Can We Run a Fitted Line Backward?

Due: Wednesday, September 23 at 11:59pm

Instructions: Please submit your *handwritten* solutions as a single PDF on Gradescope.

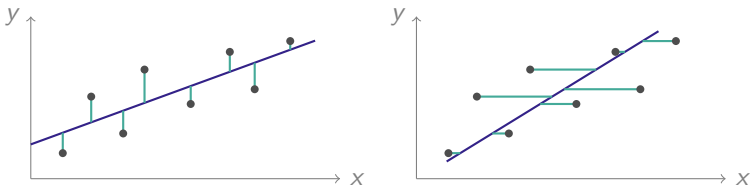
Two engineers use the same calibration measurements (x_i, y_i), $i = 1, \ldots, n$, to predict x from y. Engineer A first fits $y = wx + b$ by minimizing

$$\frac{1}{n}\sum_{i=1}^{n}\big(y_i - wx_i - b\big)^2,$$

then rearranges the fitted line to predict $x = (y - b)/w$. Engineer B directly fits $x = ay + c$ by minimizing

$$\frac{1}{n}\sum_{i=1}^{n}\big(x_i - ay_i - c\big)^2.$$

Assume the x values are not all equal, the y values are not all equal, and both fitted slopes are nonzero.



Fit y from x Fit x from y

1. Engineer A fits y from x, then rearranges the line. Find this engineer's resulting prediction rule for x as a function of y.
2. Engineer B fits x from y directly. Find this engineer's resulting prediction rule for x as a function of y.
3. Both engineers used the same measurements. Prove a necessary and sufficient geometric condition for their rules to agree for every input y.

Hint: The reading's formula for the least-squares slope and intercept.

Write $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. Define the centered vectors $\mathbf{x}_{\text{centered}}, \mathbf{y}_{\text{centered}} \in \mathbb{R}^n$ by

$$(\mathbf{x}_{\text{centered}})_i = x_i - \bar{x}, \qquad (\mathbf{y}_{\text{centered}})_i = y_i - \bar{y}.$$

For nonzero vectors $\mathbf{x}, \mathbf{y}$, you may use

$$\langle \mathbf{x}, \mathbf{y} \rangle = \cos(\theta)\|\mathbf{x}\|_2\|\mathbf{y}\|_2,$$

where θ is the angle between the vectors.

Takeaway: *A line fitted to predict y from x does not usually give the best fit when used backward. Fitting in the other direction minimizes a different error. The two rules agree only when all measurements lie on one line.*