17. Lecture 17: 5.3 Orthogonal Transformations

A linear transformation $T:\mathbf{R}^n \to \mathbf{R}^n$ is called an **orthogonal transformation** if for all $\mathbf{u}$,$\mathbf{v}$

$$T(\mathbf{u})\cdot T(\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}. \tag{17.1}$$

Note that in particular that by taking $\mathbf{v}=\mathbf{u}$ and recalling that $\mathbf{u}\cdot\mathbf{u} = \|\mathbf{u}\|^2$ it follows that

$$\|T(\mathbf{u})\| = \|\mathbf{u}\|. \tag{17.2}$$

The book takes (17.2) as definition of orthogonal but (17.1) also follows from (17.2). In fact

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u} = \|\mathbf{u}\|^2+\|\mathbf{v}\|^2+2\,\mathbf{u}\cdot\mathbf{v}. \tag{17.3}$$

Similarly

$$\|\mathbf{u}-\mathbf{v}\|^2 = (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = \mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}-\mathbf{u}\cdot\mathbf{v}-\mathbf{v}\cdot\mathbf{u} = \|\mathbf{u}\|^2+\|\mathbf{v}\|^2-2\,\mathbf{u}\cdot\mathbf{v}. \tag{17.4}$$

Subtracting (17.4) from (17.3) gives

$$\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 = 4\,\mathbf{u}\cdot\mathbf{v}. \tag{17.5}$$

Applying this to $T(\mathbf{u})$ and $T(\mathbf{v})$ in place of $\mathbf{u}$ and $\mathbf{v}$ also gives

$$\|T(\mathbf{u})+T(\mathbf{v})\|^2 - \|T(\mathbf{u})-T(\mathbf{v})\|^2 = 4\,T(\mathbf{u})\cdot T(\mathbf{v}). \tag{17.6}$$

Using (17.2) applied $\mathbf{u}+\mathbf{v}$ respectively $\mathbf{u}-\mathbf{v}$ in place of $\mathbf{u}$ we get

$$\|T(\mathbf{u}+\mathbf{v})\|^2 = \|\mathbf{u}+\mathbf{v}\|^2, \quad \text{and} \quad \|T(\mathbf{u}-\mathbf{v})\|^2 = \|\mathbf{u}-\mathbf{v}\|^2.$$

It follows that the left of (17.5) must be equal to the left of (17.6) which implies that the right hand sides are also equal so (17.1) follows.

Recall that the dot product satisfy

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\phi, \tag{17.7}$$

where ϕ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Since the orthogonal transformations preserve both the length of vectors (17.2) and the dot product of vectors (17.1) it follows (17.7) and (17.7) applied to $T(\mathbf{u})$ and $T(\mathbf{v})$ in place of $\mathbf{u}$ and $\mathbf{v}$ respectively, that the cosine of the angle between $T(\mathbf{u})$ and $T(\mathbf{v})$ is the same as the cosine of the angle between $\mathbf{u}$ and $\mathbf{v}$, i.e.

$$\frac{T(\mathbf{u})\cdot T(\mathbf{v})}{\|T(\mathbf{u})\|\|T(\mathbf{v})\|} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos\phi.$$

Rotations and Reflections are both orthogonal transformations since they both preserve the length of vectors and hence the angle between vectors.

Orthogonal Matrix

If $T(\mathbf{x}) = Q\mathbf{x}$ is an orthogonal transformation with matrix

$$Q = \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & | \end{bmatrix}$$

then it follows from the definition (17.1) that the columns of Q are orthonormal

$$\mathbf{u}_i \cdot \mathbf{u}_j = (Q\mathbf{e}_i) \cdot (Q\mathbf{e}_j) = T(\mathbf{e}_i) \cdot T(\mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

We have

$$Q^T Q = \begin{bmatrix} - & \mathbf{u}_1^T & - \\ - & \mathbf{u}_2^T & - \\ & \vdots & \\ - & \mathbf{u}_n^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{u}_n \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \mathbf{u}_2 \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{u}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_n \cdot \mathbf{u}_1 & \mathbf{u}_n \cdot \mathbf{u}_2 & \cdots & \mathbf{u}_n \cdot \mathbf{u}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

i.e.

$$Q^T Q = I. \tag{17.8}$$

A matrix satisfying this, or equivalently that the columns are orthonormal, is called an **orthogonal matrix**. (It should perhaps have been called orthonormal but it is not.)

Recall that the **transpose** A^T of A is the matrix whose ijth entry a_{ij}^T is the jith entry a_{ji} of A: i.e the rows and columns are interchanged.

We claim that the transpose satisfy

$$\mathbf{u} \cdot (A^T \mathbf{v}) = (A\mathbf{u}) \cdot \mathbf{v}. \tag{17.9}$$

If fact if we let $\mathbf{u} = \mathbf{e}_i$ and $\mathbf{v} = \mathbf{e}_j$ run over the standard basis then the left is a_{ij}^T and the right is a_{ji} which are equal by the definition of A^T.

It follows from repeated use of (17.9) that

$$(AB\mathbf{u}) \cdot \mathbf{v} = (B\mathbf{u}) \cdot (A^T \mathbf{v}) = \mathbf{u} \cdot (B^T A^T \mathbf{v}), \tag{17.10}$$

for all $\mathbf{u}$ and $\mathbf{v}$ and since also by (17.9)

$$(AB\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot ((AB)^T \mathbf{v}), \tag{17.11}$$

it follows that the right hand side of (17.10) is equal to the right hand side of (17.11) for all $\mathbf{u}$ and $\mathbf{v}$ and hence that

$$(AB)^T = B^T A^T. \tag{17.12}$$

If Q is an orthogonal matrix, i.e. $Q^T Q = I$ it follows that

$$(Q\mathbf{u}) \cdot (Q\mathbf{v}) = (Q^T Q\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v},$$

i.e. $T(\mathbf{u}) = Q\mathbf{u}$ is an orthogonal transformation (17.1).

Orthogonal Projection

If $\mathbf{u}$ and $\mathbf{v}$ are two (column) vectors then $\mathbf{u}^T$ is a row vector and we can write the dot product in terms of the matrix product

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} \tag{17.13}$$

which follows from the row-column rule.

The orthogonal projection onto a subspace V of $\mathbf{R}^m$ with orthonormal basis $\{\mathbf{u}_1,\dots,\mathbf{u}_n\}$ is

$$\mathrm{proj}_V\,\mathbf{x} = (\mathbf{u}_1\cdot\mathbf{x})\mathbf{u}_1 + \cdots + (\mathbf{u}_n\cdot\mathbf{x})\mathbf{u}_n$$

Using (17.13) we can write this as

$$\mathrm{proj}_V\,\mathbf{x} = \mathbf{u}_1(\mathbf{u}_1^T\mathbf{x}) + \cdots + \mathbf{u}_n(\mathbf{u}_n^T\mathbf{x}) = \big(\mathbf{u}_1\mathbf{u}_1^T + \cdots + \mathbf{u}_n\mathbf{u}_n^T\big)\mathbf{x}.$$

Hence

$$\mathrm{proj}_V\,\mathbf{x} = \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} - & \mathbf{u}_1^T & - \\ - & \mathbf{u}_2^T & - \\ & \vdots & \\ - & \mathbf{u}_n^T & - \end{bmatrix} \mathbf{x}$$

We have hence found a matrix formula for the orthogonal projection:

$$P = QQ^T, \qquad \text{where} \qquad Q = \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & | \end{bmatrix}$$

Another important class of matrices are the **symmetric** matrices satisfying $A^T = A$.

Note that it follows that P is symmetric $P^T = (QQ^T)^T = (Q^T)^TQ^T = QQ^T = P$ by (17.12).
Moreover $P^2 = QQ^TQQ^T = Q(Q^TQ)Q^T = QQ^T = P$, since by (17.8) $Q^TQ = I$.
Conversely, a matrix satisfying these two properties is the matrix of an orthogonal projection.

Summary

A linear transformation $T:\mathbf{R}^n\to\mathbf{R}^n$ is called an **orthogonal transformation** if for all $\mathbf{u}$,$\mathbf{v}$

$$T(\mathbf{u})\cdot T(\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}. \tag{17.14}$$

Note that in particular that by taking $\mathbf{v}=\mathbf{u}$ and recalling that $\mathbf{u}\cdot\mathbf{u}=\|\mathbf{u}\|^2$ it follows that

$$\|T(\mathbf{u})\| = \|\mathbf{u}\|. \tag{17.15}$$

The book takes (17.15) as definition of orthogonal but (17.14) also follows from (17.15). It follows from this that an orthogonal transformation also preserve angles:

$$\frac{T(\mathbf{u})\cdot T(\mathbf{v})}{\|T(\mathbf{u})\|\|T(\mathbf{v})\|} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos\phi.$$

If $T(\mathbf{x}) = Q\mathbf{x}$ is an orthogonal transformation with matrix $Q = \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & | \end{bmatrix}$ it follows from the definition (17.14) that the columns of Q are orthonormal

$$\mathbf{u}_i\cdot\mathbf{u}_j = (Q\mathbf{e}_i)\cdot(Q\mathbf{e}_j) = T(\mathbf{e}_i)\cdot T(\mathbf{e}_j) = \mathbf{e}_i\cdot\mathbf{e}_j = \delta_{ij} = \begin{cases} 1, & \text{if } i=j, \\ 0, & \text{if } i\neq j. \end{cases}$$

This can also be written

$$Q^TQ = I. \tag{17.16}$$

A matrix satisfying this, or equivalently that the columns are orthonormal, is called an **orthogonal matrix**. (It should perhaps have been called orthonormal but it is not.)

Recall that the **transpose** A^T of A is the matrix whose ijth entry a_{ij}^T is the jith entry a_{ji} of A: i.e the rows and columns are interchanged. We claim that the transpose satisfy

$$\mathbf{u}\cdot(A^T\mathbf{v}) = (A\mathbf{u})\cdot\mathbf{v}. \tag{17.17}$$

If fact if we let $\mathbf{u}=\mathbf{e}_i$ and $\mathbf{v}=\mathbf{e}_j$ run over the standard basis then the left is a_{ij}^T and the right is a_{ji} which are equal by the definition of A^T. Repeated use of (17.17) gives

$$(AB)^T = B^TA^T. \tag{17.18}$$

If Q is an orthogonal matrix, i.e. $Q^TQ=I$ it follows that

$$(Q\mathbf{u})\cdot(Q\mathbf{v}) = (Q^TQ\mathbf{u})\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{v},$$

i.e. $T(\mathbf{u}) = Q\mathbf{u}$ is an orthogonal transformation (17.14).

If $\mathbf{u}$ and $\mathbf{v}$ are two (column) vectors then $\mathbf{u}^T$ is a row vector and we can write the dot product in terms of the matrix product using the row-column rule:

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} \tag{17.19}$$

The orthogonal projection onto a subspace V of $\mathbf{R}^m$ with orthonormal basis $\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ is

$$\operatorname{proj}_V\mathbf{x} = (\mathbf{u}_1\cdot\mathbf{x})\mathbf{u}_1 + \cdots + (\mathbf{u}_n\cdot\mathbf{x})\mathbf{u}_n$$

using (17.19) one can write this in matrix form $P\mathbf{x}$ where

$$P = QQ^T.$$

Another important class of matrices are the **symmetric** matrices satisfying $A^T = A$.
It follows from using (17.18) that P is symmetric and from using (17.16) that $P^2=P$.
Conversely, a matrix satisfying these two properties is the matrix of an orthogonal projection.