

Projections Onto a Subspace

We will construct such a basis one vector at a time, so, for now, let us assume that we have an orthonormal set $\{\vec{v}_1, \dots, \vec{v}_k\}$, and we want to find a vector \vec{v}_{k+1} such that the set $\{\vec{v}_1, \dots, \vec{v}_k, \vec{v}_{k+1}\}$ is orthonormal.

Well, if we can find any vector \vec{x} such that $\vec{x} \notin \text{Span}\{\vec{v}_1, \dots, \vec{v}_k\}$, then we can split \vec{x} into two pieces:

1. The part of \vec{x} that is in $\text{Span}\{\vec{v}_1, \dots, \vec{v}_k\}$
2. The part of \vec{x} that is orthogonal to $\text{Span}\{\vec{v}_1, \dots, \vec{v}_k\}$

We did something similar to this in Linear Algebra 1, when we looked at the values $\text{proj}_{\vec{y}}\vec{x}$ and $\text{perp}_{\vec{y}}\vec{x}$. So what we want to do is expand these definitions to now look at the projection of a vector \vec{x} onto a subspace \mathbb{S} , instead of just a vector.

Recalling that $\text{proj}_{\vec{y}}\vec{x}$ is a scalar multiple of \vec{y} , we will now define $\text{proj}_{\mathbb{S}}\vec{x}$ to be a linear combination of the basis vectors for \mathbb{S} , as this would be the generalization of a scalar multiple of one vector.

But $\text{proj}_{\vec{y}}\vec{x}$ wasn't just any scalar multiple of \vec{y} . The scalar was $\frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2}$. So, to make sure that our definitions coincide

in the case that \mathbb{S} is the span of a single vector, we will take the scalars in our expanded definition to be $\frac{\vec{x} \cdot \vec{v}_i}{\|\vec{v}_i\|^2}$.

Projections Onto a Subspace

Definition: Let \mathbb{S} be a k -dimensional subspace of \mathbb{R}^n and let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ be an orthogonal basis of \mathbb{S} . If \vec{x} is any vector in \mathbb{R}^n , the **projection of \vec{x} onto \mathbb{S}** is defined to be

$$\text{proj}_{\mathbb{S}}\vec{x} = \frac{\vec{x} \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 + \frac{\vec{x} \cdot \vec{v}_2}{\|\vec{v}_2\|^2} \vec{v}_2 + \dots + \frac{\vec{x} \cdot \vec{v}_k}{\|\vec{v}_k\|^2} \vec{v}_k$$

Note: This definition only works if \mathcal{B} is an orthogonal basis. We will not consider projections based on arbitrary bases.

Projections Onto a Subspace

Definition: The **projection of \vec{x} perpendicular to \mathbb{S}** is defined to be

$$\text{perp}_{\mathbb{S}}\vec{x} = \vec{x} - \text{proj}_{\mathbb{S}}\vec{x}$$

Now, our hope is that $\text{perp}_{\mathbb{S}}\vec{x}$ will, in fact, be an element of \mathbb{S}^{\perp} . And it turns out that it is. To verify this, we will show that $\text{perp}_{\mathbb{S}}\vec{x} \cdot \vec{v}_i = 0$ for all vectors in our basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ for \mathbb{S} .

$$\begin{aligned} \text{perp}_{\mathbb{S}}\vec{x} \cdot \vec{v}_i &= (\vec{x} - \text{proj}_{\mathbb{S}}\vec{x}) \cdot \vec{v}_i \\ &= \vec{x} \cdot \vec{v}_i - \text{proj}_{\mathbb{S}}\vec{x} \cdot \vec{v}_i \\ &= \vec{x} \cdot \vec{v}_i - \left(\frac{\vec{x} \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 + \dots + \frac{\vec{x} \cdot \vec{v}_k}{\|\vec{v}_k\|^2} \vec{v}_k \right) \cdot \vec{v}_i \\ &= \vec{x} \cdot \vec{v}_i - \left(\frac{\vec{x} \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 \cdot \vec{v}_i + \dots + \frac{\vec{x} \cdot \vec{v}_k}{\|\vec{v}_k\|^2} \vec{v}_k \cdot \vec{v}_i \right) \\ &= \vec{x} \cdot \vec{v}_i - (0 + \dots + 0 + \frac{\vec{x} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} \vec{v}_i \cdot \vec{v}_i + 0 + \dots + 0) \\ &= \vec{x} \cdot \vec{v}_i - \frac{\vec{x} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} \vec{v}_i \cdot \vec{v}_i \\ &= \vec{x} \cdot \vec{v}_i - \frac{\vec{x} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} (\|\vec{v}_i\|^2) \\ &= \vec{x} \cdot \vec{v}_i - \vec{x} \cdot \vec{v}_i \\ &= 0 \end{aligned}$$

Projections Onto a Subspace

Example

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$ be an orthogonal basis for \mathbb{S} and let $\vec{x} = \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix}$. To determine $\text{proj}_{\mathbb{S}}\vec{x}$ and $\text{perp}_{\mathbb{S}}\vec{x}$, we will

first want to do the following calculations:

$$\begin{aligned} \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} &= 4 + 10 - 6 = 8 \\ \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} &= 4 + 5 + 2 = 11 \\ \left\| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\|^2 &= 1^2 + 2^2 + 3^2 = 14 \\ \left\| \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\|^2 &= 1^2 + 1^2 + (-1)^2 = 3 \end{aligned}$$

Projections Onto a Subspace

Example

Then we have that

$$\begin{aligned}\text{proj}_S \vec{x} &= \frac{8}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \frac{11}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ &= \frac{12}{21} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \frac{77}{21} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 89/21 \\ 101/21 \\ -41/21 \end{bmatrix}\end{aligned}$$

and this means that

$$\begin{aligned}\text{perp}_S \vec{x} &= \vec{x} - \text{proj}_S \vec{x} \\ &= \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} - \begin{bmatrix} 89/21 \\ 101/21 \\ -41/21 \end{bmatrix} \\ &= \begin{bmatrix} -5/21 \\ 4/21 \\ -1/21 \end{bmatrix}\end{aligned}$$

Projections Onto a Subspace

Example

We can verify our calculations by noticing that

$$\begin{aligned}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -5/21 \\ 4/21 \\ -1/21 \end{bmatrix} &= -\frac{5}{21} + \frac{8}{21} - \frac{3}{21} = 0 \\ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} -5/21 \\ 4/21 \\ -1/21 \end{bmatrix} &= -\frac{5}{21} + \frac{4}{21} + \frac{1}{21} = 0\end{aligned}$$

Projections Onto a Subspace

Theorem 7.2.2 (Approximation Theorem)

Let \mathbb{S} be a subspace of \mathbb{R}^n . Then, for any $\vec{x} \in \mathbb{R}^n$, the unique vector $\vec{s} \in \mathbb{S}$ that minimizes the distance $\|\vec{x} - \vec{s}\|$ is $\vec{s} = \text{proj}_{\mathbb{S}}\vec{x}$.

Proof

Let $\{\vec{v}_1, \dots, \vec{v}_k\}$ be an orthonormal basis for \mathbb{S} and let $\{\vec{v}_{k+1}, \dots, \vec{v}_n\}$ be an orthonormal basis for \mathbb{S}^\perp , so that $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for \mathbb{R}^n . Then for any $\vec{x} \in \mathbb{R}^n$, there are scalars x_1, \dots, x_n such that

$$\vec{x} = x_1\vec{v}_1 + \dots + x_k\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \dots + x_n\vec{v}_n$$

Let $\vec{s} = s_1\vec{v}_1 + \dots + s_k\vec{v}_k$ be an element of \mathbb{S} . Then we also have that

$$\vec{s} = s_1\vec{v}_1 + \dots + s_k\vec{v}_k + 0\vec{v}_{k+1} + \dots + 0\vec{v}_n,$$

and so we can write $\vec{x} - \vec{s}$ as:

$$\begin{aligned} \vec{x} - \vec{s} &= (x_1\vec{v}_1 + \dots + x_k\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \dots + x_n\vec{v}_n) - (s_1\vec{v}_1 + \dots + s_k\vec{v}_k + 0\vec{v}_{k+1} + \dots + 0\vec{v}_n) \\ &= (x_1 - s_1)\vec{v}_1 + \dots + (x_k - s_k)\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \dots + x_n\vec{v}_n \end{aligned}$$

In order to minimize $\|\vec{x} - \vec{s}\|$, we will minimize the easier to calculate $\|\vec{x} - \vec{s}\|^2$.

Recall that since \mathcal{B} is an orthonormal basis, and as we have written $\vec{x} - \vec{s}$ in terms of \mathcal{B} coordinates, we can still calculate $\|\vec{x} - \vec{s}\|^2$ the usual way: we sum the square of the coefficients.

Projections Onto a Subspace

Proof

$$\|\vec{x} - \vec{s}\|^2 = (x_1 - s_1)^2 + \dots + (x_k - s_k)^2 + x_{k+1}^2 + \dots + x_n^2$$

And clearly this value is minimized by setting $s_i = x_i$, so that $x_i - s_i = 0$ for $i = 1, \dots, k$. This means we have shown that the vector \vec{s} in \mathbb{S} that minimizes the distance $\|\vec{x} - \vec{s}\|$ is

$$\vec{s} = x_1\vec{v}_1 + \dots + x_k\vec{v}_k$$

To see that this is equal to $\text{proj}_{\mathbb{S}}\vec{x}$, let's first recall that

$$\text{proj}_{\mathbb{S}}\vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \dots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

since $\{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthonormal basis for \mathbb{S} . And then we notice that, for any $i = 1, \dots, k$, we get

$$\begin{aligned} \vec{x} \cdot \vec{v}_i &= (x_1\vec{v}_1 + \dots + x_k\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \dots + x_n\vec{v}_n) \cdot \vec{v}_i \\ &= x_1(\vec{v}_1 \cdot \vec{v}_i) + \dots + x_i(\vec{v}_i \cdot \vec{v}_i) + \dots + x_k(\vec{v}_k \cdot \vec{v}_i) + x_{k+1}(\vec{v}_{k+1} \cdot \vec{v}_i) + \dots + x_n(\vec{v}_n \cdot \vec{v}_i) \\ &= 0 + \dots + 0 + x_i\|\vec{v}_i\|^2 + 0 + \dots + 0 \\ &= x_i \end{aligned}$$

And thus

$$\begin{aligned} \text{proj}_{\mathbb{S}}\vec{x} &= (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \dots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k \\ &= x_1\vec{v}_1 + \dots + x_k\vec{v}_k \\ &= \vec{s} \text{ (our minimum distance vector calculated above)} \end{aligned}$$