

Lecture 12: Orthogonal Projections and Projection Matrices

1 Orthogonal Projection

From the previous lecture,

$$\mathbb{R}^n = L \oplus L^\perp.$$

Therefore, every $x \in \mathbb{R}^n$ has a unique decomposition

$$x = y + z, \quad y \in L, \quad z \in L^\perp.$$

Definition

The vector y is called the **orthogonal projection** of x onto L .

We write

$$P_L(x) = y.$$

Thus,

$$x = P_L(x) + (x - P_L(x)),$$

where

$$P_L(x) \in L$$

and

$$x - P_L(x) \in L^\perp.$$

Geometrically,

$$P_L(x)$$

is the part of x lying inside L , while

$$x - P_L(x)$$

is the part perpendicular to L .

—

2 Computing the Projection from an Orthonormal Basis

Suppose

$$\{q_1, \dots, q_r\}$$

is an orthonormal basis of L .

Theorem 2.1 (Projection Formula). *For every $x \in \mathbb{R}^n$,*

$$P_L(x) = \sum_{j=1}^r \langle q_j, x \rangle q_j.$$

Proof

Define

$$y = \sum_{j=1}^r \langle q_j, x \rangle q_j.$$

Clearly,

$$y \in L.$$

We will show that

$$x - y \in L^\perp.$$

It is enough to show that $x - y$ is orthogonal to every basis vector q_k .

For $k = 1, \dots, r$,

$$\begin{aligned} \langle x - y, q_k \rangle &= \left\langle x - \sum_{j=1}^r \langle q_j, x \rangle q_j, q_k \right\rangle \\ &= \langle x, q_k \rangle - \sum_{j=1}^r \langle q_j, x \rangle \langle q_j, q_k \rangle. \end{aligned}$$

Since the q_j 's are orthonormal,

$$\langle q_j, q_k \rangle = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

Thus

$$\langle x - y, q_k \rangle = \langle x, q_k \rangle - \langle q_k, x \rangle.$$

Since we are working over $\mathbb{R}$,

$$\langle x, q_k \rangle = \langle q_k, x \rangle.$$

Therefore,

$$\langle x - y, q_k \rangle = 0.$$

Hence $x - y$ is orthogonal to every vector in L , so

$$x - y \in L^\perp.$$

Thus

$$x = y + (x - y)$$

is the unique decomposition of x into one vector in L and one vector in $L^\perp$.

Therefore,

$$P_L(x) = y = \sum_{j=1}^r \langle q_j, x \rangle q_j.$$

□

—

Important Consequence: Independence of the Basis Although the formula

$$P_L(x) = \sum_j \langle q_j, x \rangle q_j$$

uses an orthonormal basis, the final vector does ****not**** depend on which orthonormal basis of L we choose.

The reason is that $P_L(x)$ is characterized uniquely by

$$P_L(x) \in L$$

and

$$x - P_L(x) \in L^\perp.$$

Any orthonormal basis used in the projection formula must produce the same unique vector.

—

3 Basic Properties of Orthogonal Projection

Theorem 3.1. *Let $L \subseteq \mathbb{R}^n$ be a linear subspace. Then:*

1. $P_L(x) \in L$. 2. $x - P_L(x) \in L^\perp$. 3. If $x \in L$, then

$$P_L(x) = x.$$

4. P_L is linear.

5. P_L is idempotent:

$$P_L(P_L(x)) = P_L(x).$$

6. The Pythagorean identity holds:

$$\|x\|_2^2 = \|P_L(x)\|_2^2 + \|x - P_L(x)\|_2^2.$$

4 Projection Is the Closest Point

The geometric interpretation of a projection can be made precise.

Theorem 4.1 (Best Approximation Property). *Let $L \subseteq \mathbb{R}^n$ be a linear subspace and let $x \in \mathbb{R}^n$. Then*

$$P_L(x)$$

is the unique point in L closest to x .

In other words,

$$P_L(x) = \arg \min_{y \in L} \|x - y\|_2.$$

Proof

Take any $y \in L$.

Write

$$x - y = (x - P_L(x)) + (P_L(x) - y).$$

Now,

$$x - P_L(x) \in L^\perp,$$

while

$$P_L(x) - y \in L.$$

Therefore,

$$x - P_L(x) \perp P_L(x) - y.$$

By the Pythagorean theorem,

$$\|x - y\|_2^2 = \|x - P_L(x)\|_2^2 + \|P_L(x) - y\|_2^2.$$

Hence

$$\|x - y\|_2^2 \geq \|x - P_L(x)\|_2^2.$$

Therefore,

$$\|x - y\|_2 \geq \|x - P_L(x)\|_2.$$

Equality occurs if and only if

$$\|P_L(x) - y\|_2^2 = 0,$$

which occurs if and only if

$$y = P_L(x).$$

Thus $P_L(x)$ is the unique minimizer. $\square$

—

5 Projection onto the Orthogonal Complement

Since

$$x = P_L(x) + (x - P_L(x))$$

and

$$x - P_L(x) \in L^\perp,$$

we immediately obtain

$$P_{L^\perp}(x) = x - P_L(x).$$

Therefore,

$$\boxed{P_{L^\perp} = I - P_L.}$$

This also gives another useful identity.

Theorem

For every subspace $L \subseteq \mathbb{R}^n$,

$$\boxed{(L^\perp)^\perp = L.}$$

Proof

First,

$$L \subseteq (L^\perp)^\perp$$

because every vector in L is orthogonal to every vector in $L^\perp$.

For the reverse inclusion, take

$$x \in (L^\perp)^\perp.$$

Then

$$P_{L^\perp}(x) = 0.$$

But

$$P_{L^\perp}(x) = x - P_L(x).$$

Therefore,

$$x - P_L(x) = 0,$$

so

$$x = P_L(x) \in L.$$

Thus

$$(L^\perp)^\perp \subseteq L.$$

Therefore,

$$(L^\perp)^\perp = L.$$

□

—

6 Matrix Representation of an Orthogonal Projection

Let

$$\{q_1, \dots, q_r\}$$

be an orthonormal basis of L , and form the matrix

$$Q = \begin{pmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_r \\ | & | & \cdots & | \end{pmatrix} \in \mathbb{R}^{n \times r}.$$

Since the columns are orthonormal,

$$\boxed{Q^\top Q = I_r.}$$

For $x \in \mathbb{R}^n$,

$$Q^\top x = \begin{pmatrix} \langle q_1, x \rangle \\ \vdots \\ \langle q_r, x \rangle \end{pmatrix}.$$

Therefore,

$$QQ^\top x = \sum_{j=1}^r \langle q_j, x \rangle q_j.$$

But the right-hand side is exactly $P_L(x)$.

Hence the matrix representing P_L is

$$\boxed{P_L = QQ^\top.}$$

Equivalently,

$$P_L = \sum_{j=1}^r q_j q_j^\top.$$

7 Properties of the Projection Matrix

Let

$$P = QQ^\top.$$

Then

$$P^\top = (QQ^\top)^\top = QQ^\top = P.$$

Thus

$$\boxed{P^\top = P.}$$

So an orthogonal projection matrix is symmetric.

Also,

$$P^2 = QQ^\top QQ^\top.$$

Since

$$Q^\top Q = I_r,$$

we obtain

$$P^2 = Q(Q^\top Q)Q^\top = QI_r Q^\top = QQ^\top = P.$$

Thus

$$\boxed{P^2 = P.}$$

So an orthogonal projection matrix is idempotent.

Furthermore,

$$P_{L^\perp} = I - P_L,$$

so the matrix projecting onto $L^\perp$ is

$$\boxed{I - QQ^\top.}$$

8 Example

Let

$$L = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

An orthonormal basis is

$$q_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad q_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus

$$Q = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}.$$

The projection matrix is

$$P_L = QQ^\top = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For example, if

$$x = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix},$$

then

$$P_L x = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

The residual is

$$x - P_L x = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Indeed,

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

is orthogonal to both basis vectors of L .

—