1   Overview

Main ideas:

1. two characterizations of the orgthogonal projection of a vector onto a subspace

2. orthogonal decomposition theorem

3. best approximation theorem, distance from a point to a subspace

4. formula for a projection onto a subspace using an orthonormal basis for the subspace

Examples in text:

1. write a vector as a sum of a vector in a given subspace and a vector in its orthogonal complement

2. write a vector as a sum of a vector in a given subspace and a vector in its orthogonal complement

3. write a vector in $\mathbb{R}^3$ as a sum of a vector in a given subspace and a vector in its orthogonal complement

4. find the closest point in a 2D subspace of $\mathbb{R}^3$ to a point in $\mathbb{R}^3$

5. compute the distance between a point and a plane in $\mathbb{R}^3$

2   Discussion and Worked Examples

Recall that one of the motivations for introducing a dot product on $\mathbb{R}^n$ was to be able to discuss approximation of vectors; in order for approximation of vectors to make sense we need the notions of close and far, i.e. a notion of distance, which comes from the dot product. In particular, suppose we wish to approximate a vector v with a vector in a certain subspace W. How could we find the vector in W that is closest to v?

Draw a line through the origin and a vector not on the line. Find and mark the point on the line that is closest to the tip of the vector you drew. That point corresponds to a vector through the origin. What do you notice about this vector? It is the orthogonal projection of the first vector onto the line.

This is true in general: given a vector v in $\mathbb{R}^n$ and a subspace W of $\mathbb{R}^n$, the vector in W that is closest to v is the orthogonal projection of v onto W. This is the best approximation of v in W.

Definition. Let v be a vector in $\mathbb{R}^n$, and let W be a subspace of $\mathbb{R}^n$. The *orthogonal projection of v onto W*, denoted $\text{proj}_W v$, is the unique vector in W such that $v - \text{proj}_W(v)$ is orthogonal to W (i.e. in $W^{\perp}$).

(This means that v can be written uniquely as the sum of a vector in W and a vector in $W^{\perp}$.)

Note that our definition needs some justification. It is not a priori clear that the projection vector exists or that it is unique. We will justify the definition after looking at an example.

Example Find the orthogonal projection of $v = (-1, 4, 3)$ onto the subspace W of $\mathbb{R}^3$ spanned by the orthogonal vectors $w_1 = (1, 1, 0)$ and $w_2 = (-1, 1, 0)$.

Let's find the projections of v onto w_1 and w_2 and add them together. Certainly that will yield a vector (call it $\tilde{w}$) in W, but it's not clear that it will be the orthogonal projection of v onto W.

$$\text{proj}_W v \ \overset{??}{=} \ \tilde{w} \ = \ \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} \, w_1 \ + \ \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle} \, w_2 \ = \ \frac{3}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \ + \ \frac{5}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \ = \ \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}$$

In order for this to be the projection vector, we need $v - \tilde{w}$ to be orthogonal to W. Well,

$$v - \tilde{w} \;=\; \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} - \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

This vector is clearly orthogonal to w_1 and w_2, and thus to all of W. So we have indeed found the projection of v onto W:

$$\mathrm{proj}_W v \;=\; \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle}\, w_1 \;+\; \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle}\, w_2 \;=\; \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}$$

The orthogonal decomposition of v with respect to W and $W^\perp$ is:

$$\begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} \;=\; \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

It turns out that it really matters that we started with an *orthogonal* basis $\{w_1, w_2\}$ for W. The reason why will become clear when we justify our definition of the orthogonal projection.

Note. There is a nice picture in the textbook illustrating how adding the projections onto orthogonal basis vectors does yield the desired projection vector.

Now we return to the question of proving the existence and uniqueness of the orthogonal projection of a vector onto a subspace, in order to justify the definition we gave. We will appeal to a result proven in Section 6.4, namely that any subspace of $\mathbb{R}^n$ has an orthogonal basis.

Let $\{w_1, \ldots, w_p\}$ be an orthogonal basis for W. Then the sum of the projections of v onto w_1, w_2, $\ldots$, and w_p, namely the vector

$$\tilde{w} \;=\; \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle}\, w_1 \;+\; \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle}\, w_2 \;+\; \ldots \;+\; \frac{\langle v, w_p \rangle}{\langle w_p, w_p \rangle}\, w_p$$

is clearly in W, and we notice that

$$\langle \tilde{w}, w_1 \rangle \;=\; \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} \langle w_1, w_1 \rangle + 0 + \cdots + 0 \;=\; \langle v, w_1 \rangle$$

since w_1 is orthogonal to w_2, $\ldots$, w_p. So,

$$\langle v - \tilde{w}, w_1 \rangle \;=\; \langle v, w_1 \rangle - \langle \tilde{w}, w_1 \rangle \;=\; \langle v, w_1 \rangle - \langle v, w_1 \rangle \;=\; 0$$

i.e. $v - \tilde{w}$ is orthogonal to w_1. Similarly $\tilde{w}$ is orthogonal to w_2, $\ldots$, and w_p. Since $\{w_1, \ldots, w_p\}$ spans W, this implies that $v - \tilde{w}$ is in $W^\perp$. Thus we can conclude that there is a vector (namely $\tilde{w}$) in W such that $v - \tilde{w}$ is in $W^\perp$.

It remains to show that this vector is unique. Suppose w' is another vector in W such that $v - w'$ is in $W^\perp$. Then $(v - \tilde{w}) - (v - w')$ is in $W^\perp$, since both $v - \tilde{w}$ and $v - w'$ are. But on the other hand $(v - \tilde{w}) - (v - w') = w' - w$, so it is clearly in W. Thus $(v - \tilde{w}) - (v - w') = 0$, i.e. $w' = \tilde{w}$.

Thus we have justified our definition of the orthogonal projection of a vector onto a subspace, and we have essentially proven the **Orthogonal Decomposition Theorem**: given a subspace W of $\mathbb{R}^n$, a vector v can be uniquely expressed as the sum of a vector in W (the projection of v onto W) and a vector in $W^\perp$ and that the projection can be computed by adding the projections of v onto basis vectors for W.

(The only part of the theorem that we haven't explicitly proven is that the component of v in $W^\perp$ is unique, but this follows immediately from the uniqueness of the projection.)

Our intuition from the $\mathbb{R}^2$ case leads us to think that this is the vector in W that is closest to v. We now turn our attention to justifying this.

Theorem (Best Approximation Theorem). *Let W be a subspace of $\mathbb{R}^n$ and v a vector in $\mathbb{R}^n$. The projection of v onto W is the unique vector in W that is closest to v.*

Proof. Let w be any vector in W. Then

$$v - w \;=\; (v - \text{proj}_W v) \;+\; (\text{proj}_W v - w)$$

We know that $(v - \text{proj}_W v)$ is in $W^\perp$, and certainly $(\text{proj}_W v - w)$ is in W. Thus, $(v - \text{proj}_W v)$ and $(\text{proj}_W v - w)$ are orthogonal. By the Pythagorean Theorem,

$$\|v - w\|^2 \;=\; \|v - \text{proj}_W v\|^2 \;+\; \|\text{proj}_W v - w\|^2$$

Thus the smallest $\|v - w\|$ can be is $\|v - \text{proj}_W v\|$, and this occurs precisely when $w = \text{proj}_W v$. $\qquad\square$

Example Find the best approximation to $v = (-1, 4, 3)$ in the subspace $W = \text{Span}\big((1, 1, 0), (-1, 1, 0)\big)$. How far from v is it?

By the Best Approximation Theorem, there is a unique vector in W that is closest to v, and it is the projection of v onto W. We computed this in the previous example. Thus $\text{proj}_W v = (-1, 4, 0)$ is the best approximation to v in W.

The distance between v and $\text{proj}_W v$ is the norm of their difference.

$$v - \text{proj}_W v \;=\; \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} - \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

Thus the distance between v and $\text{proj}_W v$ is

$$\|v - \text{proj}_W v\| \;=\; \sqrt{0 + 0 + 9} \;=\; 3$$

Since the closest vector to v in W is 3 units away from v, we say that W is 3 units away from v. This is how we define the distance between a vector (or a point) and a subspace.