Applied Linear Algebra

OTTO BRETSCHER

http://www.prenhall.com/bretscher

Chapter 5
Orthogonality and Least Squares

Chia-Hui Chang
Email: chia@csie.ncu.edu.tw
National Central University, Taiwan

5.1 ORTHONORMAL BASES AND ORTHOGONAL PROJECTIONS

Not all bases are created equal.

Definition. *5.1.1*

Qrthogonality, length, unit vectors
a. Tow vectors $\vec{v}$ and $\vec{w}$ in R^n are called perpendicular or orthogonal if $\vec{v} \cdot \vec{w} = 0$.

b. The length (or magnitude or norm) of a vector $\vec{v}$ in R^n is $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$.

c. A vector $\vec{u}$ in R^n is called a unit vector if its length is 1, (i.e., $\|\vec{u}\| = 1$, or $\vec{u} \cdot \vec{u} = 1$).

Explanation:

If $\vec{v}$ is a nonzero vector in R^n, then

$$\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v}$$

is a unit vector.

Definition. *5.1.2* **Orthonormal vectors**
The vector $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}$ *in* R^n *are called orthonormal if they are all unit vectors and orthogonal to one another:*

$$\vec{v_i} \cdot \vec{v_j} = \begin{cases} 1 & if \quad i=j, \\ 0 & if \quad i \neq j. \end{cases}$$

Example. *1.*

The vectors $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}$ in R^n are orthonormal.

Example. *2.*

For any scalar α, the vectors $\begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \begin{bmatrix} -\sin\alpha \\ \cos\alpha \end{bmatrix}$ are orthonormal.

Example. *3. The vectors*

$$\vec{v_1} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix}$$

in R^4 are orthonormal. Can you find a vector $\vec{v_4}$ in R^4 such that all the vectors $\vec{v_1}$,$\vec{v_2}$,$\vec{v_3}$,$\vec{v_4}$ are orthonormal.

The following properties of orthonormal vectors are often useful:

Fact 5.1.3

a. Orthonormal vectors are linearly independent.

b. Orthonormal vectors $\vec{v_1}$, ..., $\vec{v_n}$ in R^n form a basis of R^n.

Proof

a. Consider a relation

$$c_1\vec{v_1}+c_2\vec{v_2}+\cdots+c_i\vec{v_i}+\cdots+c_m\vec{v_m}=\vec{0}$$

Let us form the dot product of each side of this equation with $\vec{v_i}$:

$$(c_1\vec{v_1}+c_2\vec{v_2}+\cdots+c_i\vec{v_i}+\cdots+c_m\vec{v_m})\cdot\vec{v_i} = \vec{0}\cdot\vec{v_i}=0.$$

Because the dot product is distributive.

$$c_i(\vec{v_i}\cdot\vec{v_i})=0$$

Therefore, $c_i = 0$ for all $i = 1, \ldots, m$.

b. Any n linearly independent vectors in R^n form a basis of R^n.

Definition. *5.1.4* **Orthogonal complement** *Consider a subspace V of R^n. The orthogonal complement $V^{\perp}$ of V is the set of those vectors $\vec{x}$ in R^n that are orthogonal to all vectors in V:*

$$V^{\perp}=\{\vec{x} \text{ in } R^n\text{: } \vec{v}\cdot\vec{x}=0, \text{ for all } \vec{v} \text{ in } V \}.$$

Fact 5.1.5 If V is a subspace of R^n, then its orthogonal complement $V^{\perp}$ is a subspace of R^n as well.

Proof

We will verify that $V^{\perp}$ is closed under scalar multiplication and leave the verification of the two other properties as Exercise 23. Consider a vector $\vec{w}$ in $V^{\perp}$ and a scalar k. We have to show that $k\vec{w}$ is orthogonal to all vectors $\vec{v}$ in V. Pick an arbitrary vector $\vec{v}$ in V. Then, $(k\vec{w})\cdot\vec{v}=k(\vec{w}\cdot\vec{v})=0$, as claimed.

Orthogonal projections

See Figure 5.

The **orthogonal projection** of a vector $\vec{x}$ onto one-dimentaional subspace V with basis $\vec{v}_1$ (unit vector) is computed by:

$$proj_V \vec{x} = \vec{w} = (\vec{v_1} \cdot \vec{x})\vec{v_1}$$

Now consider a subspace V with arbitrary dimension m. Suppose we have an orthonormal basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ of V. Find $\vec{w}$ in V such that $\vec{x} - \vec{w}$ is in $V^{\perp}$. Let

$$\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$

It is required that

$$\vec{x - w} = \vec{x} - c_1\vec{v}_1 - c_2\vec{v}_2 - \cdots - c_m\vec{v}_m$$

is perpendicular to V; i.e.:

$$\vec{v}_i \cdot (\vec{x-w}) = \vec{v}_i \cdot (\vec{x} - c_1\vec{v}_1 - c_2\vec{v}_2 - \cdots - c_m\vec{v}_m)$$

$$= \vec{v}_i \cdot \vec{x} - c_1(\vec{v}_i \cdot \vec{v}_1) - \cdots - c_i(\vec{v}_i \cdot \vec{v}_i) - \cdots - c_m(\vec{v}_i \cdot \vec{v}_m)$$

$$= \vec{v}_i \cdot \vec{x} - c_i = 0$$

The equation holds if $c_i = \vec{v}_i \cdot \vec{x}$.
Therefore, there is a unique $\vec{w}$ in V such that $\vec{x} - \vec{w}$ is in $V^{\perp}$, namely,

$$\vec{w} = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + (\vec{v}_2 \cdot \vec{x})\vec{v}_2 + \cdots + (\vec{v}_m \cdot \vec{x})\vec{v}_m$$

Fact 5.1.6 Orthogonal projection
Consider a subspace V of R^n with orthonormal basis $\vec{v_1}$,$\vec{v_2}$,...,$\vec{v_m}$. For any vector $\vec{x}$ in R^n, there is a unique vector $\vec{w}$ in V such that $\vec{x}$-$\vec{w}$ is in $V^{\perp}$. This vector $\vec{w}$ is called the orthogonal projection of $\vec{x}$ onto V, denoted by $proj_V\vec{x}$. We have the formula

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

The transformation $T(\vec{x}) = proj_V\vec{x}$ from R^n to R^n is linear.

Example. *4*

Consider the subspace V=im(A) of R^4. where

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

Find $proj_V \vec{x}$, for

$$\vec{x} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 7 \end{bmatrix}.$$

Solution

The two columns of A form a basis of V. Since they happen to be orthogonal, we can construct an orthonormal basis of V merely by dividing these two vectors by their length (2 for both vectors):

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

Then,

$proj_V \vec{x}$=$(\vec{v_1} \cdot \vec{x})\vec{v_1}$+$(\vec{v_2} \cdot \vec{x})\vec{v_2}$=6$\vec{v_1}$+ 2$\vec{v_2}$=

$$\begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \\ 4 \end{bmatrix}.$$

To check this answer, verify that $\vec{x}$-$proj_V \vec{x}$ is perpendicular to both $\vec{v_1}$ and $\vec{v_2}$.

What happens when we apply Fact 5.1.6 to the subspace $V = R^n$ of R^n with orthonormal basis $\vec{v_1}$, $\vec{v_2}, \cdots, \vec{v_n}$? Clearly, $proj_V \vec{x} = \vec{x}$, for all $\vec{x}$ in R_n. Therefore,

$$\vec{x} = (\vec{v_1} \cdot \vec{x})\vec{v_1} + \cdots + (\vec{v_n} \cdot \vec{x})\vec{v_n},$$

for all $\vec{x}$ in R^n. See Figure 7.

Fact 5.1.7

Consider an orthonormal basis $\vec{v_1}, \cdots, \vec{v_n}$ of R^n. Then,

$$\vec{x} = (\vec{v_1} \cdot \vec{x})\vec{v_1} + \cdots + (\vec{v_n} \cdot \vec{x})\vec{v_n},$$

for all $\vec{x}$ in R^n.

This is useful for compute the B-coordinate, since $c_i = \vec{v}_i \cdot \vec{x}$.

Example. *5*

By using paper and pencil, express the vector $\vec{x}= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ as a linear combination of

$$\vec{v_1} = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \vec{v_2} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \vec{v_3} = \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}.$$

Solution

Since $\vec{v_1}$,$\vec{v_2}$,$\vec{v_3}$ is an orthonormal basis of R^3, we have

$\vec{x} = (\vec{v_1} \cdot \vec{x})\vec{v_1} + (\vec{v_2} \cdot \vec{x})\vec{v_2} + (\vec{v_3} \cdot \vec{x})\vec{v_3} = 3\vec{v_1} + \vec{v_2} + 2\vec{v_3}$.

From Pythagoras to Cauchy

Example. *6*

Consider a line L in R^3 and a vector $\vec{x}$ in R^3. What can you say about the relationship between the lengths of the vectors $\vec{x}$ and $proj_L\vec{x}$?

Solution

Applying the Pythagorean theorem to the shaded right triangle in Figure 8, we find that

$$\| proj_L\vec{x} \| \leq \| \vec{x} \| .$$

The statement is an equality if (and only if) $\vec{x}$ is on L.

Does this inequality hold in higher dimensional cases? We have to examine whether the Pythagorean theorem holds in R^n.

Fact 5.1.8 Pythagorean theorem
Consider two vectors $\vec{x}$ and $\vec{y}$ in R^n. The equation

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

holds if (and only if) $\vec{x}$ and $\vec{y}$ are orthogonal. (See Figure 9.)

Proof The verification is straightforward:
$\| \vec{x} + \vec{y} \|^2 = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y})$
$= \vec{x} \cdot \vec{x} + 2(\vec{x} \cdot \vec{y}) + \vec{y} \cdot \vec{y}$
$= \| \vec{x} \|^2 + 2(\vec{x} \cdot \vec{y}) + \| \vec{y} \|^2$
$= \| \vec{x} \|^2 + \| \vec{y} \|^2$
if (and only if) $\vec{x} \cdot \vec{y} = 0$.

Fact 5.1.9 Consider a subspace V of R^n and a vector $\vec{x}$ in R^n. Then,

$$\| proj_V \vec{x} \| \leq \| \vec{x} \|.$$

The statement is an equality if(and only if) $\vec{x}$ is in V.

Proof we can write $\vec{x} = proj_V \vec{x} + (\vec{x} - proj_V \vec{x})$and apply the Pythagorean theorem(see Figure 10):

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

It follows that $\| proj_V \vec{x} \| \leq \| \vec{x} \|$, as claimed.

Let V be a one-dimensional subspace of R^n spanned by a(nonzero) vector $\vec{y}$. We introduce the unit vector

$$\vec{u} = \frac{1}{\|\vec{y}\|}\vec{y}$$

in V. (See Figure 11.)
We know that

$$proj_V\vec{x} = (\vec{u}\cdot\vec{x})\vec{u} = \frac{1}{\|\vec{y}\|^2}(\vec{y}\cdot\vec{x})\vec{y}.$$

for any $\vec{x}$ in R^n. Fact 5.1.9 tells us that

$$\|\ \vec{x}\ \| \geq \|\ proj_V\vec{x}\ \| = \|\ \frac{1}{\|\vec{y}\|^2}(\vec{y}\cdot\vec{x})\vec{y}\ \| = \frac{1}{\|\vec{y}\|^2}|\vec{y}\cdot\vec{x}|\ \|\ \vec{y}\ \|.$$

To justify the last step, note that $\|\ k\vec{v}\ \| = |k|\ \|\ \vec{v}\ \|$, for all vectors $\vec{v}$ in R^n and all scalars k. (See Exercise 25(a).) We conclude that

$$\frac{|\vec{x}\cdot\vec{y}|}{\|\vec{y}\|} \leq \|\ \vec{x}\ \|.$$

Fact 5.1.10 Cauchy-Schwarz inequality
If $\vec{x}$ and $\vec{y}$ are vectors in R^n, then

$$|\vec{x} \cdot \vec{y}| \leq \| \vec{x} \| \| \vec{y} \|.$$

The statement is an equality if (and only if) $\vec{x}$ and $\vec{y}$ are parallel.

Definition. *5.1.11*

Angle between two vectors Consider two nonzero vectors $\vec{x}$ and $\vec{y}$ in R^n. The angle α between these vectors is defined as

$$\cos \alpha = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|}.$$

Note that α is between 0 and π, by definition of the inverse cosine functiion.

Example. *7*

Find the angle between the vectors

$$\vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Solution

$$\cos\alpha = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{1}{1\cdot 2} = \frac{1}{2}$$

$$\alpha = \frac{\pi}{3}$$

Correlation

Consider two characteristics of a population, with deviation vectors $\vec{x}$ and $\vec{y}$. There is a positive correlation between the two characteristics if (and only if) $\vec{x} \cdot \vec{y} > 0$.

Definition. *5.1.12*

Correlation coefficient
The correlation coefficient r between two characteristics of a population is the cosine of the angle α between the deviation vectors $\vec{x}$ and $\vec{y}$ for the two characteristics:

$$r = cos(\alpha) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|}$$

Exercise 5.1: 7, 9, 12, 19, 23, 24, 25, 28