Least Squares

∇f gradient
element wise derivatives of f

$$f(\beta) = \|y - X\beta\|^2$$
$$= (y - X\beta)^t (y - X\beta)$$
$$= y^t y - 2y^t X \beta + \beta^t X^t X \beta$$

$$\nabla f(\beta) = -2X^T y + 2X^t X \beta$$

root of gradient

$$0 = -2X^T y + 2X^t X \beta$$
$$\beta = (X^t X)^{-1} X^t y$$

$$2X_y = 2X^t X \beta$$
$$\frac{2X_y}{2X^t X} = \beta$$
$$\beta = (X^t X)^{-1} X^t y$$

Matrix Opperations in R

$\beta = (X^t * X)^{-1} * X^t * Y$

β = solve(t(X) %*% X) %*% t(X) %*% y

Averages

$$\bar{y} = (1_n^t 1_n)^{-1} 1_n^t y$$

Centering

Centering y

$$y - 1_n \bar{y}$$
$$\tilde{y} = y - (1_n^t 1_n)^{-1} 1^t 1 y = \{I - 1_n (1_n^t 1_n)^{-1} 1^t\} y$$

Row Centered vs. Column Centered

$\tilde{X} = \{I - 1_n (1_n^t 1_n)^{-1} 1_n^t\} X$ Col centered

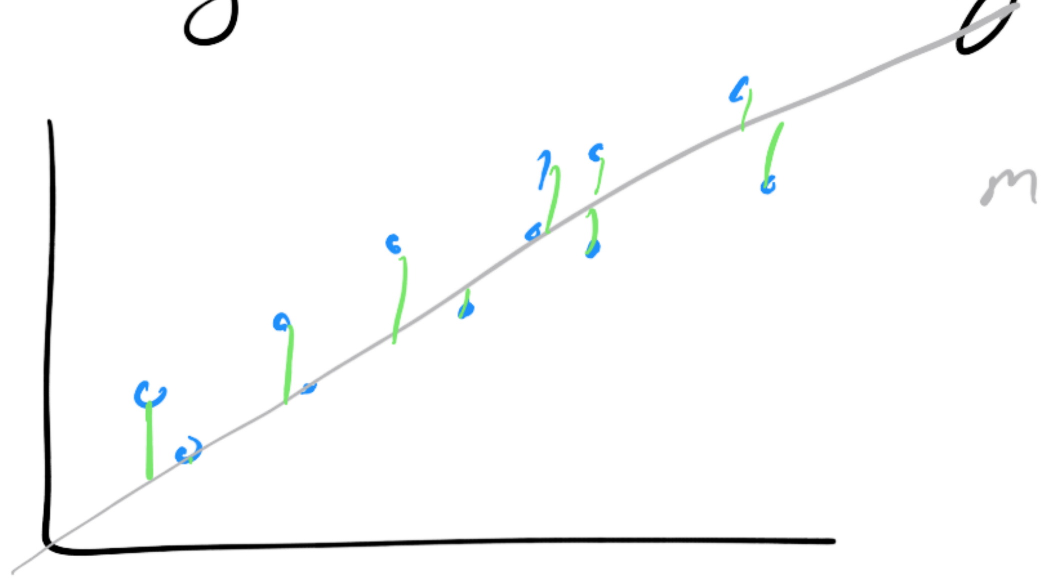
$\tilde{X} = X \{I - 1_n (1_n^t 1_n)^{-1} 1_n^t\}$ Row Centered

Variance

$$s^2 = \frac{1}{n-1} \|y - 1_n \bar{y}\|^2 = \frac{1}{n-1} \tilde{y}^t \tilde{y}$$

Single Parameter Basics

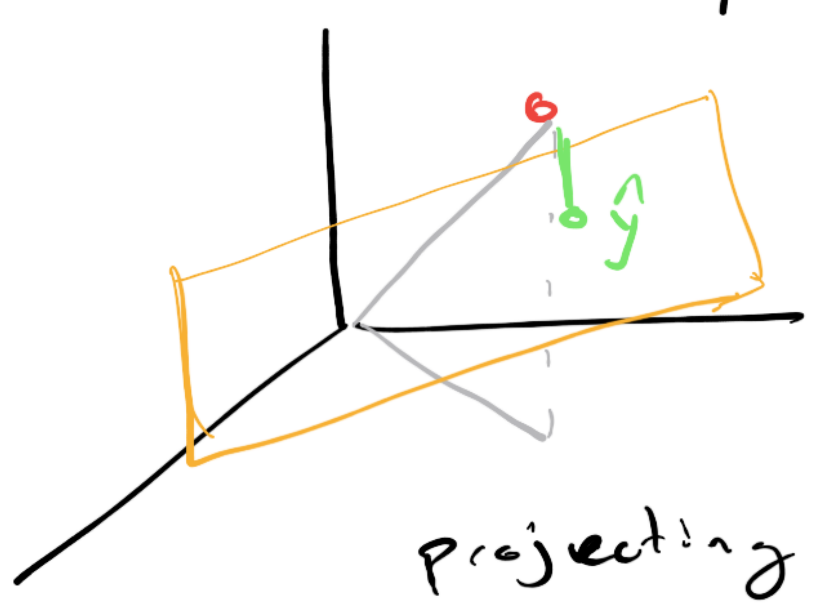
Thinking about minimizing residuals



minimize residuals

Projection in n Dimensional space

$$\|y - \{\beta_0 J_n + \beta_1 X\}\|^2$$
$$\Gamma = \{\beta_0 J_n + \beta_1 X \mid \binom{\beta_0}{\beta_1} \in \mathbb{R}^2\}$$



projecting y onto Γ $\hat{y}$

Centered Regression Through origin

$$\hat{\beta} = \frac{\langle \tilde{y}, \tilde{x} \rangle}{\langle \tilde{x}, \tilde{x} \rangle} = \hat{\rho}_{yx} \frac{\hat{\sigma}_y}{\hat{\sigma}_x}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$