Symbolic derivation of intercept and slope.
Instructor: Nam Sun Wang

Given x and y, we wish to find the intercept (a_0) and slope (a_1) such that the line fits though the given points as closely as possible.

$$y_i = a_0 + a_1 \cdot x_i + error_i$$

We wish to find a_0 and a_1 that minimize the sum of squared error (least squares regression)

$$sse = \sum_i error^2 = \sum_i \left(y_i - a_0 - a_1 \cdot x_i\right)^2$$

When sse is minimized, derivatives of sse wrt a_0 and a_1 are 0.

$$\frac{d}{d a_0} sse = 0 \longrightarrow 0 = \sum_i \left(y_i - a_0 - a_1 \cdot x_i\right)$$

$$\frac{d}{d a_1} sse = 0 \longrightarrow 0 = \sum_i \left(y_i - a_0 - a_1 \cdot x_i\right) \cdot x_i$$

Thus, we have two equations to solve for the two unknowns (a_0 & a_1). Let Mathcad crank.

Given $0 = \sum y - n \cdot a_0 - a_1 \cdot \sum x$

$$0 = \sum (x \cdot y) - a_0 \cdot \sum x - a_1 \cdot \sum \left(x^2\right)$$

$$\text{Find}\left(a_0, a_1\right) \rightarrow \begin{bmatrix} \dfrac{\left[-\sum(x) \cdot \sum(x \cdot y) + \sum(y) \cdot \sum\left(x^2\right)\right]}{\left[-\sum(x)^2 + \sum\left(x^2\right) \cdot n\right]} \\ \dfrac{1}{\left[-\sum(x)^2 + \sum\left(x^2\right) \cdot n\right]} \cdot \left(\sum(x \cdot y) \cdot n - \sum(y) \cdot \sum(x)\right) \end{bmatrix}$$

Define my intercept and slope functions in terms of x and y.

intercept: $$\text{Intercept}(x, y) := \frac{\sum(y) \cdot \sum\left(\overrightarrow{(x \cdot x)}\right) - \sum(x) \cdot \sum\left(\overrightarrow{(x \cdot y)}\right)}{\sum\left(\overrightarrow{(x \cdot x)}\right) \cdot (\text{last}(x) + 1) - \left(\sum(x)\right)^2}$$

slope: $$\text{Slope}(x, y) := \frac{\sum\left(\overrightarrow{(x \cdot y)}\right) \cdot (\text{last}(x) + 1) - \sum(y) \cdot \sum(x)}{\sum\left(\overrightarrow{(x \cdot x)}\right) \cdot (\text{last}(x) + 1) - \left(\sum(x)\right)^2}$$

Linear formulation.

$$\frac{d}{d a_0} sse = 0 \longrightarrow 0 = \sum y - n \cdot a_0 - a_1 \cdot \sum x$$

$$\frac{d}{d a_1} sse = 0 \longrightarrow 0 = \sum (x \cdot y) - a_0 \cdot \sum x - a_1 \cdot \sum (x \cdot x)$$

Express the above equations in a linear format.

$$\begin{bmatrix} n & \sum x \\ \sum x & \sum\left(x^2\right) \end{bmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{bmatrix} \sum y \\ \sum (x \cdot y) \end{bmatrix} \longrightarrow \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{bmatrix} n & \sum x \\ \sum x & \sum\left(x^2\right) \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum y \\ \sum (x \cdot y) \end{bmatrix}$$

If we add another term, say a quadratic term, we extend the same methodology.

$$y_i = a_0 + a_1 \cdot x_i + a_2 \cdot (x_i)^2 + error_i$$

We wish to find a_0 and a_1 that minimize the sum of squared error (least squares regression)

$$sse = \sum_i error^2 = \sum_i \left[y_i - a_0 - a_1 \cdot x_i - a_2 \cdot (x_i)^2 \right]^2$$

When sse is minimized, derivatives of sse wrt a_0 and a_1 are 0.

$$\frac{d}{d a_0} sse = 0 \longrightarrow 0 = \sum_i \left[y_i - a_0 - a_1 \cdot x_i - a_2 \cdot (x_i)^2 \right]$$

$$\frac{d}{d a_1} sse = 0 \longrightarrow 0 = \sum_i \left[y_i - a_0 - a_1 \cdot x_i - a_2 \cdot (x_i)^2 \right] \cdot x_i$$

$$\frac{d}{d a_2} sse = 0 \longrightarrow 0 = \sum_i \left[y_i - a_0 - a_1 \cdot x_i - a_2 \cdot (x_i)^2 \right] \cdot (x_i)^2$$

Solve the above coupled set of three equations for a_0, a_1, and a_2.

$$0 = \sum y - n \cdot a_0 - a_1 \cdot \sum x - a_2 \cdot \sum (x^2)$$

$$0 = \sum (x \cdot y) - a_0 \cdot \sum x - a_1 \cdot \sum (x^2) - a_2 \cdot \sum (x^3)$$

$$0 = \sum (x^2 \cdot y) - a_0 \cdot \sum (x^2) - a_1 \cdot \sum (x^3) - a_2 \cdot \sum (x^4)$$

We can also express these equations in a linear format.

$$\begin{bmatrix} n & \sum x & \sum (x^2) \\ \sum x & \sum (x^2) & \sum (x^3) \\ \sum (x^2) & \sum (x^3) & \sum (x^4) \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum (x \cdot y) \\ \sum (x^2 \cdot y) \end{bmatrix} \longrightarrow \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} n & \sum x & \sum (x^2) \\ \sum x & \sum (x^2) & \sum (x^3) \\ \sum (x^2) & \sum (x^3) & \sum (x^4) \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum y \\ \sum (x \cdot y) \\ \sum (x^2 \cdot y) \end{bmatrix}$$

Example:

$data := READPRN(random2_{dat})$ $\quad x := data^{\langle 0 \rangle}$ $\quad y := data^{\langle 1 \rangle}$ $\quad i := 0 .. last(x)$

$a_0 := Intercept(x, y)$ $\quad a_0 = 2.108$ ←compare to Mathcad→ $intercept(x, y) = 2.108$

$a_1 := Slope(x, y)$ $\quad a_1 = 0.198$ ←compare to Mathcad→ $slope(x, y) = 0.198$