Ira Nirenberg
Chicago Laboratory Schools
iranirenberg@yahoo.com

Linear Least Squares

Presented at:

New Orleans Math & Science Conference July 24 & 25 2006

Linear least Squares

Reasons for learning the derivation of the Linear Least Squares the algorithm.

a) The derivation gives students a better "feel" for what the calculator actually does and therefore provides some measure of mathematical empowerment.

b) The derivation can be understood using algebra 2 concepts along with a 3 dimensional coordinate system, and shows real world utility to mathematics. Calculus can also be used, but is not necessary.

c) Concepts the derivation requires:

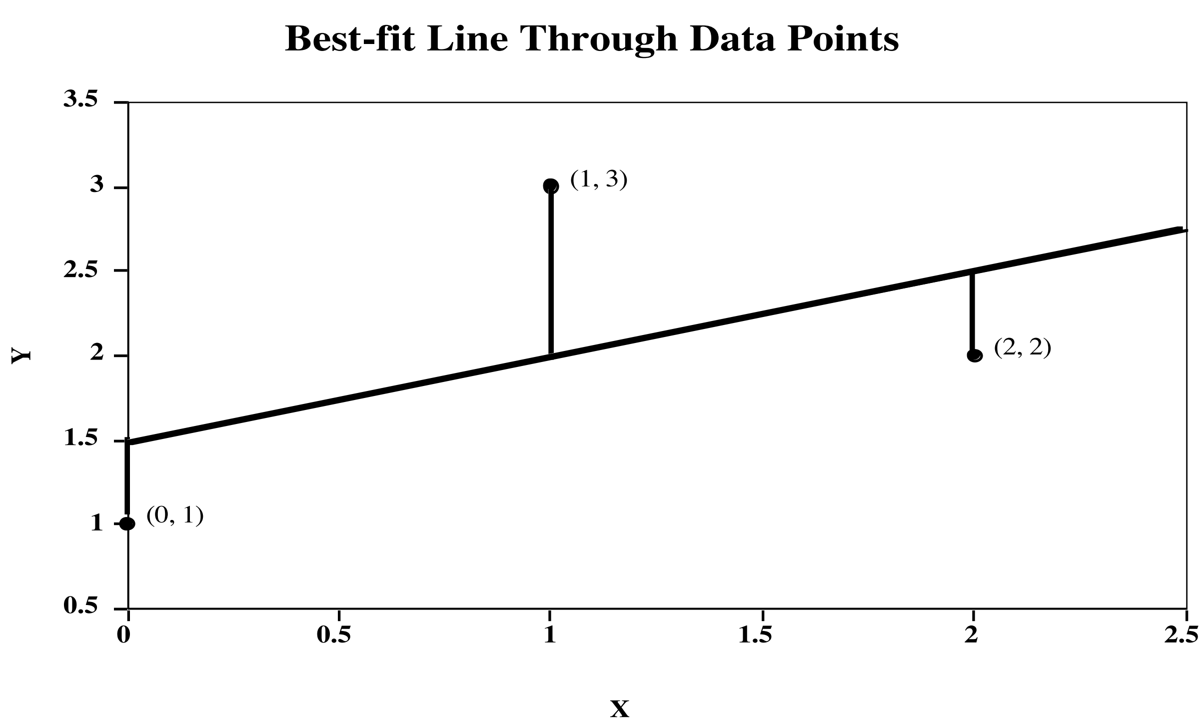
1. Understanding residuals
2. Multiplying trinomials.
3. Finding the vertex of a parabola.
4. Solving linear simultaneous equations
5. Using conic sections
6. Using summation notation (optional in specific cases)
7. Using partial derivatives (optional)

Goal: To find a "best-fit" linear equation for a scatter plot.

Method: Minimize the square of the residuals between the best-fit line and the data.

A Concrete Example:

Find the best-fit line for the three points, (0,1), (1,3), (2,2).



Residual: A residual is the difference (+/-) between the y-coordinates of the best-fit line and the y-coordinates of the data points. That is, $Y_{best\text{-}fit} - Y_{data}$, or,

Res = $(mx + b - Y_{data})$.

To minimize the square of the residulas we have:

$Res^2 = (mx + b - Y_{data})^2$

For the three points given above we need to minimize the quantity:

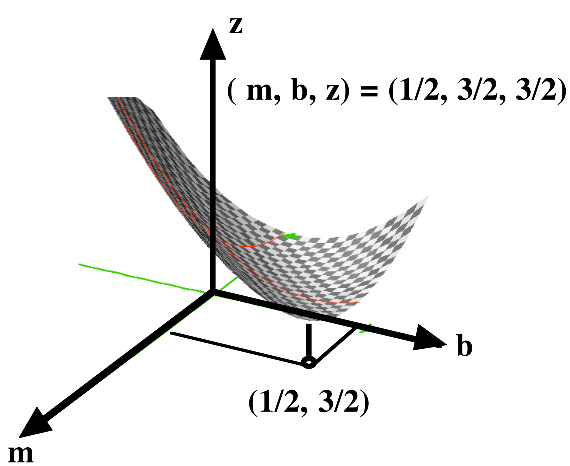
Sum of Res squared (SS) = $[(mx_1 + b - 1)^2 + (mx_2 + b - 3)^2 + (mx_3 + b - 2)^2]$

Substituting, squaring, and combining terms yields:

$SS = 5m^2 + 3b^2 + 6mb - 14m - 12b + 14$

The above equation has 2 independent variables, m and b, and 1 dependent variable, SS, which we will call z. Therefore, z = f(m, b).

z = f(m, b), gives a surface in 3 dimensional space.



Since the vertex of the surface has the minimum z value, the pair of coordinates (m,b) give the slope and intercept for the best-fit line.

How do we find the vertex?

Imagine the projection of this surface into the b-z plane and the m-z plane.

Projection onto m-z plane:

Think of the surface as a solid and begin slicing it parallel to the m-z plane until the lowest point (the vertext) emerges. This will occur for some fixed value of b. The vertex will therefore be a function of m only since the curve is defined in a plane parallel to the m-z plane. But the vertex of what? A little thought (and the equation) should show it is a parabola. Remember b is considered a constant.

Projection onto the b-z plane:

Here we take parallel slices to the b-z plane until the lowest point emerges. This occurs for some fixed value of m. The same thinking as before will produce a second parabola in the b-z plane, with the vertex a function of b only.

Two equations have now been generated, both involving the variables m and b.

1. SS = $5m^2$ + m(6b - 14) + c = 0, where c equals the constant ($3b^2$ - 12b + 14).

2. SS = $3b^2$ + b(6m - 12) + c' = 0, where c' equals the constant ($5m^2$ -14m + 14).

Recalling the equation $Ax^2 + Bx + c = 0$, we have V_m = -B/2A and V_b = -B'/2A' where prime notation designates different constants of the quadratic. Thus,

5m + 3b = 7 from equation 1

m + b = 2 from equation 2

Solving these simultaneously yields:

m = 1/2 , b = 3/2.