}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T$.
  - If $\mathbf{x}$ is any vector in $\mathbb{R}^n$ and $\mathbf{x}'$ is its projection onto Span($\mathbf{a}$), then the reflection of $\mathbf{x}$ about the line Span($\mathbf{a}$) is

$$\begin{aligned}
\mathbf{x} + 2(\mathbf{x}' - \mathbf{x}) &= 2\mathbf{x}' - \mathbf{x} \\
&= \frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T\mathbf{x} - I_n\mathbf{x} \\
&= (\frac{2}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T - I_n)\mathbf{x} \\
&= -H(\mathbf{a})\mathbf{x}.
\end{aligned}$$

  - $H(\mathbf{a})\mathbf{x}$ is the reflection of $-H(\mathbf{a})\mathbf{x}$ about the origin.

- Given an integer $n \geq 2$, indices $i, j \in \{1, \ldots, n\}$ such that $i < j$, and real numbers c and s such that $c^2 + s^2 = 1$, we define the *Givens matrix* $G_{i,j}(c, s)$ as follows:

$$
G_{i,j}(c,s) = \begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array}
\overset{\begin{array}{ccccccccc} & & & i & & & j & & \end{array}}{
\begin{bmatrix}
1 & 0 & \dots 0 & 0 & 0 & \dots 0 & 0 & 0 & \dots 0 \\
0 & 1 & \dots 0 & 0 & 0 & \dots 0 & 0 & 0 & \dots 0 \\
\vdots & \vdots & \ddots \vdots & \vdots & \vdots & \ddots \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \dots 1 & 0 & 0 & \dots 0 & 0 & 0 & \dots 0 \\
0 & 0 & \dots 0 & c & 0 & \dots 0 & -s & 0 & \dots 0 \\
0 & 0 & \dots 0 & 0 & 1 & \dots 0 & 0 & 0 & \dots 0 \\
\vdots & \vdots & \dots \vdots & \vdots & \vdots & \ddots \vdots & \vdots & \vdots & 0 \\
0 & 0 & \dots 0 & 0 & 0 & \dots 1 & 0 & 0 & \dots 0 \\
0 & 0 & \dots 0 & s & 0 & \dots 0 & c & 0 & \dots 0 \\
0 & 0 & \dots 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \vdots \\
0 & 0 & \dots 0 & 0 & 0 & 0 & 0 & 0 & \dots 1
\end{bmatrix}}
$$

- Given an integer $n \geq 2$, indices $i, j \in \{1, \ldots, n\}$ such that $i < j$, and real numbers c and s such that $c^2 + s^2 = 1$, we define the *Givens matrix* $G_{i,j}(c, s)$ as follows:

$$
G_{i,j}(c,s) \quad = \quad
\begin{array}{c}
\\ \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc}
 & & & & i & & & & j & & & \\
\end{array} \\
\left[
\begin{array}{cccccccccccc}
1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & c & 0 & \ldots & 0 & -s & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & 0 \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & s & 0 & \ldots & 0 & c & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & 0 & & 0 & 0 & 1 & & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & 0 & 0 & & 0 & 0 & 0 & \ldots & 1
\end{array}
\right]
\end{array}
$$

- $G_{i,j}(c, s)$ is orthogonal because its columns form an orthonormal basis of $\mathbb{R}^n$.

- Since $c^2 + s^2 = 1$, we see that there exists a real number (angle in radians) θ such that $c = \cos\theta$ and $s = \sin\theta$.

- Since $c^2 + s^2 = 1$, we see that there exists a real number (angle in radians) θ such that $c = \cos\theta$ and $s = \sin\theta$.
- With this set-up, we see that $G_{i,j}(c, s)$ represents rotation by angle θ in the $x_i x_j$-plane.

- Since $c^2 + s^2 = 1$, we see that there exists a real number (angle in radians) θ such that $c = \cos\theta$ and $s = \sin\theta$.
- With this set-up, we see that $G_{i,j}(c, s)$ represents rotation by angle θ in the $x_i x_j$-plane.
- This is particularly easy to see in the case when $n = 2$. In that case, we have that

$$G_{1,2}(c, s) = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

which is precisely the standard matrix of counterclockwise rotation by angle θ.

Theorem 2.1

Let $Q \in \mathbb{R}^{n\times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^TQ = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof.

Theorem 2.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof. By Proposition 1.1, we have that (a), (b), and (c) are equivalent.

Theorem 2.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof. By Proposition 1.1, we have that (a), (b), and (c) are equivalent. Moreover, since $(Q^T)^T = Q$, we have that (c) and (d) are equivalent. This proves that (a), (b), (c), and (d) are equivalent.

Theorem 2.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). Next, (b) and (d) together imply (e).

Theorem 2.1

Let $Q \in \mathbb{R}^{n\times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). Next, (b) and (d) together imply (e). On the other hand, if (e) holds, then using "(a) $\Longrightarrow$ (b)" applied to Q^{-1}, we see that Q^{-1} is invertible and satisfies $(Q^{-1})^{-1} = (Q^T)^{-1}$. Consequently, $Q^{-1} = Q^T$, and it follows that (b) holds.

Theorem 2.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);
(b) Q is invertible and satisfies $Q^{-1} = Q^T$;
(c) $QQ^T = I_n$;
(d) Q^T is orthogonal;
(e) Q is invertible and Q^{-1} is orthogonal;
(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;
(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Proof (continued). Next, (b) and (d) together imply (e). On the other hand, if (e) holds, then using "(a) $\Longrightarrow$ (b)" applied to Q^{-1}, we see that Q^{-1} is invertible and satisfies $(Q^{-1})^{-1} = (Q^T)^{-1}$. Consequently, $Q^{-1} = Q^T$, and it follows that (b) holds.

So far, we have established that (a), (b), (c), (d), and (e) are equivalent.

Proof (continued). Let us now show that (a) and (f) are equivalent.

Proof (continued). Let us now show that (a) and (f) are equivalent. Set $Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix}$.

Proof (continued). Let us now show that (a) and (f) are equivalent. Set $Q = \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_n \end{bmatrix}$. Then

$$Q^T Q = \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{q}_1 \cdot \mathbf{q}_1 & \mathbf{q}_1 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_1 \cdot \mathbf{q}_n \\ \mathbf{q}_2 \cdot \mathbf{q}_1 & \mathbf{q}_2 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_2 \cdot \mathbf{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n \cdot \mathbf{q}_1 & \mathbf{q}_n \cdot \mathbf{q}_2 & \dots & \mathbf{q}_n \cdot \mathbf{q}_n \end{bmatrix}.$$

So, $Q^T Q = I_n$ if and only if $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ is an orthonormal set. In view of Proposition 1.2, it now follows that (a) and (f) are equivalent. Analogously, (d) and (g) are equivalent. Q.E.D.

Theorem 2.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

- Making new orthogonal matrices out of old ones:

Proposition 2.2

If $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ are orthogonal, then so is their product $Q_1 Q_2$.

Proposition 2.3

Let $Q_1 \in \mathbb{R}^{m\times m}$ and $Q_2 \in \mathbb{R}^{n\times n}$ be orthogonal matrices. Then

$$Q = \left[\begin{array}{c|c} Q_1 & O_{m\times n} \\ \hline O_{n\times m} & Q_2 \end{array}\right]$$

is an orthogonal matrix in $\mathbb{R}^{(m+n)\times(m+n)}$.

Proposition 2.2

If $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ are orthogonal, then so is their product $Q_1 Q_2$.

Proof.

Proposition 2.2

If $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ are orthogonal, then so is their product $Q_1 Q_2$.

Proof. Assume $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ are orthogonal.

Proposition 2.2

If $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ are orthogonal, then so is their product $Q_1 Q_2$.

Proof. Assume $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ are orthogonal. Then $Q_1^T Q_1 = I_n$ and $Q_2^T Q_2 = I_n$, and consequently,

$$(Q_1 Q_2)^T (Q_1 Q_2) = Q_2^T \underbrace{Q_1^T Q_1}_{=I_n} Q_2 = Q_2^T Q_2 = I_n.$$

So, $Q_1 Q_2$ is indeed orthogonal.

Proposition 2.3

Let $Q_1 \in \mathbb{R}^{m\times m}$ and $Q_2 \in \mathbb{R}^{n\times n}$ be orthogonal matrices. Then

$$Q = \left[\begin{array}{c|c} Q_1 & O_{m\times n} \\ \hline O_{n\times m} & Q_2 \end{array}\right]$$

is an orthogonal matrix in $\mathbb{R}^{(m+n)\times(m+n)}$.

Proof. By hypothesis, we have that $Q_1^T Q_1 = I_m$ and $Q_2^T Q_2 = I_n$. We now compute:

$$\begin{aligned}
Q^T Q &= \left[\begin{array}{c|c} Q_1^T & O_{m\times n} \\ \hline O_{n\times m} & Q_2^T \end{array}\right] \left[\begin{array}{c|c} Q_1 & O_{m\times n} \\ \hline O_{n\times m} & Q_2 \end{array}\right] \\
&= \left[\begin{array}{c|c} Q_1^T Q_1 + O_{m\times n} O_{n\times m} & Q_1^T O_{m\times n} + O_{m\times n} Q_2 \\ \hline O_{n\times m} Q_1 + Q_2^T O_{n\times m} & O_{n\times m} O_{m\times n} + Q_2^T Q_2 \end{array}\right] \\
&= \left[\begin{array}{c|c} Q_1^T Q_1 & O_{m\times n} \\ \hline O_{n\times m} & Q_2^T Q_2 \end{array}\right] \\
&= \left[\begin{array}{c|c} I_m & O_{m\times n} \\ \hline O_{n\times m} & I_n \end{array}\right] \\
&= I_{m+n}.
\end{aligned}$$

So, Q is indeed an orthogonal matrix. Q.E.D.

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

1. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
2. for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
3. for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof.

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof. (a) For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have that

$$(Q\mathbf{x}) \cdot (Q\mathbf{y}) \;=\; (Q\mathbf{x})^T (Q\mathbf{x}) \;=\; \mathbf{x}^T \underbrace{Q^T Q}_{=I_n} \mathbf{y} \;=\; \mathbf{x}^T \mathbf{y} \;=\; \mathbf{x} \cdot \mathbf{y}.$$

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof. (a) For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have that

$$(Q\mathbf{x}) \cdot (Q\mathbf{y}) \;=\; (Q\mathbf{x})^T(Q\mathbf{x}) \;=\; \mathbf{x}^T \underbrace{Q^T Q}_{=I_n} \mathbf{y} \;=\; \mathbf{x}^T \mathbf{y} \;=\; \mathbf{x} \cdot \mathbf{y}.$$

(b) For $\mathbf{x} \in \mathbb{R}^n$, we have that

$$||Q\mathbf{x}|| \;=\; \sqrt{(Q\mathbf{x}) \cdot (Q\mathbf{x})} \;\overset{(a)}{=}\; \sqrt{\mathbf{x} \cdot \mathbf{x}} \;=\; ||\mathbf{x}||.$$

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

a. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
b. for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
c. for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

Proof (continued).

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

- (a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
- (b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
- (c) for all $i, j \in \{1, \ldots, n\}$, $|q_{i,j}| \leq 1$.

Proof (continued). (c) By Theorem 2.1, the columns of Q form an orthonormal basis. In particular, all columns of Q are of length one, and it follows that all entries of Q have absolute value at most 1.

Theorem 2.4

Let $Q = [q_{i,j}]_{n\times n}$ be an orthogonal matrix in $\mathbb{R}^{n\times n}$. Then all the following hold:

a. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
b. for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
c. for all $i, j \in \{1, \ldots, n\}$, $|q_{i,j}| \leq 1$.

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

a. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
b. for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
c. for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Suppose that $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, and that $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ is given by $f_Q(\mathbf{x}) = Q\mathbf{x}$.
  - So, $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ is the linear transformation whose standard matrix is Q.

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

a. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;
b. for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;
c. for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Suppose that $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, and that $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ is given by $f_Q(\mathbf{x}) = Q\mathbf{x}$.
  - So, $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ is the linear transformation whose standard matrix is Q.
- By Theorem 2.4(b), f_Q preserves vector length.

Theorem 2.4

Let $Q = [q_{i,j}]_{n \times n}$ be an orthogonal matrix in $\mathbb{R}^{n \times n}$. Then all the following hold:

(a) for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$;

(b) for all $\mathbf{x} \in \mathbb{R}^n$, $||Q\mathbf{x}|| = ||\mathbf{x}||$;

(c) for all $i, j \in \{1, \dots, n\}$, $|q_{i,j}| \leq 1$.

- Suppose that $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, and that $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ is given by $f_Q(\mathbf{x}) = Q\mathbf{x}$.
    - So, $f_Q : \mathbb{R}^n \to \mathbb{R}^n$ is the linear transformation whose standard matrix is Q.
- By Theorem 2.4(b), f_Q preserves vector length.
- On the other hand, recall that for non-zero vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have that $\mathbf{x} \cdot \mathbf{y} = ||\mathbf{x}|| \, ||\mathbf{y}|| \cos\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. So, Theorem 2.4(a) effectively states that f_Q preserves angles between vectors.

- Scalar product, coordinate vectors, and matrices of linear transformations

- Scalar product, coordinate vectors, and matrices of linear transformations

Proposition 3.1

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle \quad = \quad [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}.$$

Proof.

- Scalar product, coordinate vectors, and matrices of linear transformations

Proposition 3.1

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle \;=\; [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}.$$

Proof. We prove the result for the case when V is a vector space over $\mathbb{C}$. The proof for $\mathbb{R}$ is similar but easier (because there are no complex conjugates).

- Scalar product, coordinate vectors, and matrices of linear transformations

Proposition 3.1

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}.$$

Proof. We prove the result for the case when V is a vector space over $\mathbb{C}$. The proof for $\mathbb{R}$ is similar but easier (because there are no complex conjugates). Fix $\mathbf{x}, \mathbf{y} \in V$. Since $\mathcal{B}$ is an orthonormal basis for V, we see that

$$\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle \, \mathbf{u}_i \quad \text{and} \quad \mathbf{y} = \sum_{i=1}^{n} \langle \mathbf{y}, \mathbf{u}_i \rangle \, \mathbf{u}_i,$$

and consequently (next slide):

Proposition 3.1

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle \quad = \quad [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}.$$

Proof (continued).

$$[\mathbf{x}]_{\mathcal{B}} \quad = \quad \begin{bmatrix} \langle \mathbf{x}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_n \rangle \end{bmatrix} \qquad \text{and} \qquad [\mathbf{y}]_{\mathcal{B}} \quad = \quad \begin{bmatrix} \langle \mathbf{y}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{y}, \mathbf{u}_n \rangle \end{bmatrix}.$$

Proposition 3.1

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}.$$

Proof (continued).

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{x}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_n \rangle \end{bmatrix} \quad \text{and} \quad [\mathbf{y}]_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{y}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{y}, \mathbf{u}_n \rangle \end{bmatrix}.$$

We now compute (next slide):

Proof (continued).

$$
\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \Big\langle \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle\ \mathbf{u}_i, \sum_{j=1}^{n} \langle \mathbf{y}, \mathbf{u}_j \rangle\ \mathbf{u}_j \Big\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \Big\langle \langle \mathbf{x}, \mathbf{u}_i \rangle\ \mathbf{u}_i, \langle \mathbf{y}, \mathbf{u}_j \rangle\ \mathbf{u}_j \Big\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle \overline{\langle \mathbf{y}, \mathbf{u}_j \rangle}\ \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\
&= \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle \overline{\langle \mathbf{y}, \mathbf{u}_i \rangle} && \text{by orthonormality} \\
&= \begin{bmatrix} \langle \mathbf{x}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_n \rangle \end{bmatrix} \cdot \begin{bmatrix} \langle \mathbf{y}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{y}, \mathbf{u}_n \rangle \end{bmatrix} \\
&= [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}} && \text{Q.E.D.}
\end{aligned}
$$

Proposition 3.1

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$, equipped with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $|| \cdot ||$, and let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis of V. Let $\cdot$ be the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}.$$

Theorem 3.2

Let U and V be non-trivial, finite-dimensional vector spaces over $\mathbb{R}$. Assume that U is equipped with a scalar product $\langle \cdot, \cdot \rangle_U$ and the induced norm $|| \cdot ||_U$, and that V is equipped with a scalar product $\langle \cdot, \cdot \rangle_V$ and the induced norm $|| \cdot ||_V$. Let $\mathcal{B}_U = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ and $\mathcal{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be orthonormal bases for U and V, respectively, and let $f : U \to V$ be a linear transformation. Then the following two statements are equivalent:

(i) the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;[a]

(ii) for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

---

[a] However, despite Theorem 2.1, this does not necessarily mean that the matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ is orthogonal. This is because ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ is an $n \times m$ matrix, and it is possible that $m \neq n$, i.e. that ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ is not a square matrix. Only square matrices can be orthogonal!

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof.

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof. Set ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = \left[\begin{array}{ccc} \mathbf{c}_1 & \dots & \mathbf{c}_m \end{array} \right]$.

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof. Set ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_m \end{bmatrix}$. We observe that

$$\begin{aligned} \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\right)^T {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} &= \begin{bmatrix} \mathbf{c}_1^T \\ \mathbf{c}_2^T \\ \vdots \\ \mathbf{c}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_m \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{c}_1 \cdot \mathbf{c}_1 & \mathbf{c}_1 \cdot \mathbf{c}_2 & \dots & \mathbf{c}_1 \cdot \mathbf{c}_m \\ \mathbf{c}_2 \cdot \mathbf{c}_1 & \mathbf{c}_2 \cdot \mathbf{c}_2 & \dots & \mathbf{c}_2 \cdot \mathbf{c}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{c}_m \cdot \mathbf{c}_1 & \mathbf{c}_m \cdot \mathbf{c}_2 & \dots & \mathbf{c}_m \cdot \mathbf{c}_m \end{bmatrix}. \end{aligned}$$

So, we see that (i) holds if and only if $\left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\right)^T {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = I_m$.

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). By Proposition 3.1, for all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\mathbf{x}]_{\mathcal{B}_U} \cdot [\mathbf{y}]_{\mathcal{B}_U}$;

(2) $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [f(\mathbf{x})]_{\mathcal{B}_V} \cdot [f(\mathbf{y})]_{\mathcal{B}_V}$.

(i) the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;

(ii) for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). By Proposition 3.1, for all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\mathbf{x}]_{\mathcal{B}_U} \cdot [\mathbf{y}]_{\mathcal{B}_U}$;

(2) $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [f(\mathbf{x})]_{\mathcal{B}_V} \cdot [f(\mathbf{y})]_{\mathcal{B}_V}$.

Now, for all $\mathbf{x}, \mathbf{y} \in U$, we have that

$$\begin{aligned}
\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V &\overset{(2)}{=} [f(\mathbf{x})]_{\mathcal{B}_V} \cdot [f(\mathbf{y})]_{\mathcal{B}_V} \\
&= \left([f(\mathbf{x})]_{\mathcal{B}_V}\right)^T \, [f(\mathbf{y})]_{\mathcal{B}_V} \\
&= \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \, [\mathbf{x}]_{\mathcal{B}_U}\right)^T \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \, [\mathbf{y}]_{\mathcal{B}_U}\right) \\
&= [\mathbf{x}]_{\mathcal{B}_U}{}^T \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\right)^T \, {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \, [\mathbf{y}]_{\mathcal{B}_U}.
\end{aligned}$$

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). Reminder: For all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\mathbf{x}]_{\mathcal{B}_U} \cdot [\mathbf{y}]_{\mathcal{B}_U}$;

(2) $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [f(\mathbf{x})]_{\mathcal{B}_V} \cdot [f(\mathbf{y})]_{\mathcal{B}_V}$.

$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = {[\mathbf{x}]_{\mathcal{B}_U}}^T \ ({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U})^T \ {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \ [\mathbf{y}]_{\mathcal{B}_U}$.

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). Reminder: For all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\mathbf{x}]_{\mathcal{B}_U} \cdot [\mathbf{y}]_{\mathcal{B}_U}$;

(2) $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [f(\mathbf{x})]_{\mathcal{B}_V} \cdot [f(\mathbf{y})]_{\mathcal{B}_V}$.

$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [\mathbf{x}]_{\mathcal{B}_U}{}^T \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \right)^T {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \, [\mathbf{y}]_{\mathcal{B}_U}$.

Suppose first that (i) holds. Then $\left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \right)^T {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = I_m$,

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). Reminder: For all $\mathbf{x}, \mathbf{y} \in U$:

(1) $\langle \mathbf{x}, \mathbf{y} \rangle_U = [\mathbf{x}]_{\mathcal{B}_U} \cdot [\mathbf{y}]_{\mathcal{B}_U}$;

(2) $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [f(\mathbf{x})]_{\mathcal{B}_V} \cdot [f(\mathbf{y})]_{\mathcal{B}_V}$.

$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = [\mathbf{x}]_{\mathcal{B}_U}{}^T \; ({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U})^T \; {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} \; [\mathbf{y}]_{\mathcal{B}_U}$.

Suppose first that (i) holds. Then $({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U})^T \; {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = I_m$, and consequently, for all $\mathbf{x}, \mathbf{y} \in U$, we have that

$$\begin{aligned} \langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V &= [\mathbf{x}]_{\mathcal{B}_U}{}^T \; \underbrace{({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U})^T \; {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}}_{=I_m} \; [\mathbf{y}]_{\mathcal{B}_U} \\ &= [\mathbf{x}]_{\mathcal{B}_U}{}^T \; [\mathbf{y}]_{\mathcal{B}_U} \\ &= [\mathbf{x}]_{\mathcal{B}_U} \cdot [\mathbf{y}]_{\mathcal{B}_U} \\ &\overset{(1)}{=} \langle \mathbf{x}, \mathbf{y} \rangle_U. \end{aligned}$$

Thus, (ii) holds.

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). Reminder: ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_m \end{bmatrix}$.

Suppose now that (ii) holds. Then for all $i, j \in \{1, \dots, m\}$, we have that

$$\begin{aligned}
\mathbf{e}_i^m \cdot \mathbf{e}_j^m &= [\mathbf{u}_i]_{\mathcal{B}_U} \cdot [\mathbf{u}_j]_{\mathcal{B}_U} \\
&\overset{(1)}{=} \langle \mathbf{u}_i, \mathbf{u}_j \rangle_U \\
&\overset{(ii)}{=} \langle f(\mathbf{u}_i), f(\mathbf{u}_j) \rangle_V \\
&\overset{(2)}{=} [f(\mathbf{u}_i)]_{\mathcal{B}_V} \cdot [f(\mathbf{u}_j)]_{\mathcal{B}_V} \\
&= \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\ [\mathbf{u}_i]_{\mathcal{B}_U}\right) \cdot \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\ [\mathbf{u}_j]_{\mathcal{B}_U}\right) \\
&= \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\ \mathbf{e}_i^m\right) \cdot \left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\ \mathbf{e}_j^m\right) \\
&= \mathbf{c}_i \cdot \mathbf{c}_j.
\end{aligned}$$

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). Reminder: ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_m \end{bmatrix}$; for all $i, j \in \{1, \dots, m\}$, $\mathbf{e}_i^m \cdot \mathbf{e}_j^m = \mathbf{c}_i \cdot \mathbf{c}_j$.

1. the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;
2. for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

Proof (continued). Reminder: ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = \begin{bmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_m \end{bmatrix}$; for all $i, j \in \{1, \dots, m\}$, $\mathbf{e}_i^m \cdot \mathbf{e}_j^m = \mathbf{c}_i \cdot \mathbf{c}_j$.

This implies that $\{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ form an orthonormal set of vectors in $\mathbb{R}^n$, i.e. (i) holds.

Theorem 3.2

Let U and V be non-trivial, finite-dimensional vector spaces over $\mathbb{R}$. Assume that U is equipped with a scalar product $\langle \cdot, \cdot \rangle_U$ and the induced norm $|| \cdot ||_U$, and that V is equipped with a scalar product $\langle \cdot, \cdot \rangle_V$ and the induced norm $|| \cdot ||_V$. Let $\mathcal{B}_U = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ and $\mathcal{B}_V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be orthonormal bases for U and V, respectively, and let $f : U \to V$ be a linear transformation. Then the following two statements are equivalent:

(i) the columns of the $n \times m$ matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ form an orthonormal set of vectors in $\mathbb{R}^n$;[a]

(ii) for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.

---

[a]However, despite Theorem 2.1, this does not necessarily mean that the matrix ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ is orthogonal. This is because ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ is an $n \times m$ matrix, and it is possible that $m \neq n$, i.e. that ${}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}$ is not a square matrix. Only square matrices can be orthogonal!