Let $\vec{a} = (a_1, \ldots, a_n)^T$ be a vector. This vector defines a line, through the origin, in the direction of a. Let $\vec{b} = (b_1, \ldots, b_n)^T$ be any vector.

The projection of $\vec{b}$ onto $\vec{a}$, we want the point p on the line closest to $\vec{b}$. The key to doing this is to use orthogonality.

What is a projection?

The projection of $\vec{b}$ onto $\vec{a}$, $\text{proj}_{\vec{a}}\,\vec{b}$, is some multiple of $\vec{a}$, call it $\vec{p} = \hat{x}\vec{a}$. We need to find the number $\hat{x}$.

Note that

$$\vec{b} - \vec{p} = \vec{b} - \hat{x}\vec{a} = \vec{e}.$$

Since the error is perpendicula to $\vec{a}$,

$$\vec{a} \cdot (\vec{b} - \hat{x}\vec{a}) = \vec{a} \cdot \vec{b} - \hat{x}\left\|a\right\|^2 = 0.$$

This gives us that

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \hat{x}\left\|a\right\|^2 \\ \frac{\vec{a} \cdot \vec{b}}{\left\|\vec{a}\right\|^2} &= \hat{x} \end{aligned}$$

We now have that

$$\text{proj}_{\vec{a}}\,\vec{b} = \hat{x}\vec{a} = \frac{\vec{a} \cdot \vec{b}}{\left\|a\right\|^2}\,\vec{a}.$$

What is a Projection Matrix

We've noted that the project is

$$\vec{p} = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|^2}\,\vec{a}.$$

We need to understand what matrix P yields

$$P\vec{b} = \vec{p}.$$

What's multiplying $\vec{b}$. It's

$$\frac{aa^T}{\|a\|^2},$$

which defines the matrix P.

Projection onto a Subspace

We begin with k linearly independent vectors in $\mathbb{R}^n$, $\{a_1, \dots, a_k\}$.

Q: Find

$$p = \hat{x_1} a_1 + \hat{x_2} a_2 + \cdots + \hat{x_k} a_k$$

closest to a given vector b.

To do this, we calculate a bunch of 1-dimensional projections simultaneously:

$$a_1 \cdot (b - A\hat{x}) = 0$$
$$\vdots$$
$$a_n \cdot (b - A\hat{x}) = 0$$

which one can put into matrix form

$$p = \hat{x_1} a_1 + \cdots + \hat{x_n} a_n = A\hat{x}$$

via $A^T A\hat{x} = A^T b$, which has a solution$p = A\hat{x} = A(A^T A)^{-1} A^T b$.

1 Example

Consider the collection of vectors $\mathscr{S} = \{(1,-1,0)^T, (1,0,1)^T\}$. Project the vector $\vec{b} = (2,3,-4)^T$, onto the plane spanned by the vectors in $\mathscr{S}$. We want to find $\vec{p}$ such that

$$\vec{p} = \hat{x_1}\begin{pmatrix}1\\-1\\0\end{pmatrix} + \hat{x_2}\begin{pmatrix}1\\0\\1\end{pmatrix},$$

where $\vec{p} = \text{proj}_{\vec{p}}\vec{b}$.

Take

$$A = \begin{pmatrix}1 & 1\\-1 & 0\\0 & 1\end{pmatrix}$$

We proceed by calculating the left and right-hand sides of the normal equations:

$$A^TA = \begin{pmatrix}1 & -1 & 0\\1 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\-1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}$$

and

$$A^T\vec{b} = \begin{pmatrix}1 & -1 & 0\\1 & 0 & 1\end{pmatrix}\begin{pmatrix}2\\3\\-4\end{pmatrix} = \begin{pmatrix}-1\\-2\end{pmatrix}$$

We then want $\hat{x}$ such that

$$\begin{pmatrix}-1 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}\hat{x_1}\\\hat{x_2}\end{pmatrix} = \begin{pmatrix}-1\\-2\end{pmatrix}$$

which we obtain via row reduction

$$\left(\begin{array}{cc|c}1 & 2 & -2\\-1 & 1 & -1\end{array}\right) \xrightarrow{r_2+r_1} \begin{pmatrix}1 & 2 & -2\\0 & 3 & -3\end{pmatrix}$$

which yields $\hat{x_2} = -1$ and, via back-substitution, $\hat{x_1} = 0$.