Ordinary Least Square = Maximum likelihood estimators

$$y \approx X\beta \qquad \|y - X\beta\|_2^2 = \sum_{i=1}^{n} (y_i - x_i^t \beta)^2$$

parameters for our random variable

$$\{y_i, x_i\}_{i=1}^{n} \qquad y_i \sim N(x_i^t \beta, \sigma^2)$$

random variable

$$\text{Likelihood} = \prod_{i=1}^{n} p(y_i \mid x_i, \beta, \sigma^2)$$

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y_i - x_i^t\beta)^2}{2\sigma^2}}$$

MLE

$$\text{loglik}(\beta, \sigma^2) = \sum_{i=1}^{n} -\frac{1}{2} \log \sigma^2 - \frac{(y_i - x_i^t \beta)^2}{2\sigma^2}$$

$$= -\frac{n}{2} \log \sigma^2 - \frac{\sum_{i=1}^{n} (y_i - x_i^t \beta)^2}{2\sigma^2}$$

$$\hat{\beta}_{mle} = \underset{\beta \in \mathbb{R}^p}{\text{argmin}} \sum_{i=1}^{n} (y_i - x_i^t \beta)^2$$

$$\hat{\sigma}^2_{mle} = \frac{RSS}{n}$$

objective function

Score

$$J(\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{RSS}{2\sigma^2}$$

$$J'(\sigma^2) = \frac{-n}{\sigma^2} + \frac{RSS}{(\sigma^2)^2}$$

Value of log liklihood at evaluated at MLE

$$\text{loglik}(\beta,\sigma^2) = -\frac{n}{2}\log\sigma^2 - \frac{\sum_{i=1}^{n}(y_i - x_i^t\beta)^2}{2\sigma^2}$$

$$\text{loglik}(\hat\beta_{MLE},\hat\sigma^2_{MLE}) = -\frac{n}{2}\log\frac{RSS}{n} - \frac{\cancel{RSS}}{2\cdot(\cancel{\frac{RSS}{n}})}$$

$$= -\frac{n}{2}\log\frac{RSS}{n} - \frac{n}{2} = -\frac{n}{2}\left(\log\frac{RSS}{n} + 1\right)$$

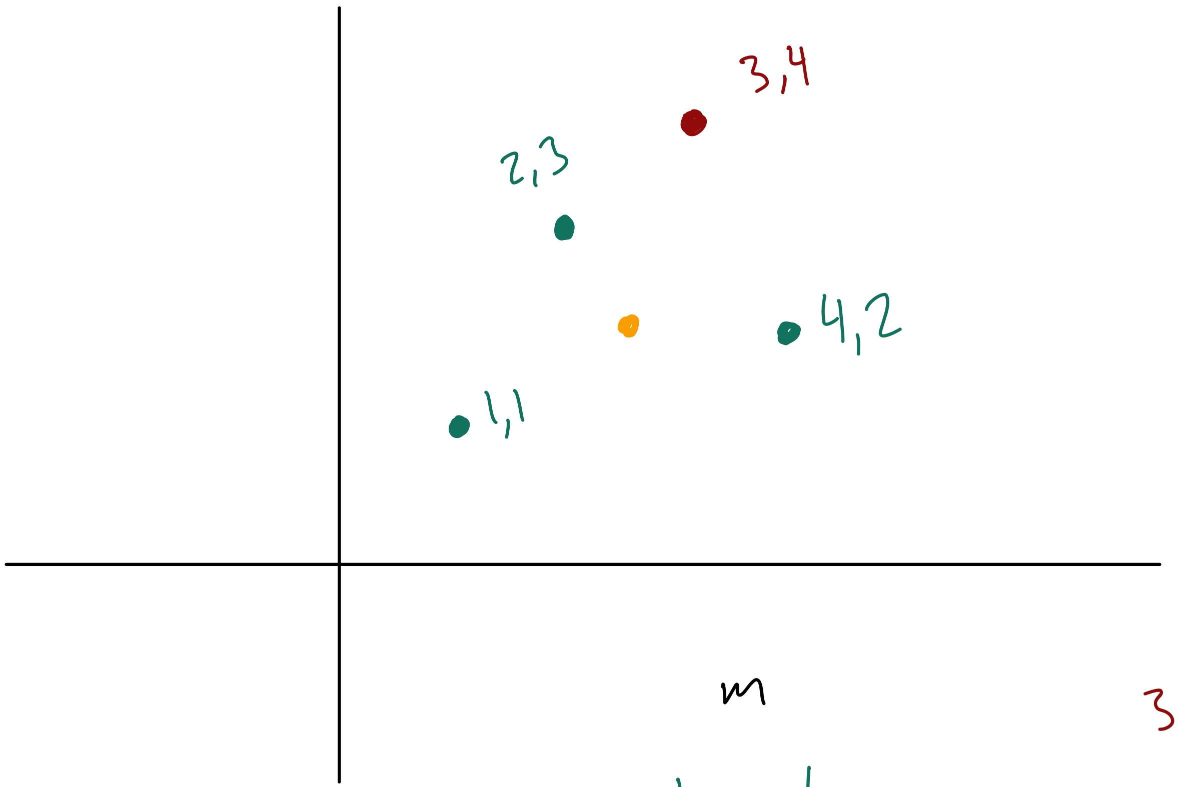
$$Y_{n\times 1} \sim N_n(X\beta, \sigma^2 I_n)$$ Gaussian Linear Regression

$$\hat\beta = (X^TX)^{-1}X^Ty \qquad \hat y_{n\times 1} = X\hat\beta = \underbrace{X(X^TX)^{-1}X^T}_{n\times n}\underbrace{y}_{n\times 1}$$

$\downarrow$ projection matrix

F-stat

Large F-stat means our larger model contributes significantly enough



m

3,4

1 , 1
2 , 3
4 , 2

$$\sum \| (3,4) - m \|^2$$

$$= \sqrt{(3-1)^2 + (4-1)^2} + \sqrt{(3-2)^2 + (4-3)^2} + \sqrt{(3-4)^2 + (4-2)^2} = \sqrt{13} + \sqrt{2} + \sqrt{5}$$

$$\| (3,4) - (7,6) \|^2 = \sqrt{(3-7)^2 + (4-6)^2} = \sqrt{16 + 4} = \sqrt{20}$$

```
sum=0
for each center in m1:
    norm_vec = X
    norm_vec[:,0] -= center[0]
    norm_vec[:,1] -= center[1]
    sum += exp thing with ||norm_vec||^2
```