10/28/2021

Projections, Best Approximation Theorem, Regression

$$y = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

H is a subspace equal to $span\{u_1, u_2\}$. Project y onto the subspace H.

$$proj_H y = proj_{u_1} y + proj_{u_2} y = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 =$$

$$\frac{(-1)(1) + (4)1 + 3(0)}{1^2 + 1^2 + 0} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{(-1)(-1) + 4(1) + 3(0)}{(-1)^2 + 1^2 + 0} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} =$$

$$\frac{3}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{5}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} = y_{\parallel}$$

Orthogonal component (perpendicular distance from the point to the plane (subspace))

$$y_{\perp} = y - y_{\parallel} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} - \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

The Best Approximation Theorem

Let W be a subspace of R^n, and let y be any vector in R^n, and let $y_{\parallel}$ be the orthogonal project of y onto W. Then $y_{\parallel}$ is the closest point in W to y, in the sense that $\|y - y_{\parallel}\| < \|y - v\|$ where v is any other point in W.

Basis for R^2 is $u_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, u_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$. Write $x = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$ as a vector in this basis.

The coefficients for the basis vectors are the projection coefficients onto that basis vector.

$$c_1 \begin{bmatrix} 2 \\ -3 \end{bmatrix} + c_2 \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$$

$$c_1 = \frac{x \cdot u_1}{u_1 \cdot u_1}, c_2 = \frac{x \cdot u_2}{u_2 \cdot u_2}$$

$$c_1 = \frac{2(9) + (-3)(-7)}{2^2 + (-3)^2} = \frac{18 + 21}{4 + 9} = \frac{39}{13} = 3$$

$$c_2 = \frac{(9(6) + (-7)(4))}{6^2 + 4^2} = \frac{54 - 28}{36 + 16} = \frac{26}{52} = \frac{1}{2}$$

$$[x]_B = \begin{bmatrix} 3 \\ \frac{1}{2} \end{bmatrix}$$

$$P_B[x]_B = x$$

$$\begin{bmatrix} 2 & 6 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$$

$$P_B^{-1} = \frac{1}{8+18}\begin{bmatrix} 4 & -6 \\ 3 & 2 \end{bmatrix} = \frac{1}{26}\begin{bmatrix} 4 & -6 \\ 3 & 2 \end{bmatrix}$$

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{26}\begin{bmatrix} 4 & -6 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 9 \\ -7 \end{bmatrix} = \frac{1}{26}\begin{bmatrix} 36+42 \\ 27-14 \end{bmatrix} = \frac{1}{26}\begin{bmatrix} 78 \\ 13 \end{bmatrix} = \begin{bmatrix} 3 \\ \frac{1}{2} \end{bmatrix}$$

Least Squares Solutions

If A is an mxn matrix and b is in R^m, a least-squares solution of Ax=b is a vector $x_{\|}$ in R^n such that $\|b - Ax_{\|}\| < \|b - Ax\|$ for any x in R^n.

Typically, we use this theorem when the system is overdetermined: more equations than variables.

$$Ax = b$$

Multiplying both sides of the equation by the transpose serves to project A (and b) onto a subspace.

$$A^T A x = A^T b$$

$A^T A$ is now nxn. And $A^T b$ is nx1. If we solve for x in this equation, then we'll get the best approximation to b in the original system that we can get.

This equation is called the Normal Equation.

To solve this exactly, $A^T A$ must be invertible. For this to be true, the original columns of A must be independent.

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 2 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$Ax = b, or\ get\ the\ best\ approximation$$

Use the normal equation to estimate the best x to approximate b.

$$A^T A x = A^T b$$

$$\begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 2 \end{bmatrix} x = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

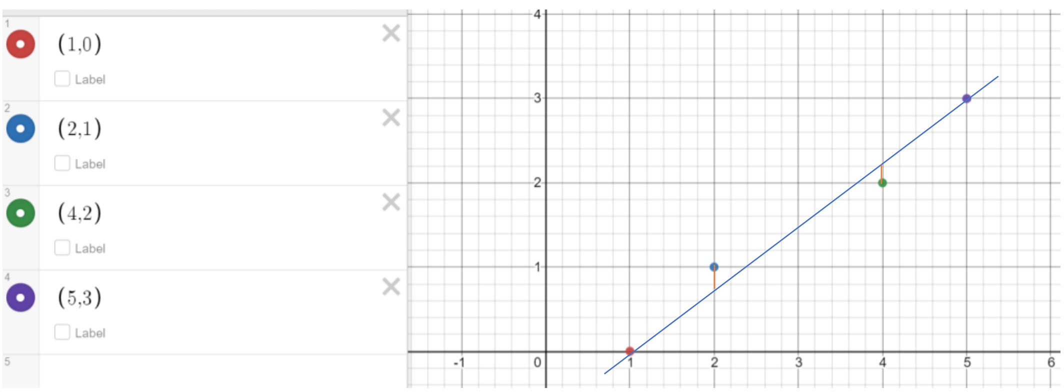
$$\begin{bmatrix} 17 & 2 \\ 2 & 8 \end{bmatrix} x = \begin{bmatrix} 19 \\ 22 \end{bmatrix}$$

$$x = \frac{1}{136 - 4} \begin{bmatrix} 8 & -2 \\ -2 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 22 \end{bmatrix} = \begin{bmatrix} \frac{9}{11} \\ \frac{28}{11} \end{bmatrix} \approx \begin{bmatrix} 0.818 \\ 2.545 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{9}{11} \\ \frac{28}{11} \end{bmatrix} = \begin{bmatrix} \frac{36}{11} \approx 3.27 \\ \frac{56}{11} \approx 5.09 \\ \frac{65}{11} \approx 5.91 \end{bmatrix} \approx \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Find the best fit (least-squares) regression line to the equation $\beta_0 + \beta_1 x = y$ for the points:

$$\{(1,0), (2,1), (4,2), (5,3)\}$$



$$\begin{aligned} \beta_0 + \beta_1 &= 0 \\ \beta_0 + 2\beta_1 &= 1 \\ \beta_0 + 4\beta_1 &= 2 \\ \beta_0 + 5\beta_1 &= 3 \end{aligned}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = (A^T A)^{-1} A^T y$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} -0.6 \\ 0.7 \end{bmatrix}$$

$$y = -0.6 + 0.7x$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$