Real Inner Product Spaces and
Orthogonal Transformations

Math 422

Definition 1 *A vector space V equipped with a real valued function $< , >: V \times V \to \mathbb{R}$ is a* real inner product space *if and only if the following conditions hold for all $a, b, c \in V$ and $t \in \mathbb{R}$:*

1. $\langle a, b \rangle = \langle b, a \rangle$
2. $\langle a+b, c \rangle = \langle a+c, b+c \rangle$
3. $\langle ta, b \rangle = \langle a, tb \rangle = t \langle a, b \rangle$
4. $\langle a, a \rangle \geq 0$ *and* $\langle a, a \rangle = 0$ *if and only if* $a = 0$.

The function $< , >$ on an inner product space V, is called an inner product *on V.*

Real valued functions on a vector space are called *forms*. Axiom (1) says that the form $\langle \, , \, \rangle$ is symmetric. Taken together, axioms (1), (2) and (3) tell us that $\langle \, , \, \rangle$ is *bilinear*, i.e., for each fixed element $x \in V$, the forms $\langle x, \, \rangle : \mathbb{R} \to \mathbb{R}$ and $\langle \, ,x \rangle : \mathbb{R} \to \mathbb{R}$ are linear. Axiom (4) says that $\langle \, , \, \rangle$ is *positive definite* and *non-degenerate*. In these terms, we say that *an inner product is a symmetric, bilinear, positive definite and non-degenerate form on $V \times V$.*

Definition 2 *Let V be an inner product space. The* norm *on V induced by the inner product is a form $\| \, \| : V \to \mathbb{R}$ given by $\|a\| = \sqrt{\langle a, a \rangle}$. The* distance *between two vectors $a, b \in V$ relative to the given inner product is defined by $d(a, b) = \|a - b\|$.*

Definition 3 *The* Euclidean inner product *(often called the "dot product") of two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ is defined by $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n$. In this case we use the notation $\mathbf{x} \bullet \mathbf{y}$ to denote the inner product.*

Theorem 4 ***(The polarization identity)*** *Let $\mathbf{x}, \mathbf{y}$ be elements of an inner product sapce V. Then*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} \|\mathbf{x} + \mathbf{y}\|^2 - \frac{1}{4} \|\mathbf{x} - \mathbf{y}\|^2. \tag{1}$$

Proof.

$$\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \mathbf{2} \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \|x\|^2 + \mathbf{2} \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$$

and

$$\|\mathbf{x} - \mathbf{y}\|^2 = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \mathbf{2} \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \|x\|^2 - \mathbf{2} \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2.$$

Therefore

$$\begin{aligned} \frac{1}{4} \|\mathbf{x} + \mathbf{y}\|^2 - \frac{1}{4} \|\mathbf{x} - \mathbf{y}\|^2 &= \frac{1}{4} \left(\|x\|^2 + \mathbf{2} \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \right) - \frac{1}{4} \left(\|x\|^2 - \mathbf{2} \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \right) \\ &= \langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

■

Let $\mathbf{x} = [x_1 \cdots x_n]^T$, $\mathbf{y} = [y_1 \cdots y_n]^T$ be coordinate matrices in the standard basis for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Observe that the dot product can be expressed as a matrix product as follows:

$$\mathbf{x} \bullet \mathbf{y} = x_1 y_1 + \cdots + x_n y_n = [x_1 \cdots x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \mathbf{x}^T \mathbf{y}.$$

So if A is an $n \times n$ square matrix,

$$A\mathbf{x} \bullet \mathbf{y} = (A\mathbf{x})^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \mathbf{x} \bullet A^T \mathbf{y}. \tag{2}$$

Definition 5 *A square matrix A is* <u>*orthogonal*</u> *if $A^T = A^{-1}$.*

Theorem 6 *Let A be an $n \times n$ matrix. The following statements are all equivalent:*

1. *A is an orthogonal matrix.*
2. *$\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$, i.e., multiplication by A preserves Euclidean norm.*
3. *$A\mathbf{x} \bullet A\mathbf{y} = \mathbf{x} \bullet \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, i.e., multiplication by A preserves Euclidean inner product.*

Proof. (1) $\Rightarrow$ (2). Let $\mathbf{x} \in \mathbb{R}^n$. Using the fact that $A^T A = I$ and the identity in equation (2),

$$\|A\mathbf{x}\|^2 = A\mathbf{x} \bullet A\mathbf{x} = \mathbf{x} \bullet A^T A\mathbf{x} = \mathbf{x} \bullet \mathbf{x} = \|\mathbf{x}\|^2 .$$

(2) $\Rightarrow$ (3). Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. By the polarization identity (1) and assumption (2),

$$A\mathbf{x} \bullet A\mathbf{y} = \frac{1}{4}\|A\mathbf{x}+A\mathbf{y}\|^2 - \frac{1}{4}\|A\mathbf{x}-A\mathbf{y}\|^2 = \frac{1}{4}\|A(\mathbf{x}+\mathbf{y})\|^2 - \frac{1}{4}\|A(\mathbf{x}-\mathbf{y})\|^2$$

$$= \frac{1}{4}\|\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}\|\mathbf{x}-\mathbf{y}\|^2 = \mathbf{x} \bullet \mathbf{y}.$$

(3) $\Rightarrow$ (1). Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. By assumption (3) and equation (2),

$$0 = A\mathbf{x} \bullet A\mathbf{y} - \mathbf{x} \bullet \mathbf{y} = \mathbf{x} \bullet A^T A\mathbf{y} - \mathbf{x} \bullet I\mathbf{y} = \mathbf{x} \bullet \left(A^T A\mathbf{y} - I\mathbf{y}\right) = \mathbf{x} \bullet \left(A^T A - I\right)\mathbf{y}.$$

Since this holds for all $\mathbf{x} \in \mathbb{R}^n$, it holds in particular when $\mathbf{x} = \left(A^T A - I\right)\mathbf{y}$.Thus

$$0 = \left(A^T A - I\right)\mathbf{y} \bullet \left(A^T A - I\right)\mathbf{y} = \left\|\left(A^T A - I\right)\mathbf{y}\right\|^2$$

in which case

$$\left(A^T A - I\right)\mathbf{y} = \mathbf{0}.$$

This is a homogeneous system of linear equations satisfied by all $\mathbf{y} \in \mathbb{R}^n$. In particular, this holds for $\mathbf{y} = \mathbf{e}_i$, in which case the i^{th} column of $\left(A^T A - I\right)$ is $\mathbf{0}$. Since this is true for all $i = 1, \ldots, n$, the matrix $A^T A - I = \mathbf{0}$ and A is orthogonal. ■

Definition 7 *A linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ is an* <u>*orthogonal transformation*</u> *if and only if its matrix in the standard basis is an orthogonal matrix.*

Example 8 *Rotations about the origin and reflections in lines through the origin are orthogonal transformations of the plane.*

Unlike the determinant, orthogonality is to some extent basis dependent. Since the purpose of the Euclidean inner product is to measure Euclidean length and angle, it makes sense to study orthogonality relative to coordinate matrices of vectors in an *orthonormal basis,* i.e., a basis of pairwise orthogonal unit vectors. Consider $\mathbb{R}^n$ with its standard basis and the Euclidean inner product. Consider an orthonormal basis

$$\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$$

and construct the associated transition matrix P from $\mathcal{B}$-coordinates to standard coordinates:

$$P = [\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_n].$$

Then by orthonormality of the $\mathbf{u}_i$'s we have

$$P^T P = \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix} [\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_n] = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \cdots & \mathbf{u}_1^T\mathbf{u}_n \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & & \mathbf{u}_2^T\mathbf{u}_n \\ \vdots & & \ddots & \vdots \\ \mathbf{u}_n^T\mathbf{u}_1 & \mathbf{u}_n^T\mathbf{u}_2 & \cdots & \mathbf{u}_n^T\mathbf{u}_n \end{bmatrix} = I$$

so that $P^{-1} = P^T$; thus P and P^{-1} are orthogonal matrices. By Theorem 6, multiplication by P and P^{-1} preserves the Euclidean inner product. Therefore, if L is an orthogonal transformation and P^{-1} is an orthogonal transition matrix from standard basis to orthonormal basis $\mathcal{B}$, then

$$[L]_{\mathcal{B}} [L]_{\mathcal{B}}^T = (P^T [L] P)(P^T [L] P)^T = P^T [L] P P^T [L]^T P = I,$$

which proves that $[L]_{\mathcal{B}}$ is orthogonal. In summary:

Theorem 9 *The matrix of an orthogonal transformation in any orthonormal basis is an orthogonal matrix.*

Corollary 10 *If $L : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation, then*

$$\det(L) = \pm 1.$$

Proof. Since the standard basis is an orthonormal basis, $[L]$ is an orthogonal matrix. Furthermore, $\det(A) = \det\left(A^T\right)$ and $\det(AB) = \det A \det B$ for all square matrices A and B. Therefore

$$[\det(L)]^2 = \det([L]) \det\left([L]^T\right) = \det\left([L][L]^T\right) = \det(I) = 1.$$

■