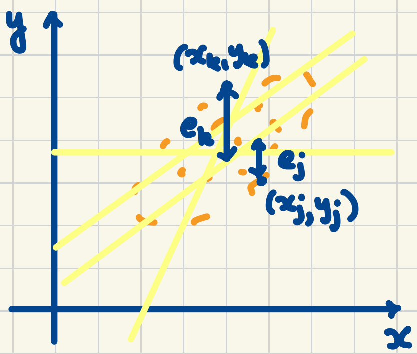
M339G: January 30th, 2026.

Simple Linear Regression [cont'd].

The Model.

$Y = \beta_0 + \beta_1 X + \varepsilon$ w/ $\varepsilon \perp X$
and $\varepsilon \sim N(0, \sigma^2)$



data set:

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

e_k, $k = 1..n$
RESIDUALS

Every line of fit would have the form

$$\boxed{\hat{y} = \hat{b}_0 + \hat{b}_1 x}$$

w/ $\hat{b}_0$ and $\hat{b}_1$ "candidate" coefficients

$$SSE = RSS = \sum_{k=1}^{n} e_k^2 \longrightarrow \min$$

$$\sum_k (y_k - \boxed{\hat{y}_k})^2 \longrightarrow \min$$

$$\sum_k (y_k - \hat{b}_0 - \hat{b}_1 x_k)^2 \xrightarrow{\hat{b}_0, \hat{b}_1} \min$$

For unbiasedness: $\sum_k e_k = 0$

Differentiate w/ respect to $\hat{b}_0$ ✓ and $\hat{b}_1$.

$$\frac{\partial RSS}{\partial \hat{b}_0} = -2 \cdot \sum_{k=1}^{n} (y_k - \hat{b}_0 - \hat{b}_1 x_k) = 0$$

$$\sum_k y_k = \sum_k (\hat{b}_0 + \hat{b}_1 x_k) = n \cdot \hat{b}_0 + \hat{b}_1 \sum_k x_k \quad /:n$$

$$\underbrace{\frac{1}{n} \sum_k y_k} = \hat{b}_0 + \hat{b}_1 \cdot \underbrace{\left(\frac{1}{n} \sum_k x_k\right)}$$

$$\boxed{\bar{y} = \hat{b}_0 + \hat{b}_1 \bar{x}}$$ normal equation

$(\bar{x}, \bar{y})$ is on the least·squares line.

$$\frac{\partial RSS}{\partial \hat{b}_1} = \underbrace{-2 \sum_k \left((y_k - \hat{b}_0 - \hat{b}_1 x_k) \cdot x_k\right)}_{=0} = 0$$

$$\sum_k x_k y_k - \hat{b}_0 \sum_k x_k - \hat{b}_1 \sum_k x_k^2 = 0$$

$$\hat{b}_1 = \frac{\cancel{\frac{1}{n}} \sum (x_k - \bar{x})(y_k - \bar{y})}{\cancel{\frac{1}{n}} \sum (x_k - \bar{x})^2} = \frac{n \cdot \sum x_k y_k - \sum x_k \cdot \sum y_k}{n \cdot \sum x_k^2 - (\sum x_k)^2}$$

$$= \frac{\sum x_k y_k - \frac{1}{n} \sum x_k \cdot \sum y_k}{\sum x_k^2 - \frac{1}{n} (\sum x_k)^2}$$

Intuition:

$$\hat{\beta}_1 = \text{"} \frac{Cov[X,Y]}{Var[X]} \text{"} = \text{"} \frac{\rho_{X,Y} \cdot \cancel{\sigma_X} \cdot \sigma_Y}{\sigma_X^{\cancel{2}}} \text{"}$$

$$= \text{"} \rho_{XY} \frac{\sigma_Y}{\sigma_X} \text{"}$$

$$\boxed{\hat{\beta}_1 = \frac{S_{XY}}{S_X^2} \qquad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \cdot \bar{X}}$$

Our Estimators

$$\boxed{\hat{\sigma}^2 = \frac{RSS}{n-2}}$$

From the $\varepsilon \perp X$ requirement, we get the other normal equation:

$$\boxed{\sum_{k=1}^{n} e_k x_k = 0}$$