Worksheet 17: Orthogonal Transformations (§5.3)

Definition: An $n \times n$ matrix A is said to be *symmetric* if $A^\top = A$, and *orthogonal* if $A^\top = A^{-1}$, i.e., if A is invertible and its inverse is the same as its transpose.

Problem 1.

(a) Show that if A and B are orthogonal $n \times n$ matrices, then the matrices $A^\top$, A^{-1}, and AB are also orthogonal.

(b) Show that for any matrix A, both $A^\top A$ and $AA^\top$ are symmetric. Expressed as a dot product, what is the (i, j)-entry of $A^\top A$? How about the (i, j)-entry of $AA^\top$?

Solution:

(a) These facts following immediately from the definition using the identities $(A^\top)^\top = A$, $(AB)^\top = B^\top A^\top$, and $(A^{-1})^\top = (A^\top)^{-1}$.

(b) $(A^\top A)^\top = A^\top (A^\top)^\top = A^\top A$ and $(AA^\top)^\top = (A^\top)^\top A^\top = AA^\top$, so both $A^\top A$ and $AA^\top$ are symmetric. The (i, j)-entry of $A^\top A$ is the dot product of the ith and jth columns of A, and the (i, j)-entry of $AA^\top$ is the dot product of the ith and jth rows of A.

Definition: A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be an *orthogonal* transformation if it preserves the dot product, i.e., if for all vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$T(\vec{v}) \cdot T(\vec{w}) \;=\; \vec{v} \cdot \vec{w}.$$

Problem 2. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, and let A be the standard matrix of T. Prove that the following are equivalent:*

(a) T *preserves length*, i.e., $\|T(v)\| = \|v\|$ for all $v \in \mathbb{R}^n$.†

(b) T *preserves distance*, i.e., $\|T(v) - T(w)\| = \|v - w\|$ for all $v, w \in \mathbb{R}^n$.

(c) T is an orthogonal transformation, i.e., T preserves the dot product.

(d) T maps any orthonormal basis of $\mathbb{R}^n$ to an orthonormal basis of $\mathbb{R}^n$.

(e) T maps the standard basis of $\mathbb{R}^n$ to an orthonormal basis of $\mathbb{R}^n$.

(f) The columns of A form an orthonormal basis of $\mathbb{R}^n$.

(g) $A^\top A = I_n$.

---

*Hint: prove $(a) \Leftrightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (f) \Leftrightarrow (g) \Leftrightarrow (h) \Leftrightarrow (i)$ and $(g) \Rightarrow (c) \Rightarrow (a)$.

†This is the textbook's definition of *orthogonal transformation*. We prefer the one that is given in terms of the dot product, but part of what you are proving here is that the two definitions are equivalent.

(h) $AA^\top = I_n$. $\Big\}$ (i.e., A is an orthogonal matrix)

(i) The rows of A form an orthonormal basis of $\mathbb{R}^n$.

Solution:
$(a) \Leftrightarrow (b)$: If T is linear and preserves lengths, then for all $v, w \in \mathbb{R}^n$, $\|T(v) - T(w)\| = \|T(v-w)\| = \|v - w\|$. Conversely, if T is linear and preserves distances, then for all $v \in \mathbb{R}^n$, $\|T(v)\| = \|T(v) - 0\| = \|T(v) - T(0)\| = \|v - 0\| = \|v\|$.
$(a \wedge b) \Rightarrow (c)$: Let $v, w \in \mathbb{R}^n$. Expanding each side of $\|T(v-w)\| = \|v-w\|$ in terms of the dot product and using the facts that $\|T(v)\| = \|v\|$ and $\|T(w)\| = \|w\|$ to simplify, we find that $T(v) \cdot T(w) = v \cdot w$.
$(c) \Rightarrow (d)$: Assuming (c), if $u_i \cdot u_j = \delta_{ij}$ for each i, j, then also $T(u_i) \cdot T(u_j) = \delta_{ij}$ for each i, j.
$(d) \Rightarrow (e)$: Immediate, since the standard basis of $\mathbb{R}^n$ is orthonormal.
$(e) \Rightarrow (f)$: Follows from the fact that $A = [\, T(\vec{e}_1) \ \cdots \ T(\vec{e}_n) \,]$.
$(f) \Leftrightarrow (g)$: If we write $A = [\, \vec{u}_1 \ \cdots \ \vec{u}_n \,]$, the (i,j)-entry of $A^\top A$ is $\vec{u}_i \cdot \vec{u}_j$.
$(g) \Leftrightarrow (h)$: Follows from Theorem 2.4.8 in the text.
$(h) \Leftrightarrow (i)$: If we write $A = \begin{bmatrix} \text{—} & \vec{u}_1^\top & \text{—} \\ & \vdots & \\ \text{—} & \vec{u}_n^\top & \text{—} \end{bmatrix}$, the (i,j)-entry of $AA^\top$ is $\vec{u}_i \cdot \vec{u}_j$.
$(g) \Rightarrow (c)$: If $A^\top A = I_n$, then $T(v) \cdot T(w) = Av \cdot Aw = v^\top A^\top A w = v^\top w = v \cdot w$.
$(c) \Rightarrow (a)$: If T preserves the dot product, then for all v, $\|T(v)\|^2 = T(v) \cdot T(v) = v \cdot v = \|v\|^2$.

Problem 3.

(a) Give some examples of orthogonal transformations. reflections, rotations

(b) For $n > 1$, give an example of a function $f : \mathbb{R}^n \to \mathbb{R}^n$ that preserves length but is *not* an orthogonal transformation. (Why does this not contradict Problem 2?) $f(\vec{x}) = \|\vec{x}\|\vec{e}_1$. This doesn't contradict Problem 2 because f is not linear.

Problem 4. Show that if $T : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, then the matrix of T with respect to any orthonormal basis of $\mathbb{R}^n$ is an orthogonal matrix. Is this true for *any* basis of $\mathbb{R}^n$ (i.e., not necessarily an orthonormal one)?

Solution: Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, so the standard matrix $[T]_\mathcal{E}$ of T is orthogonal by Problem 2. Let $\mathcal{U} = (\vec{u}_1, \ldots, \vec{u}_n)$ be an orthonormal basis of $\mathbb{R}^n$, and write $Q = [\vec{u}_1 \ \cdots \ \vec{u}_n]$, so that Q is an orthogonal matrix and thus so is the matrix $Q^{-1} = Q^\top$ by Problem 1. Then

$$[T]_\mathcal{U} \ = \ S_{\mathcal{E}\to\mathcal{U}}[T]_\mathcal{E} S_{\mathcal{U}\to\mathcal{E}} \ = \ Q^\top [T]_\mathcal{E} Q$$

is also orthogonal, since products of orthogonal matrices are orthogonal (again by Problem 1).

This statement could fail, however, if the basis is not orthonormal. For instance, if $\mathcal{B} = (\vec{e}_1, \vec{e}_1 + \vec{e}_2)$ then the $\mathcal{B}$-matrix of the reflection of $\mathbb{R}^2$ over the x-axis is $\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ which is not an orthogonal matrix even though the reflection is an orthogonal transformation.

Problem 5. Determine all orthogonal transformations $T : \mathbb{R}^2 \to \mathbb{R}^2$.

Solution: Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is an orthogonal transformation. Then T is linear, so it is determined by its effect on the standard basis $(\vec{e}_1, \vec{e}_2)$. By Problem (2), $\vec{u}_1 = T(\vec{e}_1)$ and $\vec{u}_2 = T(\vec{e}_2)$ are unit vectors that are orthogonal to one another. If $\vec{u}_2$ is obtained from $\vec{u}_1$ by a 90° counter-clockwise rotation, then T is a rotation. If $\vec{u}_2$ is obtained from $\vec{u}_1$ by a 90° *clockwise* rotation, then T is a rotation followed by a reflection, which is itself a reflection. So T is either a rotation or a reflection. In terms of matrices, the matrix of T relative to the standard coordinates must be

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$