Least Squares: Find vector b such that

$$\underbrace{\frac{\partial L}{\partial b_0}; \frac{\partial L}{\partial b_1}; \frac{\partial L}{\partial b_2}; \cdots \frac{\partial L}{\partial b_k}}_{}$$ are all $\emptyset$

Compute gradient of L with respect to b

$$\frac{\partial L}{\partial b} = \begin{bmatrix} \frac{\partial L}{\partial b_0} \\ \frac{\partial L}{\partial b_1} \\ \vdots \\ \frac{\partial L}{\partial b_k} \end{bmatrix}$$

$$L = y^T y - 2y^T X b + b^T X^T X b$$

$$\frac{\partial x^n}{\partial x} = nx^{n-1}$$

$$\frac{\partial L}{\partial b} = \emptyset - 2y^T X + 2b^T X^T X$$

Find b such that $\frac{\partial L}{\partial b} = \emptyset$

$$\Rightarrow -2y^T X + 2\hat{b}^T X^T X = 0$$

$$\hat{b}^T X^T X = y^T X$$

$$(X^T X)\hat{b} = X^T y \Rightarrow \text{Normal Equations}$$

Given that $(X^T X)$ can be inverted:

$$\hat{b} = (X^T X)^{-1} X^T y$$