Linear Algebra 2: Lecture 14

Irena Penev

Summer 2023

In this section, we assume that $\mathbb{R}^n$ is equipped with the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

1 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). $\square$

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. If $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis of V, then obviously, it is an orthonormal set. Suppose now that $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal set. Then $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is linearly independent (by Proposition 2.3 from Lecture Notes 12). But now by Proposition 1.11(a) from Lecture Notes 7,[1] we see that $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a basis of V, and by supposition, this basis is orthonormal. $\square$

2 Orthogonal matrices

A matrix $Q \in \mathbb{R}^{n\times n}$ is *orthogonal* if it satisfies $Q^TQ = I_n$. The following theorem gives several equivalent characterizations of orthogonal matrices.

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) $QQ^T = I_n$;

---

[1] Proposition 1.11(a) from Lecture Notes 7 states that any linearly independent set of n vectors in an n-dimensional vector space is in fact a basis of that vector space.

(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.

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.

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

$$\begin{aligned} Q^TQ &= \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{q}_1 \cdot \mathbf{q}_1 & \mathbf{q}_1 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_1 \cdot \mathbf{q}_n \\ \mathbf{q}_2 \cdot \mathbf{q}_1 & \mathbf{q}_2 \cdot \mathbf{q}_2 & \dots & \mathbf{q}_2 \cdot \mathbf{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n \cdot \mathbf{q}_1 & \mathbf{q}_n \cdot \mathbf{q}_2 & \dots & \mathbf{q}_n \cdot \mathbf{q}_n \end{bmatrix}. \end{aligned}$$

So, $Q^TQ = I_n$ 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. This completes the argument. □

Propositions 2.2 and 2.3 (below) give two different ways of forming new orthogonal matrices from old ones.

Proposition 2.2. *If* $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ *are orthogonal, then so is their product* Q_1Q_2*.*

Proof. Assume $Q_1, Q_2 \in \mathbb{R}^{n \times n}$ are orthogonal. Then $Q_1^TQ_1 = I_n$ and $Q_2^TQ_2 = I_n$, and consequently,

$$(Q_1Q_2)^T(Q_1Q_2) = Q_2^T \underbrace{Q_1^TQ_1}_{=I_n} Q_2 = Q_2^TQ_2 = I_n.$$

So, Q_1Q_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 the* $(m+n)\times(m+n)$ *matrix*

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

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

Proof. By hypothesis, we have that $Q_1^TQ_1 = I_m$ and $Q_2^TQ_2 = I_n$. We now compute:

$$\begin{aligned} Q^TQ &= \left[\begin{array}{c:c} Q_1^T & O_{m\times n} \\ \hdashline O_{n\times m} & Q_2^T \end{array}\right] \left[\begin{array}{c:c} Q_1 & O_{m\times n} \\ \hdashline O_{n\times m} & Q_2 \end{array}\right] \\ &= \left[\begin{array}{c:c} Q_1^TQ_1 + O_{m\times n}O_{n\times m} & Q_1^TO_{m\times n} + O_{m\times n}Q_2 \\ \hdashline O_{n\times m}Q_1 + Q_2^TO_{n\times m} & O_{n\times m}O_{m\times n} + Q_2^TQ_2 \end{array}\right] \\ &= \left[\begin{array}{c:c} Q_1^TQ_1 & O_{m\times n} \\ \hdashline O_{n\times m} & Q_2^TQ_2 \end{array}\right] \\ &= \left[\begin{array}{c:c} I_m & O_{m\times n} \\ \hdashline O_{n\times m} & I_n \end{array}\right] \\ &= I_{m+n}. \end{aligned}$$

So, Q is indeed an orthogonal matrix. ☐

We now consider some examples of orthogonal matrices.

- Obviously, 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

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

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

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

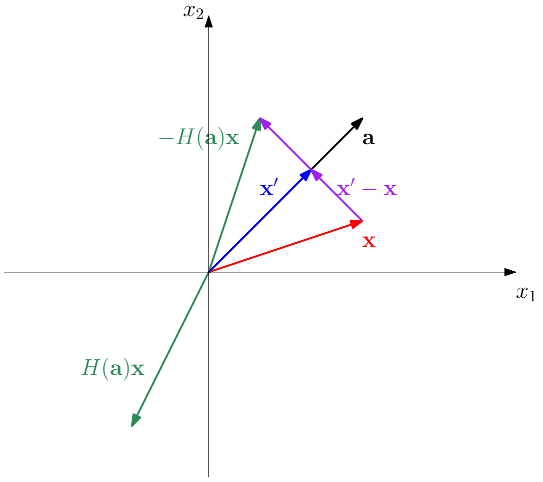
Let now us discuss the geometric meaning of this matrix. Using Theorem 2.3 from Lecture Notes 13, we see that the standard matrix of orthogonal projection onto the line $\mathrm{Span}(\mathbf{a}) = \mathrm{Col}(\begin{bmatrix} \mathbf{a} \end{bmatrix})$ is

$$\mathbf{a}(\mathbf{a}^T\mathbf{a})^{-1}\mathbf{a}^T \;=\; \mathbf{a}(\mathbf{a}\cdot\mathbf{a})^{-1}\mathbf{a}^T \;=\; \tfrac{1}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\mathbf{a}^T.$$

Now, if $\mathbf{x}$ is any vector in $\mathbb{R}^n$, and $\mathbf{x}'$ represents the orthogonal projection of $\mathbf{x}$ onto $\mathrm{Span}(\mathbf{a}) = \mathrm{Col}(\begin{bmatrix} \mathbf{a} \end{bmatrix})$, then the reflection of $\mathbf{x}$ about the line $\mathrm{Span}(\mathbf{a}) = \mathrm{Col}(\begin{bmatrix} \mathbf{a} \end{bmatrix})$ is given by

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

Thus, $-H(\mathbf{a})$ is the standard matrix of reflection about the $\mathrm{Span}(\mathbf{a})$ line. The Householder matrix $H(\mathbf{a})$ itself is the standard matrix of the linear operation that first reflects about the $\mathrm{Span}(\mathbf{a})$ line and then reflects about the origin. In the case of $\mathbb{R}^2$, this is illustrated in the picture below.



- 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:

$$G_{i,j}(c,s) \quad = \quad \begin{array}{c|cccccccccccc|}
 & & & & & i & & & & j & & & \\
\hline
 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
 & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
 & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\
 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
i & 0 & 0 & \dots & 0 & c & 0 & \dots & 0 & -s & 0 & \dots & 0 \\
 & 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\
 & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & 0 \\
 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\
j & 0 & 0 & \dots & 0 & s & 0 & \dots & 0 & c & 0 & \dots & 0 \\
 & 0 & 0 & \dots & 0 & 0 & 0 & & 0 & 0 & 1 & & 0 \\
 & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\
 & 0 & 0 & \dots & 0 & 0 & 0 & & 0 & 0 & 0 & \dots & 1 \\
\hline
\end{array}$$

More precisely, $G_{i,j}(c,s) = [g_{i,j}]_{n\times n}$ is given by:

- $g_{i,i} = g_{j,j} = c$;
- $g_{k,k} = 1$ for all $k \in \{1,\dots,n\} \setminus \{i,j\}$;
- $g_{i,j} = -s$;
- $g_{j,i} = s$;
- all other entries are zero.

It is not hard to check that the columns of $G_{i,j}(c,s)$ form an orthonormal set in $\mathbb{R}^n$, and therefore (by Proposition 1.2) an orthonormal basis of $\mathbb{R}^n$. So, by Theorem 2.1, the Givens matrix $G_{i,j}(c,s)$ really is orthogonal. Let us now give a geometric interpretation of this matrix. 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_ix_j-plane. This is particularly easy to see in the case when $n = 2$. In that case, we have that

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

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

Theorem 2.4. *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

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

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

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

(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. ☐

Remark: 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.

3 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. 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

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

and consequently,

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

We now compute:

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \Big\langle \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle\, \mathbf{u}_i, \sum_{i=1}^{n} \langle \mathbf{y}, \mathbf{u}_i \rangle\, \mathbf{u}_i \Big\rangle \\
&= \Big\langle \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle\, \mathbf{u}_i, \sum_{j=1}^{n} \langle \mathbf{y}, \mathbf{u}_j \rangle\, \mathbf{u}_j \Big\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \Big\langle \langle \mathbf{x}, \mathbf{u}_i \rangle\, \mathbf{u}_i, \langle \mathbf{y}, \mathbf{u}_j \rangle\, \mathbf{u}_j \Big\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle \overline{\langle \mathbf{y}, \mathbf{u}_j \rangle}\, \langle \mathbf{u}_i, \mathbf{u}_j \rangle \\
&\overset{(*)}{=} \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i \rangle \overline{\langle \mathbf{y}, \mathbf{u}_i \rangle} \\
&= \begin{bmatrix} \langle \mathbf{x}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_n \rangle \end{bmatrix} \cdot \begin{bmatrix} \langle \mathbf{y}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{y}, \mathbf{u}_n \rangle \end{bmatrix} \\
&= [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}},
\end{aligned}$$

where (*) follows from the fact that $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal set. □

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 $\mathbb{R}^n$;*[2]
- *(ii) for all $\mathbf{x}, \mathbf{y} \in U$, we have that $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle_V = \langle \mathbf{x}, \mathbf{y} \rangle_U$.*

---

[2] 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!

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 $\left({}_{\mathcal{B}_V}[f]_{\mathcal{B}_U}\right)^T \ {}_{\mathcal{B}_V}[f]_{\mathcal{B}_U} = I_m$.

Next, by Proposition 3.1, the following hold for all $\mathbf{x}, \mathbf{y} \in U$:

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

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

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

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

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

Thus, (ii) holds.

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

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

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