EECS 245, Winter 2026

LEC 4 Simple Linear Regression

→ Read: All of Ch. 2

Agenda

Read Ch. 2!

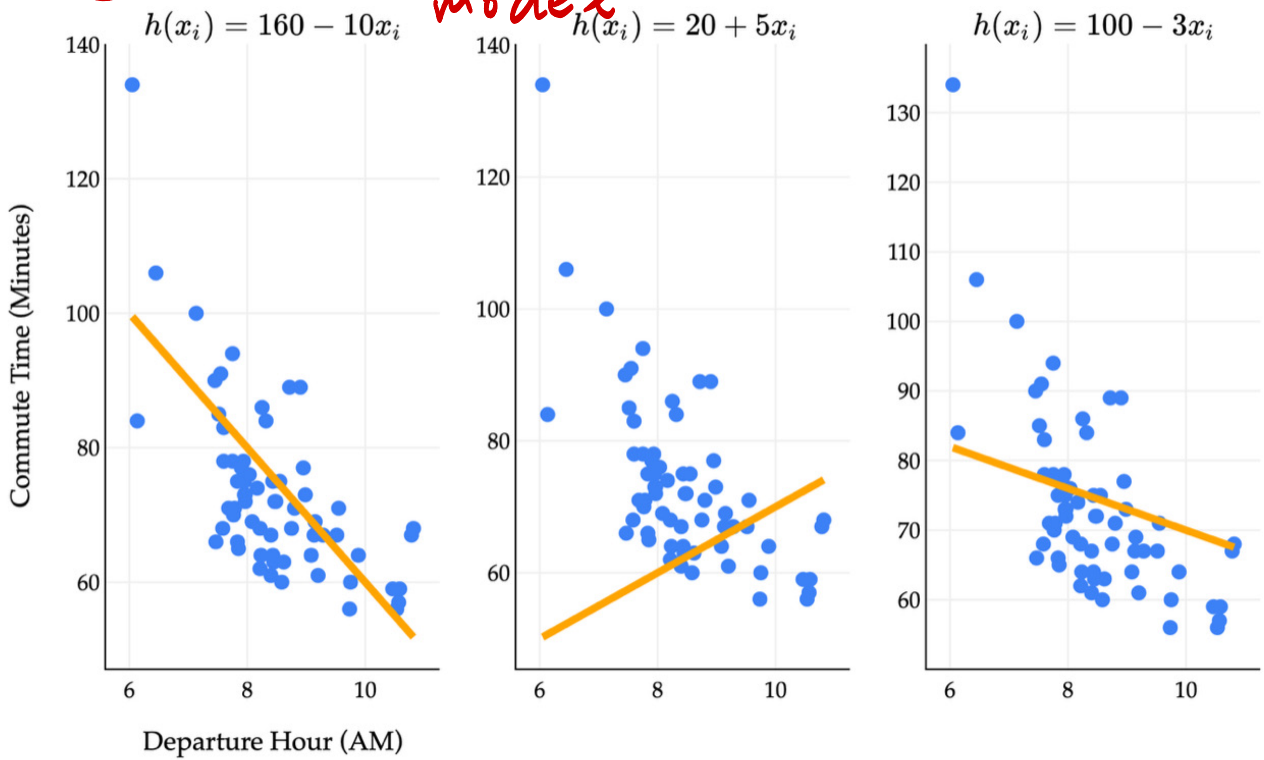
lots of smaller sections

- Modeling recipe for simple linear regression] 2.1
- Partial derivatives] 2.2
- Deriving and using the optimal parameters] 2.3
- Correlation] 2.4

Announcements

- HW 2 due Friday
- HW 1 scores and solutions available: read and reflect!
- Lab solutions posted on website

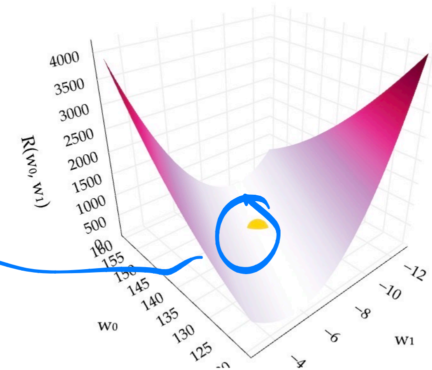
$$\underbrace{h(x_i) = w_0 + w_1 x_i}_{\text{model}}$$



$$R_{sq}(w_0, w_1) = \overbrace{\frac{1}{n}\sum_{i=1}^{n} \underbrace{\left(y_i - (w_0 + w_1 x_i)\right)^2}_{(\text{actual} - \text{pred})^2}}^{\text{mean squared error}}$$

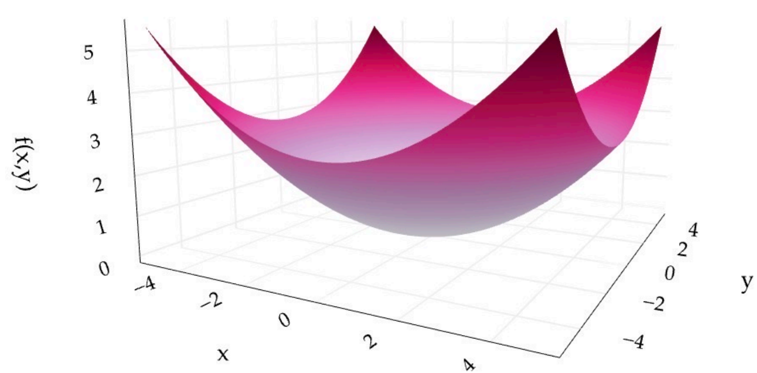
two input variables

We want to find the w_0, w_1 that minimize $R_{sq}(w_0, w_1)$

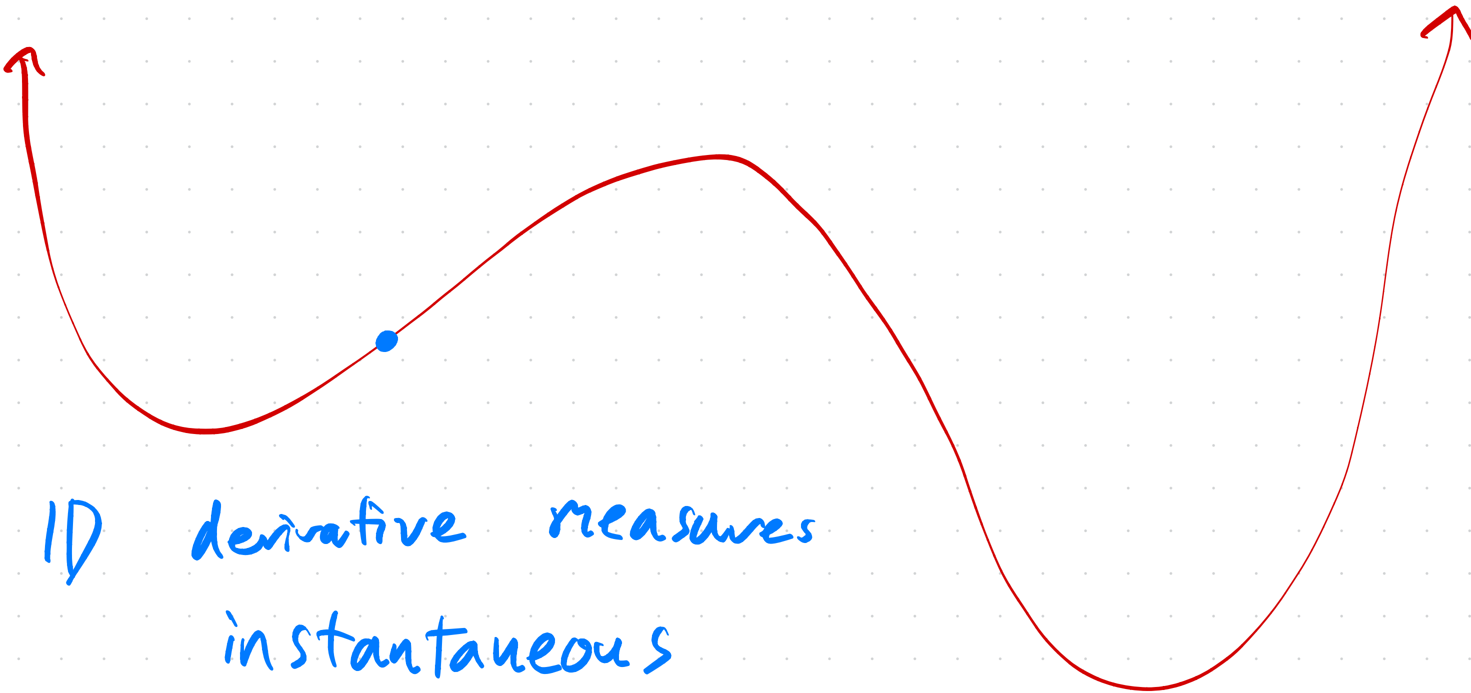


Aside: functions with multiple input variables

Ex: $$f(x,y) = \frac{x^2+y^2}{9}$$



"paraboloid"



1D derivative measures instantaneous

rate of change

"Partial derivative" of f with respect to x is computed by treating all other variables (y) as constants functions

$$\boxed{\frac{\partial f}{\partial x}}$$

curly d

$$f(x,y) = \frac{x^2+y^2}{9}$$

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x}\left[\frac{x^2+y^2}{9}\right]$$

$$= \frac{\partial}{\partial x}\left[\frac{x^2}{9} + \frac{y^2}{9}\right]$$

$$= \frac{2x}{9}$$

$$\frac{\partial f}{\partial y}(x,y) = 2y/9$$

$$\frac{\partial f}{\partial x}(x,y) = \frac{2x}{9} \qquad \frac{\partial f}{\partial y}(x,y) = \frac{2y}{9}$$

e.x. suppose $x = -3$, $y = 1/2$

$$\frac{\partial f}{\partial x}(-3, 1/2) = \frac{2(-3)}{9} = \frac{-6}{9} = -\frac{2}{3}$$

$$\frac{\partial f}{\partial y}(-3, 1/2) = 1/9$$

look at 2.2 diagram

another example

$$g(x,y) = x^3 \underbrace{- 3xy^2}_{=(-3y^2)x} + 2\sin(x)\cos(y)$$

$$\frac{\partial g}{\partial x}(x,y) = \underbrace{3x^2 - 3y^2 + 2\cos(x)\cos(y)}_{\text{function of BOTH } x \text{ and } y!}$$

$$\frac{\partial g}{\partial y}(x,y) = -6xy - 2\sin(x)\sin(y)$$

What's the point?

partial derivatives help us minimize/maximize functions!

→ idea: find the points where ALL partial derivatives are 0, solve resulting system back to the main point

$$R_{sq}(w_0, w_1) = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - (w_0 + w_1 x_i) \right)^2$$

strategy:

- compute $\frac{\partial R}{\partial w_0}$ and $\frac{\partial R}{\partial w_1}$
- set both $= 0$
- solve the resulting system of equations

chain rule!

$$R_{sq}(w_0, w_1) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - (w_0 + w_1 x_i)\right)^2$$

$$\frac{\partial R}{\partial w_0} = \frac{1}{n}\sum_{i=1}^{n} 2\left(y_i - (w_0 + w_1 x_i)\right) \underbrace{\frac{\partial}{\partial w_0}\left(y_i - (w_0 + w_1 x_i)\right)}_{= -1}$$

$$= -\frac{2}{n}\sum_{i=1}^{n}\left(y_i - (w_0 + w_1 x_i)\right)$$

$$\frac{\partial R}{\partial w_1} = -\frac{2}{n}\sum_{i=1}^{n} x_i\left(y_i - (w_0 + w_1 x_i)\right)$$

① $$\frac{\partial R}{\partial w_0} = -\frac{2}{n}\sum_{i=1}^{n}\left(y_i - (w_0 + w_1 x_i)\right) = 0$$

② $$\frac{\partial R}{\partial w_1} = -\frac{2}{n}\sum_{i=1}^{n} x_i\left(y_i - (w_0 + w_1 x_i)\right) = 0$$

now, set both $=0$ and solve the resulting system of 2 eq'ns, 2 unknowns

plan: in eq ①, isolate for $w_0^* = \underline{\qquad\qquad}$ ← optimal intercept

then, substitute into eq ② and solve for w_1^*.

look at chapter 2.3 for the details

but what you end up with is

optimal intercept

$$w_0^* = \bar{y} - w_1^* \bar{x}$$

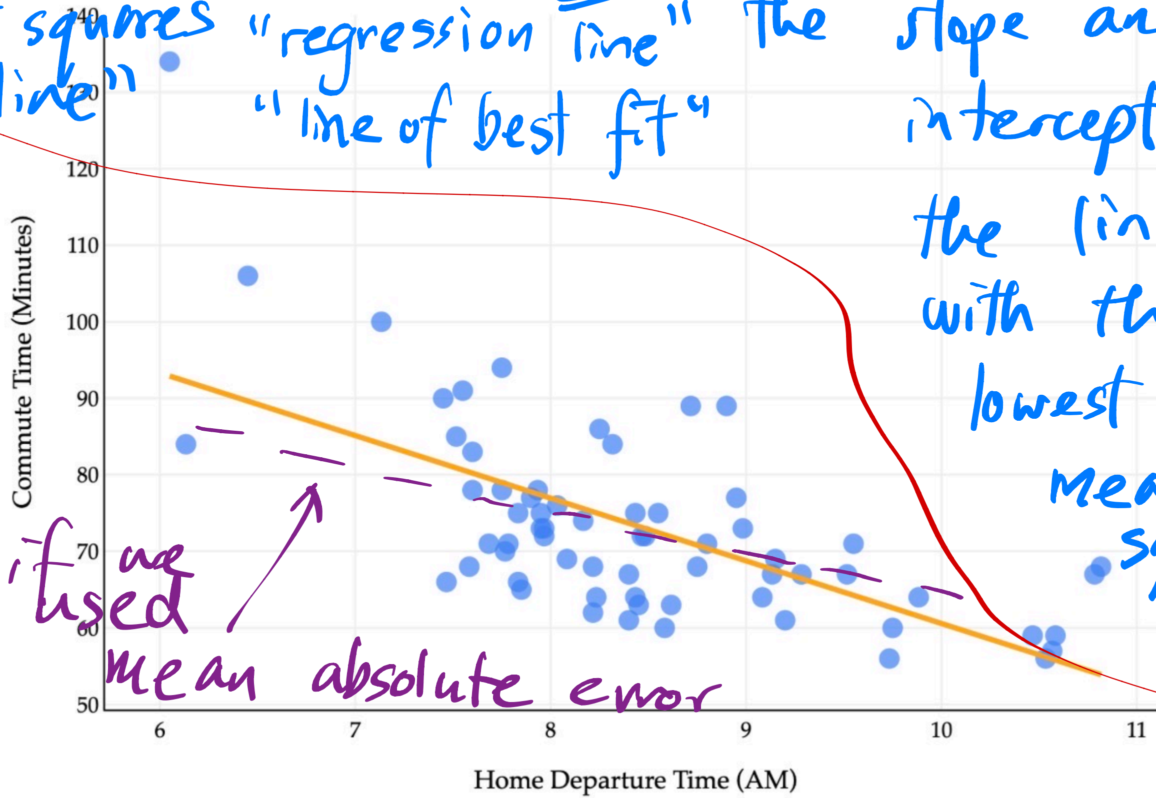
$$w_1^* = \frac{\sum_{i=1}^{n} x_i (y_i - \bar{y})}{\sum_{i=1}^{n} x_i (x_i - \bar{x})}$$

many equivalent formulas for w_1^*!

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

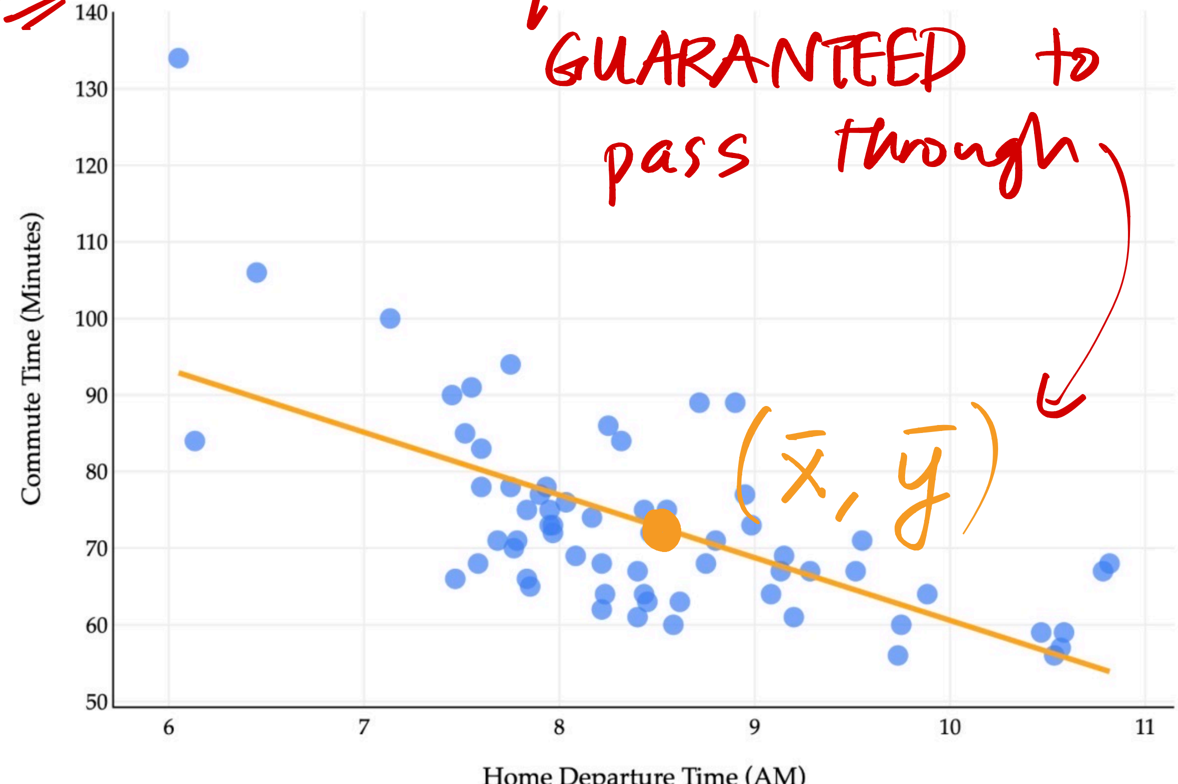
The point of it all was to find the slope and intercept of the line with the lowest possible mean squared error!

"least squares line" "regression line" "line of best fit"



if we used mean absolute error

Fact: The line that minimizes mean squared error is GUARANTEED to pass through



How to prove that fact?

optimal intercept

$$h(x_i) = w_0^* + w_1^* x_i$$

plug in $w_0^* = \bar{y} - w_1^* \bar{x}$

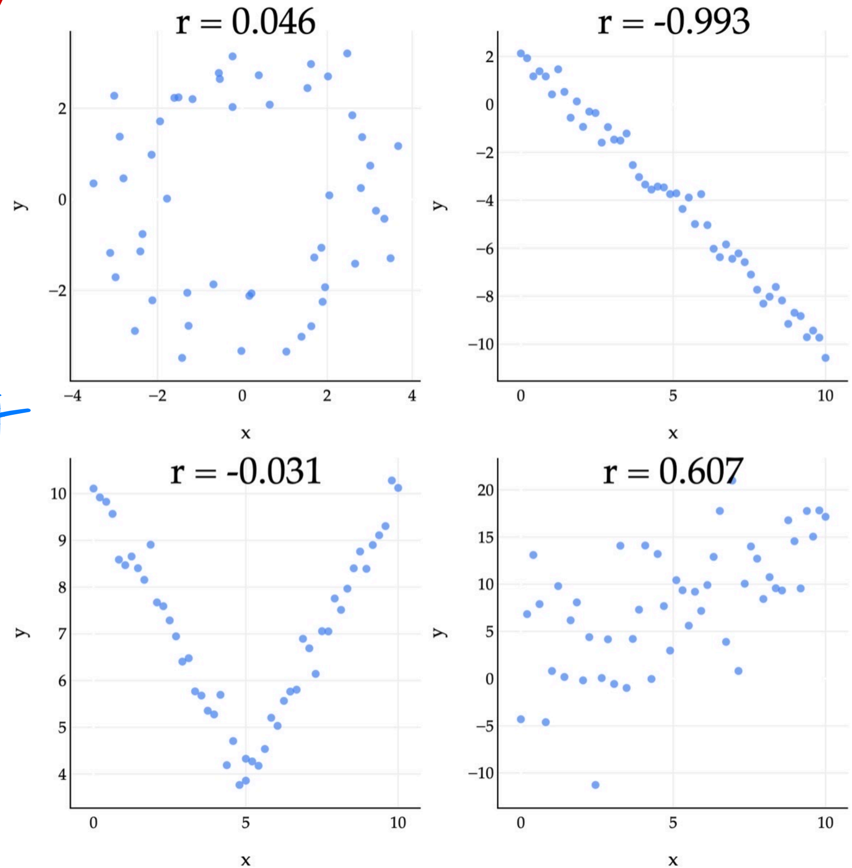
$$h(x_i) = \bar{y} - w_1^* \bar{x} + w_1^* x_i$$

plug in $x_i = \bar{x}$

$$h(\bar{x}) = \bar{y} \cancel{- w_1^* \bar{x}} \cancel{+ w_1^* \bar{x}} = \bar{y}$$

Correlation coefficient, r, $-1 \le r \le 1$

"how much does the scatter plot look like a straight line?"

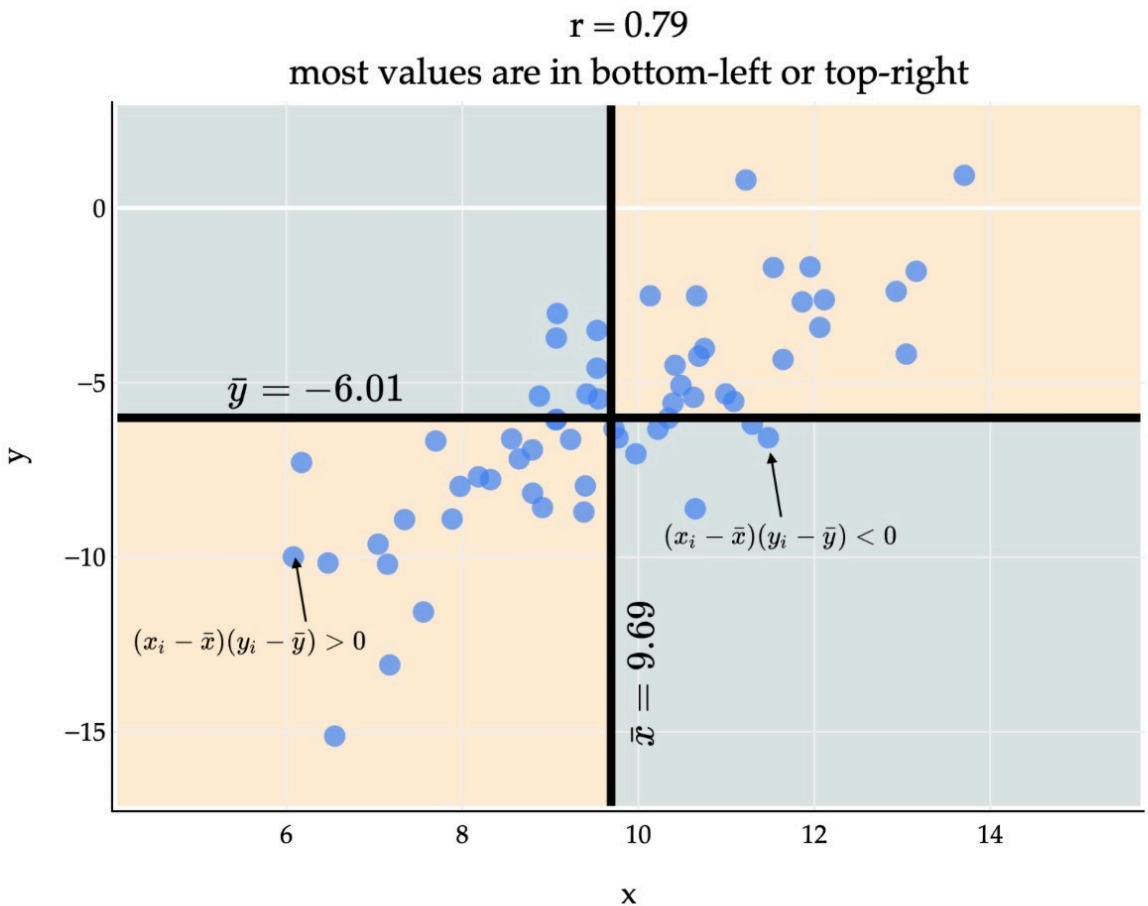


$$w_1^* = r\frac{\sigma_y}{\sigma_x}$$

optimal slope

std deviations of y and x

$$r = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{\text{average}} \underbrace{\left(\frac{x_i - \bar{x}}{\sigma_x}\right)}_{\text{\# of SDs above mean}} \left(\frac{y_i - \bar{y}}{\sigma_y}\right)$$


r = 0.79
most values are in bottom-left or top-right
$\bar{y} = -6.01$
$\bar{x} = 9.69$
$(x_i - \bar{x})(y_i - \bar{y}) > 0$
$(x_i - \bar{x})(y_i - \bar{y}) < 0$
0
−5
−10
−15
6
8
10
12
14
x
y