5.3 ORTHOGONAL TRANSFORMATIONS AND ORTHOGONAL MATRICES

Definition 5.3.1 Orthogonal transformations and orthogonal matrices

A linear transformation T from R^n to R^n is called orthogonal if it preserves the length of vectors:

$$\|T(\vec{x})\| = \|\vec{x}\|, \text{ for all } \vec{x} \text{ in } R^n.$$

If $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, we say that A is an orthogonal matrix.

EXAMPLE 1 The rotation

$$T(\vec{x}) = \begin{bmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{bmatrix} \vec{x}$$

is an <u>orthogonal transformation</u> from R^2 to R^2, and

$$A = \begin{bmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{bmatrix} \vec{x}$$

is an <u>orthogonal matrix</u>, for all angles ϕ.

EXAMPLE 2 Reflection

Consider a subspace V of R^n. For a vector $\vec{x}$ in R^n, the vector $R(\vec{x}) = 2proj_V\vec{x} - \vec{x}$ is called the reflection of $\vec{x}$ in V. (see Figure 1).

Show that reflections are orthogonal transformations.

Solution

We can write

$$R(\vec{x}) = proj_V\vec{x} + (proj_V\vec{x} - \vec{x})$$

and

$$\vec{x} = proj_V\vec{x} + (\vec{x} - proj_V\vec{x}).$$

By the pythagorean theorem, we have

$$\|R(\vec{x})\|^2 = \|proj_V\vec{x}\|^2 + \|proj_V\vec{x} - \vec{x}\|^2$$

$$= \|proj_V\vec{x}\|^2 + \|\vec{x} - proj_V\vec{x}\|^2 = \|\vec{x}\|^2.$$

Fact 5.3.2 Orthogonal transformations preserve orthogonality

Consider an orthogonal transformation T from R^n to R^n. If the vectors $\vec{v}$ and $\vec{w}$ in R^n are orthogonal, then so are $T(\vec{v})$ and $T(\vec{w})$.

Proof

By the theorem of Pythagoras, we have to show that

$$\|T(\vec{v}) + T(\vec{w})\|^2 = \|T(\vec{v})\|^2 + \|T(\vec{w})\|^2.$$

Let's see:

$\|T(\vec{v}) + T(\vec{w})\|^2 = \|T(\vec{v} + \vec{w})\|^2$ (T is linear)

$= \|\vec{v} + \vec{w}\|^2$ (T is orthogonal)

$= \|\vec{v}\|^2 + \|\vec{w}\|^2$ ($\vec{v}$ and $\vec{w}$ are orthogonal)

$= \|T(\vec{v})\|^2 + \|T(\vec{w})\|^2$.
($T(\vec{v})$ and $T(\vec{w})$ are orthogonal)

Fact 5.3.3 Orthogonal transformations and orthonormal bases

a. A linear transformation T from R^n to R^n is orthogonal iff the vectors $T(\vec{e_1})$, $T(\vec{e_2})$,...,$T(\vec{e_n})$ form an orthonormal basis of R^n.

b. An $n \times n$ matrix A is orthogonal iff its columns form an orthonormal basis of R^n.

Proof Part(a):
$\Rightarrow$ If T is orthogonal, then, by definition, the $T(\vec{e_i})$ are unit vectors, and by Fact 5.3.2, since $\vec{e_1}$, $\vec{e_2}$,...,$\vec{e_n}$ are orthogonal, $T(\vec{e_1})$, $T(\vec{e_2})$,...,$T(\vec{e_n})$ are orthogonal.

$\Leftarrow$ Conversely, suppose the $T(\vec{e_i})$ form an orthonormal basis.
Consider a vector

$$\vec{x} = x_1\vec{e_1} + x_2\vec{e_2} + \cdots + x_n\vec{e_n}$$

in R^n. Then,

$$\|T(\vec{x})\|^2 = \|x_1 T(\vec{e_1}) + x_2 T(\vec{e_2}) + \cdots + x_n T(\vec{e_n})\|^2$$

$$= \|x_1 T(\vec{e_1})\|^2 + \|x_2 T(\vec{e_2})\|^2 + \cdots + \|x_n T(\vec{e_n})\|^2$$
(by Pythagoras)

$$= x_1^2 + x_2^2 + \cdots + x_n^2$$

$$= \|\vec{x}\|^2$$

Part(b) then follows from Fact 2.1.2.

Warning: A matrix with orthogonal columns need not be orthogonal matrix.

As an example, consider the matrix $A = \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix}$.

EXAMPLE 3 Show that the matrix A is orthogonal:

$$A = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}.$$

Solution

Check that the columns of A form an orthonoraml basis of R^4.

Fact 5.3.4
Products and inverses of orthogonal matrices

a. The product AB of two orthogonal $n \times n$ matrices A and B is orthogonal.

b. The inverse A^{-1} of an orthogonal $n \times n$ matrix A is orthogonal.

Proof
In part (a), the linear transformation $T(\vec{x}) = AB\vec{x}$ preserves length, because $\|T(\vec{x})\| = \|A(B\vec{x})\| = \|B\vec{x}\| = \|\vec{x}\|$. Figure 4 illustrates property (a).

In part (b), the linear transformation $T(\vec{x}) = A^{-1}\vec{x}$ preserves length, because $\|A^{-1}\vec{x}\| = \|A(A^{-1}\vec{x})\|$.

The Transpose of a Matrix

EXAMPLE 4 Consider the orthogonal matrix

$$A = \frac{1}{7}\begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix}.$$

Form another 3×3 matrix B whose ijth entry is the jith entry of A:

$$B = \frac{1}{7}\begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}$$

Note that the rows of B correspond to the columns of A. Compute BA, and explain the result.

Solution

$$BA = \frac{1}{49}\begin{bmatrix} 2 & 6 & 3 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix}\begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix} =$$

$$\frac{1}{49}\begin{bmatrix} 49 & 0 & 0 \\ 0 & 49 & 0 \\ 0 & 0 & 49 \end{bmatrix} = I_3$$

This result is no coincidence: The ijth entry of BA is the dot product of the ith row of B and the jth column of A. By definition of B, this is just the dot product of the ith column of A and the jth column of A. Since A is orthogonal, this product is 1 if $i = j$ and 0 otherwise.

Definition 5.3.5 The transpose of a matrix; symmetric and skew-symmetric matrices
Consider an $m \times n$ matrix A.
The transpose A^T of A is the $n \times m$ matrix whose ijth entry is the jith entry of A: The roles of rows and columns are reversed.
We say that a square matrix A is symmetric if $A^T = A$, and A is called skew-symmetric if $A^T = -A$.

EXAMPLE 5 If $A = \begin{bmatrix} 1 & 2 & 3 \\ 9 & 7 & 5 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 9 \\ 2 & 7 \\ 3 & 5 \end{bmatrix}$.

EXAMPLE 6 The symmetric 2×2 matrices are those of the form $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, for example, $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$.

The symmetric 2×2 matrices form a three-dimensional subspace of $R^{2\times 2}$, with basis $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

The skew-symmetric 2×2 matrices are those of the form $A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$, for example, $A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$. These form a one-dimmensional space with basis $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Note that the transpose of a (column) vector $\vec{v}$ is a row vector: If

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \text{ then } \vec{v}^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}.$$

The transpose give us a convenient way to express the dot product of two (cloumn) vectors as a matrix product.

Fact 5.3.6
If $\vec{v}$ and $\vec{w}$ are two (column) vectors in R^n, then

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}.$$

For example,

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 2.$$

Fact 5.3.7
Consider an $n \times n$ matrix A. The matrix A is orthogonal if (and only if) $A^T A = I_n$ or, equivalently, if $A^{-1} = A^T$.

Proof
To justify this fact, write A in terms of its columns:

$$A = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \dots & \vec{v_n} \\ | & | & & | \end{bmatrix}$$

Then,

$$A^T A = \begin{bmatrix} - & \vec{v_1}^T & - \\ - & \vec{v_2}^T & - \\ & \vdots & \\ - & \vec{v_n}^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \dots & \vec{v_n} \\ | & | & & | \end{bmatrix} =$$

$$\begin{bmatrix} \vec{v_1} \cdot \vec{v_1} & \vec{v_1} \cdot \vec{v_2} & \dots & \vec{v_1} \cdot \vec{v_n} \\ \vec{v_2} \cdot \vec{v_1} & \vec{v_2} \cdot \vec{v_2} & \dots & \vec{v_2} \cdot \vec{v_n} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{v_n} \cdot \vec{v_1} & \vec{v_n} \cdot \vec{v_2} & \dots & \vec{v_n} \cdot \vec{v_n} \end{bmatrix}.$$

By Fact 5.3.3(b) this product is I_n if (and only if) A is orthogonal.

Summary 5.3.8 Orthogonal matrices
Consider an $n \times n$ matrix A. Then, the following statements are equivalent:

1. A is an orthogonal matrix.

2. The transformation $L(\vec{x}) = A\vec{x}$ preserves length, that is, $\|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x}$ in R^n.

3. The columns of A form an orthonormal basis of R^n.

4. $A^T A = I_n$.

5. $A^{-1} = A^T$.

Fact 5.3.9 Properties of the transpose
a. If A is an $m \times n$ matrix and B an $n \times p$ matrix, then

$$(AB)^T = B^T A^T.$$

Note the order of the factors.

b. If an $n \times n$ matrix A is invertible, then so is A^T, and

$$(A^T)^{-1} = (A^{-1})^T.$$

c. For any matrix A,

$$rank(A) = rank(A^T).$$

Proof

a. Compare entries:

ijth entry of $(AB)^T$=jith entry of AB
=(jth row of A)·(ith column of B)

ijth entry of B^TA^T=(ith row of B^T)·(jth column of A^T)
=(ith column of B)·(jth row of A)

b. We know that

$$AA^{-1} = I_n$$

Transposing both sides and using part(a), we find that

$$(AA^{-1})^T = (A^{-1})^T A^T = I_n.$$

By Fact 2.4.9, it follows that

$$(A^{-1})^T = (A^T)^{-1}.$$

c. Consider the row space of A (i.e., the span of the rows of A). It is not hard to show that the dimmension of this space is rank(A) (see Exercise 49-52 in section 3.3):

rank(A^T)=dimension of the span of the columns of A^T
=dimension of the span of the rows of A
=rank(A)

The Matrix of an Orthogonal projection

The transpose allows us to write a formula for the matrix of an orthogonal projection. Consider first the orthogonal projection

$$proj_L\vec{x} = (\vec{v_1} \cdot \vec{x})\vec{v_1}$$

onto a line L in R^n, where $\vec{v_1}$ is a unit vector in L. If we view the vector $\vec{v_1}$ as an $n \times 1$ matrix and the scalar $\vec{v_1} \cdot \vec{x}$ as a 1×1, we can write

$proj_L\vec{x} = \vec{v_1}(\vec{v_1} \cdot \vec{x})$
$= \vec{v_1}\vec{v_1}^T\vec{x}$
$= M\vec{x}$,

where $M = \vec{v_1}\vec{v_1}^T$. Note that $\vec{v_1}$ is an $n \times 1$ matrix and $\vec{v_1}^T$ is $1 \times n$, so that M is $n \times n$, as expected.

More generally, consider the projection

$$proj_v\vec{x} = (\vec{v_1}\cdot\vec{x})\vec{v_1} + \cdots + (\vec{v_m}\cdot\vec{x})\vec{v_m}$$

onto a subspace V of R^n with orthonormal basis $\vec{v_1}$,...,$\vec{v_m}$. We can write

$$proj_v\vec{x} = \vec{v_1}\vec{v_1}^T\vec{x} + \cdots + \vec{v_m}\vec{v_m}^T\vec{x}$$
$$= (\vec{v_1}\vec{v_1}^T + \cdots + \vec{v_m}\vec{v_m}^T)\vec{x}$$

$$= \begin{bmatrix} | & & | \\ \vec{v_1} & \ldots & \vec{v_m} \\ | & & | \end{bmatrix} \begin{bmatrix} - & \vec{v_1}^T & - \\ & \vdots & \\ - & \vec{v_m}^T & - \end{bmatrix} \vec{x}$$

Fact 5.3.10 The matrix of an orthogonal projection

Consider a subspace V of R^n with orthonormal basis $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}$. The matrix of the orthogonal projection onto V is

$$AA^T, \text{ where } A = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_m} \\ | & | & & | \end{bmatrix}.$$

Pay attention to the order of the factors (AA^T as opposed to A^TA).

EXAMPLE 7 Find the matrix of the orthogonal projection onto the subspace of R^4 spanned by

$$\vec{v_1} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v_2} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

Solution

Note that the vectors $\vec{v_1}$ and $\vec{v_2}$ are orthonormal. Therefore, the matrix is

$$AA^T = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Exercises 5.3: 1, 3, 5, 11, 13, 15, 20