M339G: January 28th, 2026.

Why $(\quad)^2$?

$$\mathbb{E}[(X-a)^2] \xrightarrow[a]{} \min$$

sol'n = ? $a = \mathbb{E}[X]$

Then, the value is $\mathbb{E}[(X-\mathbb{E}[X])^2] = Var[X]$

Reducible vs. Irreducible.

Fact:

$$\mathbb{E}[(Y - \hat{f}(X))^2 \mid X = x] = ?$$

By our model

$\boxed{Y = f(X) + \varepsilon}$ w/ ε independent from X

and $\boxed{\mathbb{E}[\varepsilon] = 0}$

$$\mathbb{E}[(f(X) + \varepsilon - \hat{f}(X))^2 \mid X = x] = \text{linearity of } \mathbb{E}$$

$$= \mathbb{E}[(f(X) - \hat{f}(X))^2 \mid X = x] +$$

$$+ \cancel{2\,\mathbb{E}[(f(X) - \hat{f}(X)) \cdot \varepsilon \mid X = x]}$$

$$+ \mathbb{E}[\varepsilon^2 \mid \cancel{X = x}]$$

$\varepsilon \perp X$ (independent)

$$= (f(x) - \hat{f}(x))^2 + \mathbb{E}[\varepsilon^2]$$

$$\mathbb{E}\left[(Y-\hat{f}(X))^2 \mid X=x\right] = \underbrace{(f(x)-\hat{f}(x))^2}_{\text{Reducible}} + \underbrace{\mathrm{Var}[\varepsilon]}_{\text{Irreducible}}$$

Simple Linear Regression.

In general. f... the ideal fit, i.e.,

$$\boxed{Y = f(X) + \varepsilon} \text{ w/ } \varepsilon \text{ independent of } X$$

and $\mathbb{E}[\varepsilon] = 0$

$$\boxed{\mathbb{E}[Y \mid X = x] = \underline{f(x)}}$$

Solve the optimization problem:

$$\mathbb{E}\left[(Y-g(X))^2 \mid X=x\right] \xrightarrow[g]{} \min$$

=> $\hat{f}$... the fitted f'tion/fit in a family of candidates g

"Def'n." In simple linear regression, the model is

$$\boxed{Y = \beta_0 + \beta_1 \cdot X + \varepsilon}$$

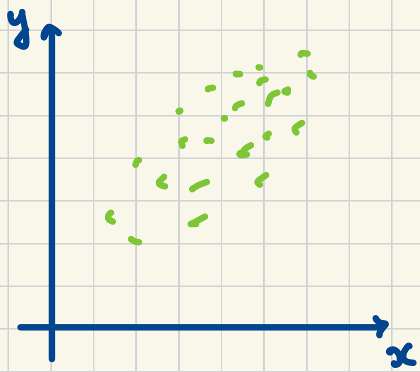
w/ ε and X are independent and $\varepsilon \sim N(0, \sigma^2)$

Method: Find ESTIMATORS

$$\boxed{\begin{matrix} \hat{\beta}_0 \text{ and } \hat{\beta}_1 \\ \beta_0 \text{ and } \beta_1 \end{matrix}} \begin{matrix} \text{and } \hat{\sigma} \\ \text{and } \sigma. \end{matrix}$$

for β_0 and β_1 and σ.

Set up:



data set:

$(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$