EECS 245 Fall 2025
Math for ML

EECS 245 Fall 2025

Math for ML

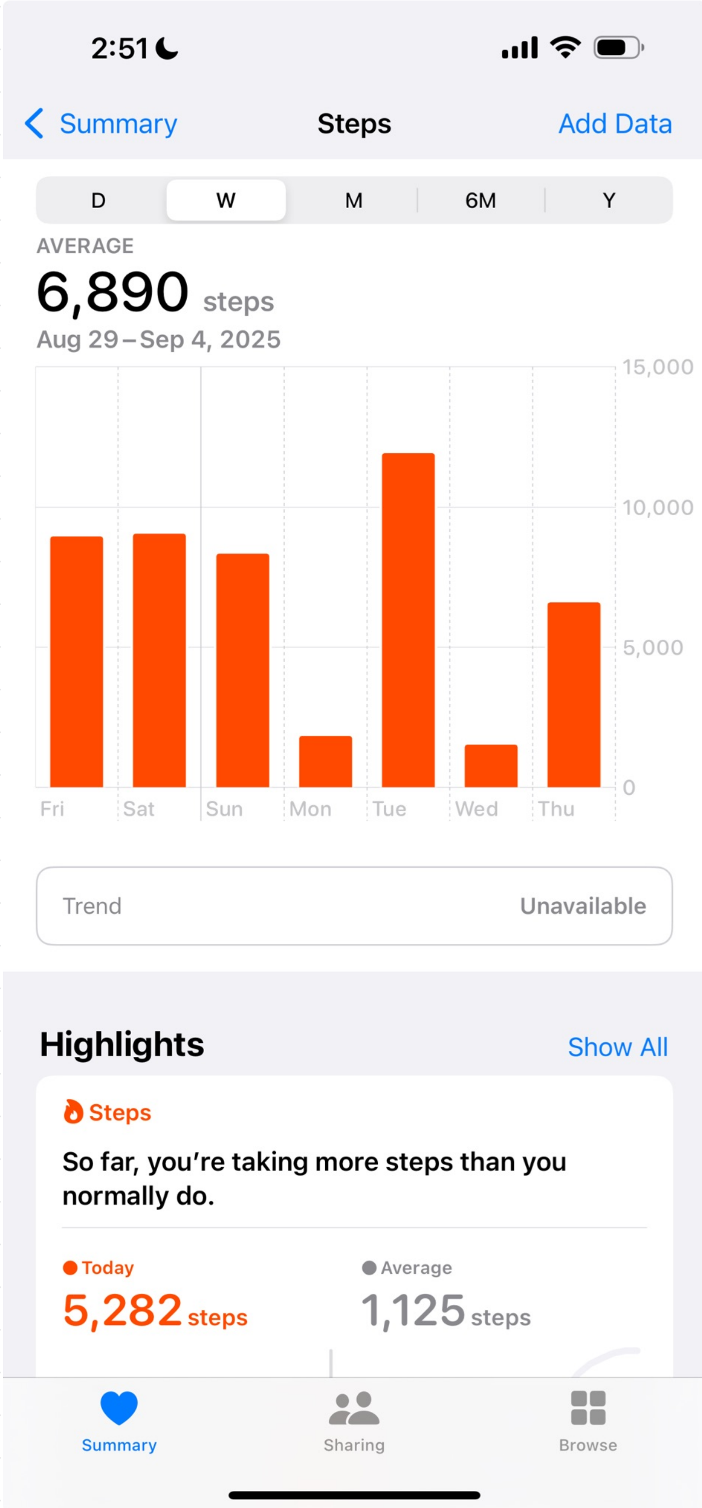
Lecture 4: Simple Linear Regression

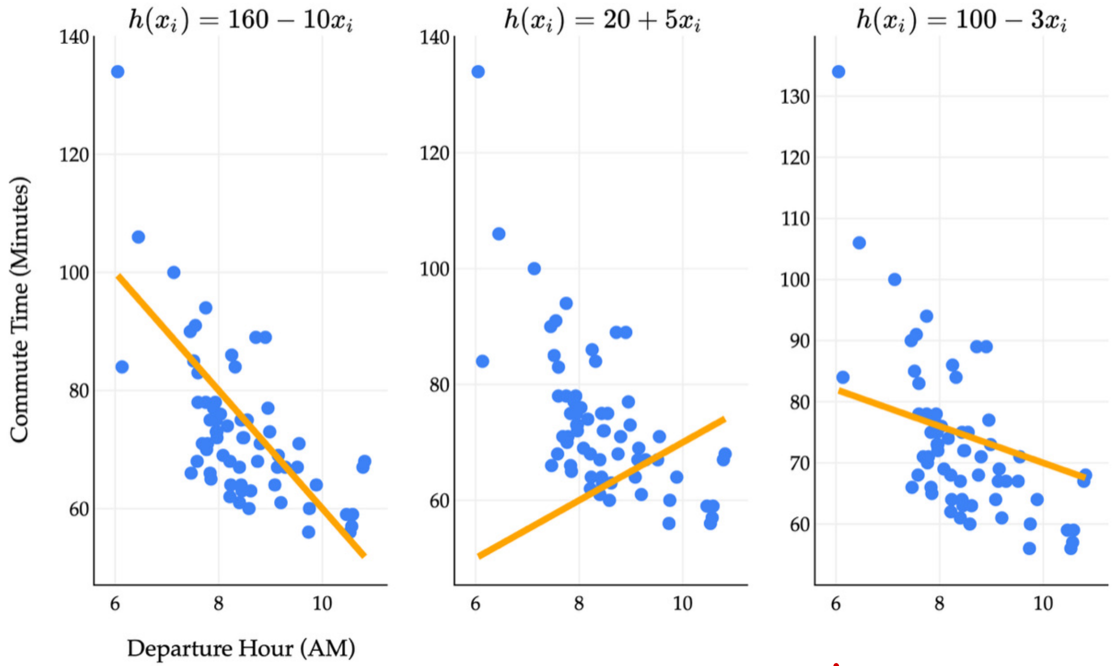
→ Read Ch. 1.4 (new!)

Agenda

1. Big ideas from Lab 2
2. Modeling recipe for simple linear regression
3. Partial derivatives
4. Deriving the optimal parameters
5. Using the optimal parameters
6. Correlation

all in Ch. 1.4, definitely read


2:51
Summary
Steps
Add Data
D
W
M
6M
Y
AVERAGE
6,890 steps
Aug 29–Sep 4, 2025
15,000
10,000
5,000
0
Fri
Sat
Sun
Mon
Tue
Wed
Thu
Trend
Unavailable
Highlights
Show All
Steps
So far, you're taking more steps than you normally do.
Today
5,282 steps
Average
1,125 steps
Summary
Sharing
Browse



Recipe

1) $h(x_i) = w_0 + w_1 x_i$ "simple linear reg"

parameters

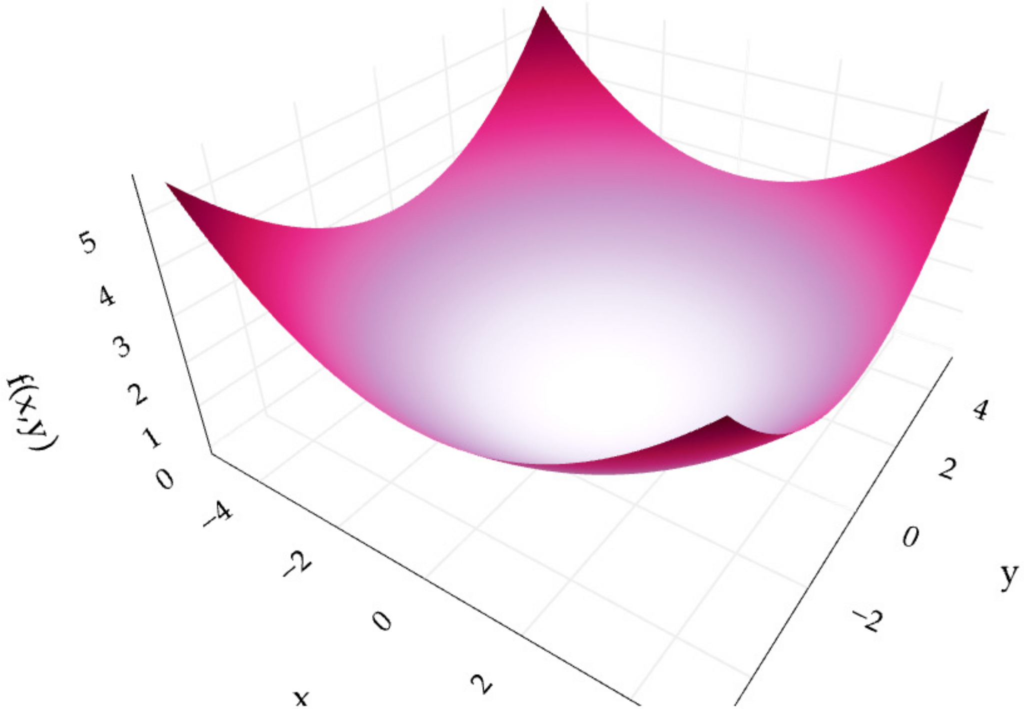
2) $(y_i - h(x_i))^2$

3) $$\underbrace{R_{sq}(w_0, w_1) = \frac{1}{n} \sum_{i=1}^{n} (y_i - (w_0 + w_1 x_i))^2}_{\text{function of } w_0,\ w_1}$$

Aside : Functions with multiple input variables

e.g. $f(x,y) = \dfrac{x^2+y^2}{9}$

"paraboloid"



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

partial derivative of f $\underbrace{\text{wrt}}$ x $= \dfrac{\partial f}{\partial x}$

"with respect to"

curly "d"

$\dfrac{\partial f}{\partial x}$ = derivative of f wrt x, while holding y constant

---

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

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

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

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

example $(-3, 0.5)$

$$\frac{\partial f}{\partial x}(-3, 0.5) = -\frac{6}{9} = -\frac{2}{3}$$

$$\frac{\partial f}{\partial y}(-3, 0.5) = \frac{2(0.5)}{9} = \frac{1}{9}$$

Another ex:

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

$$\frac{\partial g}{\partial x}(x,y) = 3x^2 - 3y^2 + 2\cos(x)\cos(y)$$

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

$$(-3x)\frac{d}{dy}y^2$$

$$= (-3x)(2y)$$

$$= -6xy$$

What's the point?

partial derivatives help us minimize / maximize functions

→ strategy: set all partial derivatives to 0, solve the resulting system

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

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

at $x=0$, $y=0$, both are 0!

back to the main plot:

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

plan:

compute $\frac{\partial R}{\partial w_0}(w_0, w_1)$, compute $\frac{\partial R}{\partial w_1}(w_0, w_1)$ set both $= 0$, solve

both will involve both params

$$\left.\begin{aligned} \overbrace{\frac{\partial R}{\partial w_0}}^{} &= 0 \\ \frac{\partial R}{\partial w_1} &= 0 \end{aligned}\right\} \text{system of 2 eq'ns, 2 unknowns}$$

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

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

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

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

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

Now, need to solve:

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

optimal intercept

plan:

→ In eq (1), isolate for $w_0^* = \underline{\qquad\qquad}$.

→ Substitute w_0^* into eq (2), isolate for $w_1^* = \underline{\qquad\qquad}$.

isolate for w_0.

$$-\frac{2}{n}\sum_{i=1}^{n}\left(y_i-(w_0+w_1x_i)\right)=0$$

$$\sum_{i=1}^{n}\Big(\underbrace{y_i}_{\text{actual}}-\underbrace{(w_0+w_1x_i)}_{\text{pred}}\Big)=0$$

"balance condition"

$\sum \text{errors} = 0$

$$\sum_{i=1}^{n}(y_i-w_0-w_1x_i)=0$$

$$\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}w_0-\sum_{i=1}^{n}w_1x_i=0$$

$$\sum_{i=1}^{n}y_i-nw_0-w_1\sum_{i=1}^{n}x_i=0$$

$$\sum_{i=1}^{n}y_i-w_1\sum_{i=1}^{n}x_i=nw_0$$

$$\frac{\sum_{i=1}^{n} y_i - w_1 \sum_{i=1}^{n} x_i}{n} = \frac{n w_0}{n}$$

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

opt. slope

optimal intercept

next: substitute

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

into

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

substitute

isolate w_1

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

$$w_1^* = \frac{\sum_{i=1}^{n} x_i(y_i-\bar{y})}{\sum_{i=1}^{n} x_i(x_i-\bar{x})} = \boxed{\frac{\sum_{i=1}^{n} (x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n} (x_i-\bar{x})^2}}$$

proof: see Ch. 1.4

$$\sum_{i=1}^{n} \underbrace{(x_i-\bar{x})(y_i-\bar{y})}_{\text{distribute}} = \sum \left(x_i(y_i-\bar{y}) - \bar{x}(y_i-\bar{y}) \right)$$

$$= \sum x_i(y_i-\bar{y}) - \bar{x} \underbrace{\sum (y_i-\bar{y})}_{=0}$$

To summarize,

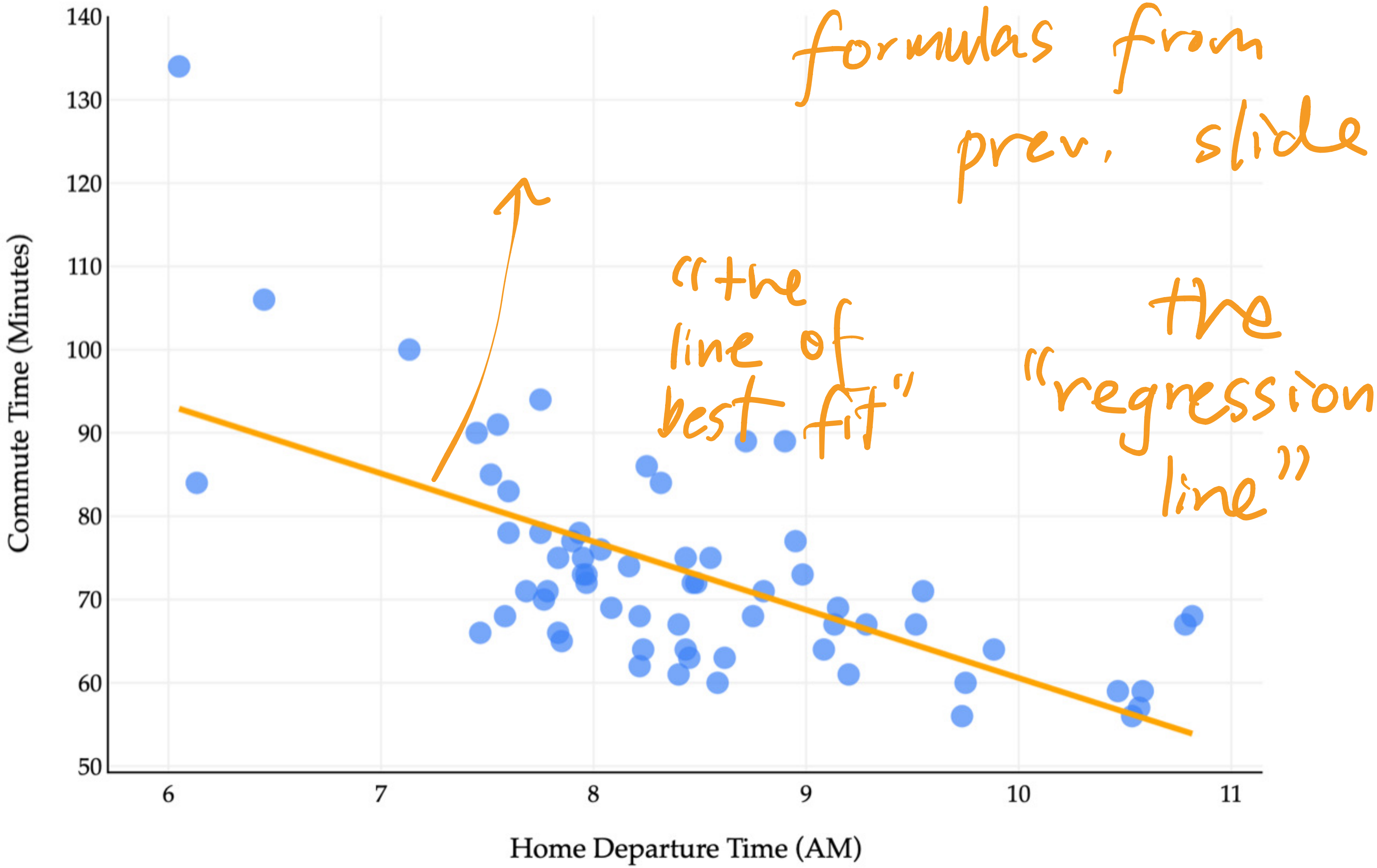
$$w_1^* = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

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

optimal parameters for $h(x_i) = w_0 + w_1 x_i$

USING SQUARED LOSS

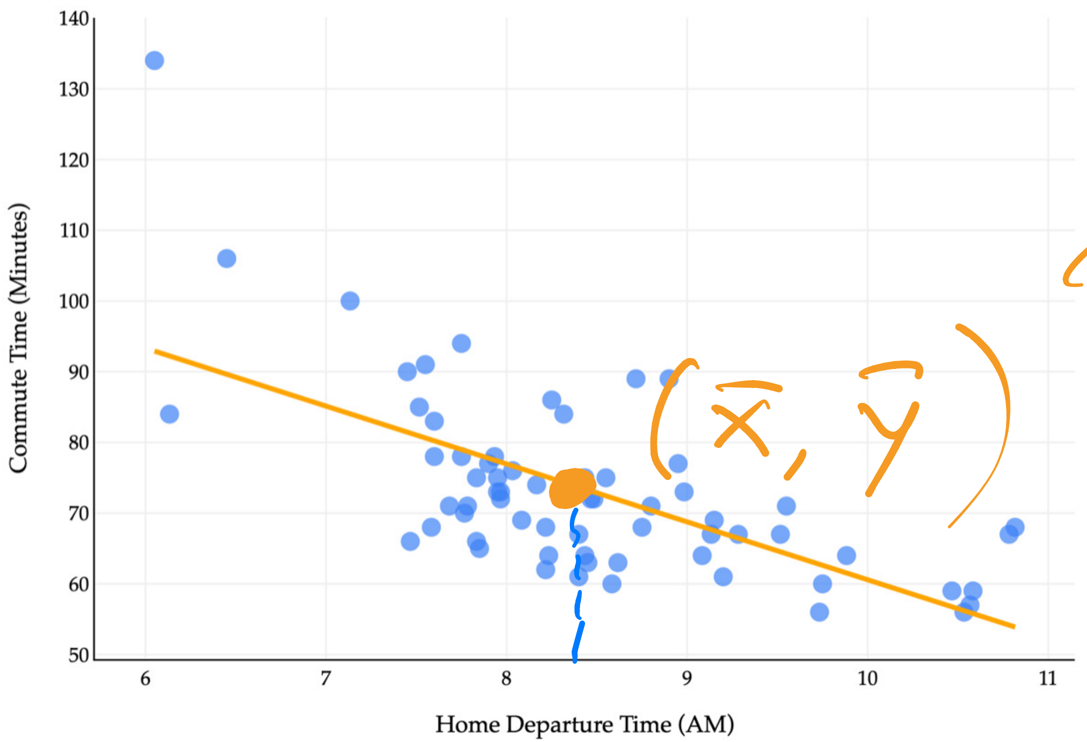
calculated using formulas from prev. slide



Fact ! The regression line

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

always passes through the point $(\bar{x}, \bar{y})$



why ?

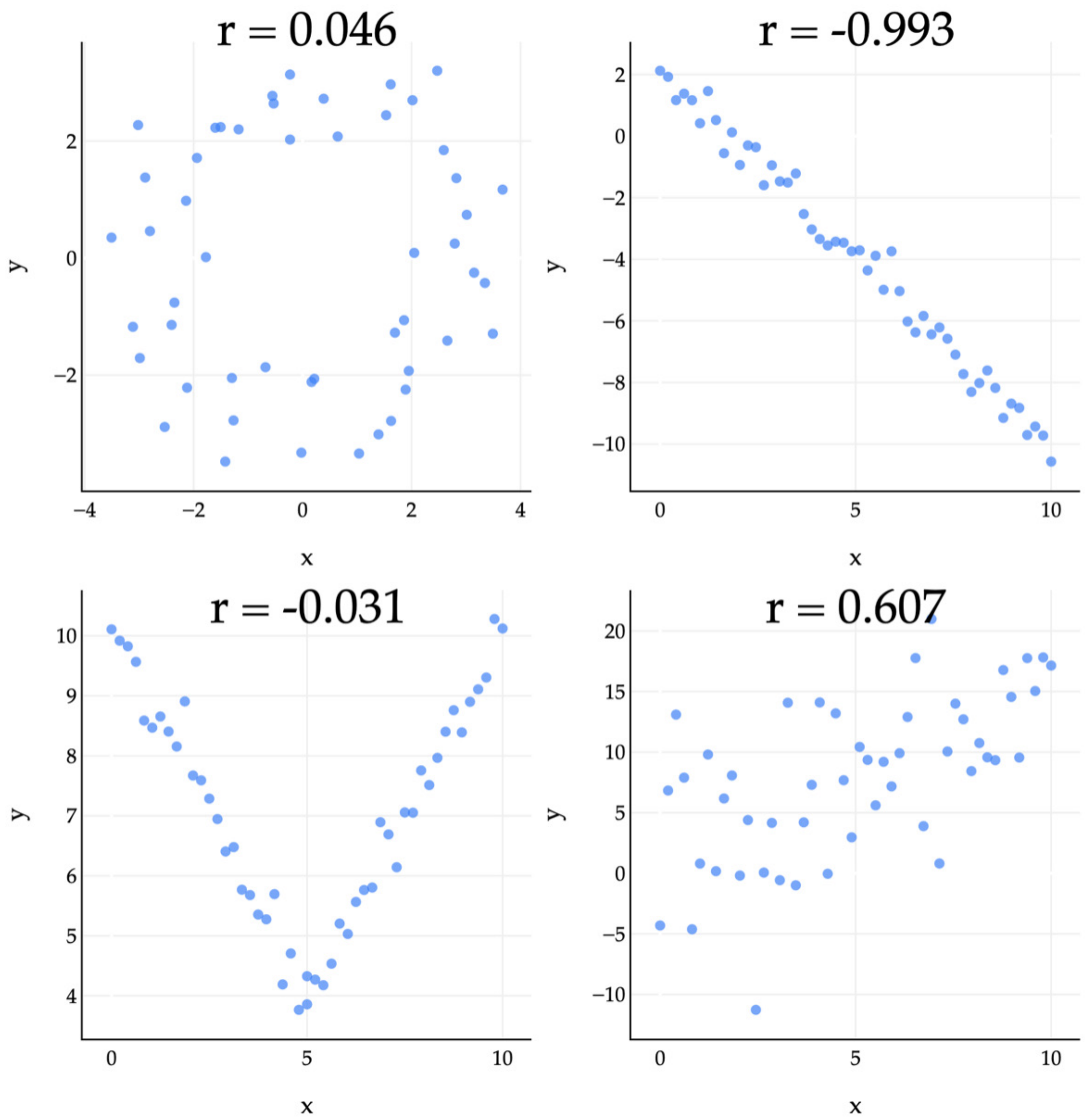
$$w_0^* = \bar{y} - w_1^* \bar{x}$$

plug in $x_i = \bar{x}$

$$h(\bar{x}) = w_0^* + w_1^* \bar{x}$$

$$= \underbrace{\bar{y} - w_1^* \bar{x} + w_1^* \bar{x}}_{} = \bar{y}$$

Correlation coefficient, r $\quad -1 \leq r \leq 1$



1)

$$r = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma_x} \right) \left(\frac{y_i - \bar{y}}{\sigma_y} \right)$$

$$\frac{x_i - \text{mean}}{SD}$$

$= \#$ SDs x_i is above the mean