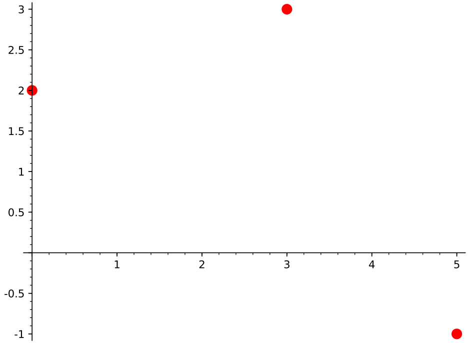
Orthogonal Projections

Dr. C. F. Rocca Jr

Western Connecticut State University

Suppose we are given a list of points and asked to find the *lesat squares regression line* for the points:



$$[(0,2),(3,3),(5,-1)]$$

We begin by setting up the problem as a matrix equation:

$$\begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} \beta \\ \mu \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

If this were consistent then all the points are on one line and we could find its equation, if not we need to do more work.

The first naive solution is to try and project the $\vec{y}$ onto each of the basis vectors in the column space and add them:

$$\vec{y} = (2,\, 3,\, -1)\,,\ \vec{x} = (0,\, 3,\, 5)\,,\ \vec{c} = (1,\, 1,\, 1)$$

so the projections are

$$proj_{\vec{x}}\vec{y} = \left(0,\, \frac{6}{17},\, \frac{10}{17}\right)\,,\ proj_{\vec{c}}\vec{y} = \left(\frac{4}{3},\, \frac{4}{3},\, \frac{4}{3}\right)$$

and we get the projection

$$\widehat{y} = proj_{\vec{x}}\vec{y} + proj_{\vec{c}}\vec{y} = \left(\frac{4}{3},\, \frac{86}{51},\, \frac{98}{51}\right)$$

Then we can solve the system:

$$\begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} \beta \\ \mu \end{pmatrix} = \begin{pmatrix} \frac{4}{3} \\ \frac{86}{51} \\ \frac{98}{51} \end{pmatrix}$$

Which gives:

$$(\beta, \mu) = \left(\frac{4}{3}, \frac{2}{17}\right)$$

Then we can solve the system:

$$\begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 5 \end{pmatrix} \cdot \begin{pmatrix} \beta \\ \mu \end{pmatrix} = \begin{pmatrix} \frac{4}{3} \\ \frac{86}{51} \\ \frac{98}{51} \end{pmatrix}$$

Which gives:

$$(\beta, \mu) = \left(\frac{4}{3}, \frac{2}{17}\right)$$

But it turns out this would be wrong.

What we need to do is first *orthogonalize* the columns of A then do as we did above. So replace $\vec{x}$ with

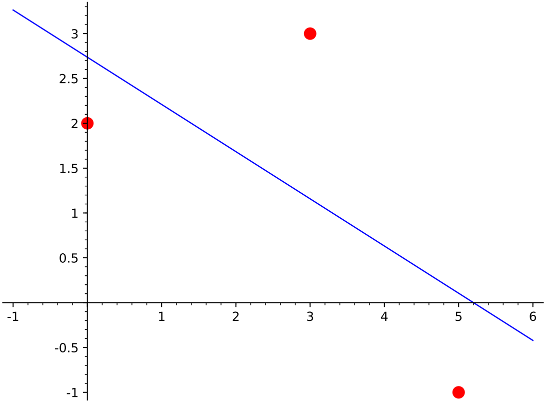
$$\widehat{x} = \vec{x} - proj_{\vec{c}}\vec{x} = \left(-\frac{8}{3}, \frac{1}{3}, \frac{7}{3}\right)$$

and then proceed as before with

$$\widehat{y} = proj_{\widehat{x}}\vec{y} + proj_{\vec{c}}\vec{y} = \left(\frac{52}{19}, \frac{22}{19}, \frac{2}{19}\right)$$

and

$$(\beta, \mu) = \left(\frac{52}{19}, -\frac{10}{19}\right)$$



3
2.5
2
1.5
1
0.5
-0.5
-1
-1
1
2
3
4
5
6

But of course there is a clever/tricky way to solve this...

Since we are looking to the **orthogonal** projection of $\vec{y}$ onto the column space of A, call it $\widehat{y}$, we get

$$A^T(\vec{y} - \widehat{y}) = \vec{0}$$

which gives us

$$A^T\vec{y} = A^T\widehat{y}.$$

But the $\widehat{y}$ we are looking for should satisfy

$$\widehat{y} = A \cdot \begin{pmatrix} \beta \\ \mu \end{pmatrix}$$

so using the previous result we get

$$A^T \vec{y} = A^T \widehat{y} = A^T A \cdot \begin{pmatrix} \beta \\ \mu \end{pmatrix},$$

or

$$\begin{pmatrix} \beta \\ \mu \end{pmatrix} = (A^T A)^{-1} A^T \vec{y}.$$

Going back to our example this is:

$$\begin{pmatrix} \beta \\ \mu \end{pmatrix} = (A^T A)^{-1} A^T \vec{y} = \begin{pmatrix} \frac{52}{19} \\ -\frac{10}{19} \end{pmatrix}$$

Which, is exactly what we got before.