FYS-STK 4155 SEPT 15

Bayes' theorem

$$D = \{[x_0, y_0] \dots [x_{n-1}, y_{n-1}]\}$$

OLS

$$p(y_i|\beta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - X_{i*}\beta)^2}{2\sigma^2}\right]$$

i.i.d.

$$p(D|\beta) = \prod_{i=0}^{n-1} p(y_i|\beta)$$

product rule

$$p(A,B) = p(A \cap B) =$$

$$p(A|B)\,p(B) = p(B|A)\,p(A)$$

Marginal distribution

$$p(A) = \sum_b p(A, B=b)$$

$$= \sum_b P(A \mid B=b)\, P(B=b)$$

conditional probability

$$P(A \mid B) = \frac{P(A, B)}{P(B)} \qquad P(B) > 0$$

$\Longrightarrow$

Bayes' theorem

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{\sum_a P(B \mid A=a)\, P(A=a)}$$

posterior probability

likelihood

prior

$$P(B \mid A) \longrightarrow P(D \mid B)$$

$$P(A \mid B) \longrightarrow P(B \mid D)$$

$$P(B \mid D) \propto P(D \mid B)\, P(B)$$

how to model $P(\beta)$?

OLS $P(\beta)$ is given by uniform distribution.

$$P(\beta) = \prod_{j=0}^{p-1} e^{-\beta_j^2/2\tau^2}$$

$$\beta \sim N(0, \tau^2)$$

$$P(\beta|D) = \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi^2\sigma^2}} \exp\left[-\frac{(y_i - x_{i*}\beta)^2}{2\sigma^2}\right] \times \prod_{j=0}^{p-1} e^{-\beta_j^2/2\tau^2}$$

$$C(\beta) = -\log P(\beta|D)$$

$$= \frac{n}{2}\log(2\pi\sigma^2) + \frac{\|(y - X\beta)\|_2^2}{2\sigma^2}$$

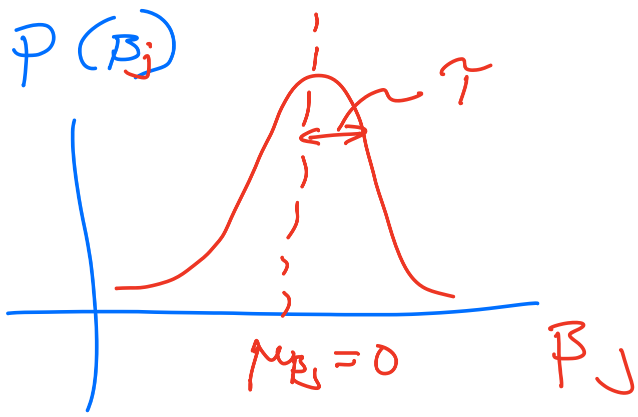
$$+ \frac{1}{2\tau^2} \|\beta\|_2^2$$

$$\frac{1}{2\tau^2} \Rightarrow \lambda \Rightarrow$$

$$C(\beta) = \frac{\|(y - X\beta)\|_2^2}{2\sigma^2} + \lambda \|\beta\|_2^2$$

= Ridge Regression ?

$\lambda > 0$



shrinking τ (increasing λ) or increasing τ (decreasing λ) affects the variance (and thereby standard deviation) of the parameters β_j

Laplace distribution

$$P(\beta_j) = e^{-|\beta_j|/\tau}$$

$$\mu_{\beta_j} = 0$$

$$P(\beta|D) \propto \prod_{i=0}^{n-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y_i - x_i\beta)^2}{2\sigma^2}\right]$$

$$\times \prod_{j=0}^{p-1} e^{|\beta_j|/\tau}$$

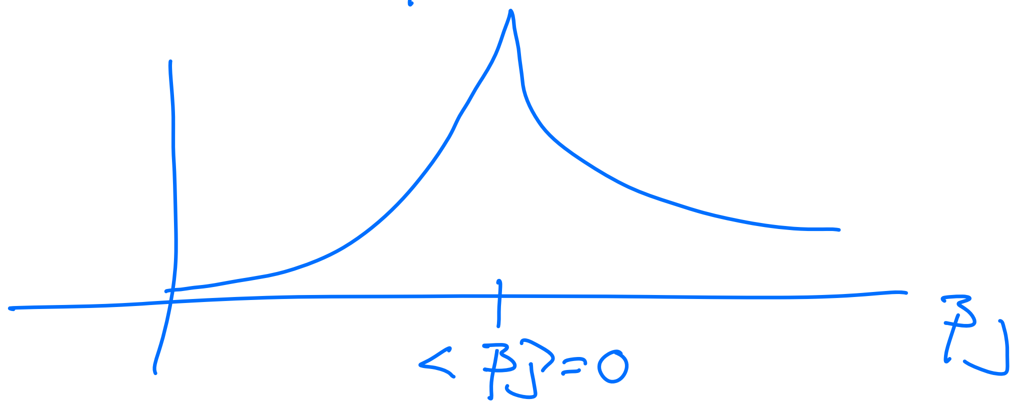
$$-\log P(\beta|D) = C(\beta)$$

Dropping constant

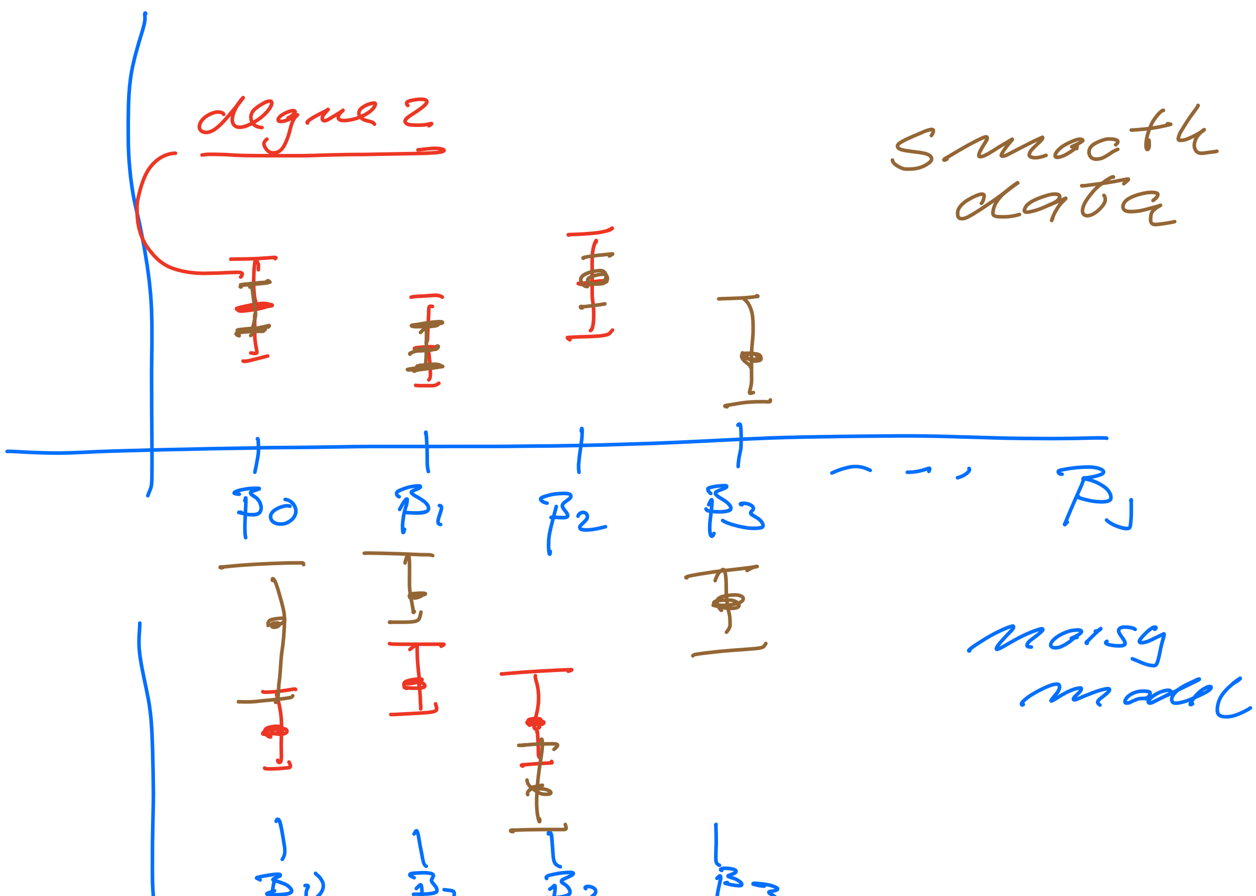
$$= \frac{||(y - X\beta)||_2^2}{2\sigma^2} + \frac{1}{\tau}||\beta||_1$$

Lasso = Least absolute shrinkage and selection operation.

$$\lambda = \frac{1}{\tau}$$



Calculations with different models (project 1)



Bias-variance tradeoff + Resampling techniques

- Bootstrap (Jackknife)
- Cross-validation

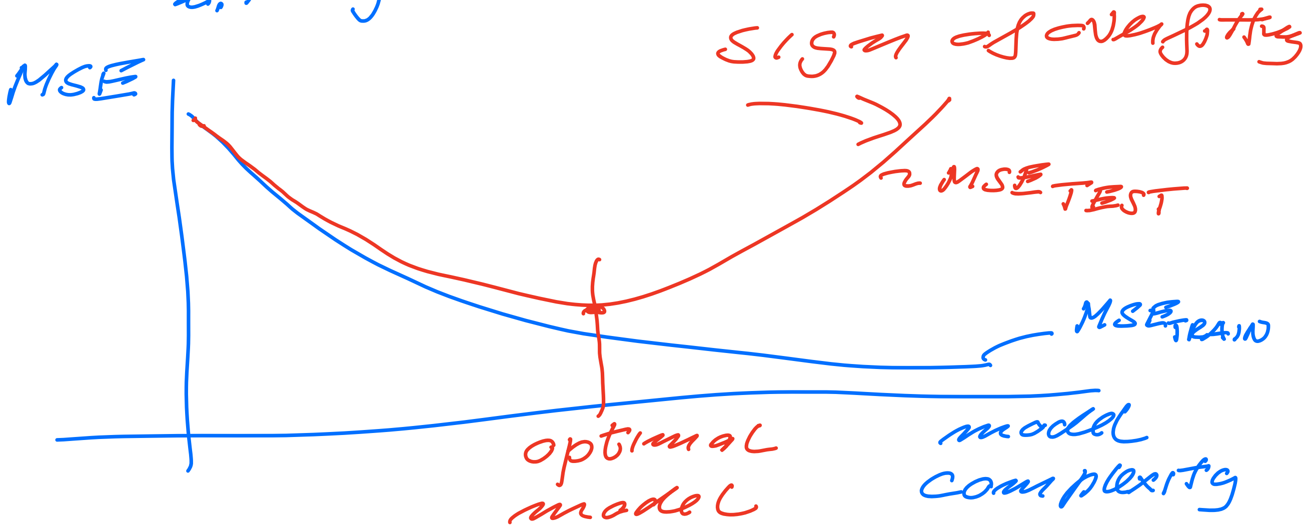
Estimate the MSE and other expectation values

$$MSE = \frac{1}{n}\sum_{i=0}^{n-1}(y_i - \tilde{y}_i)^2$$

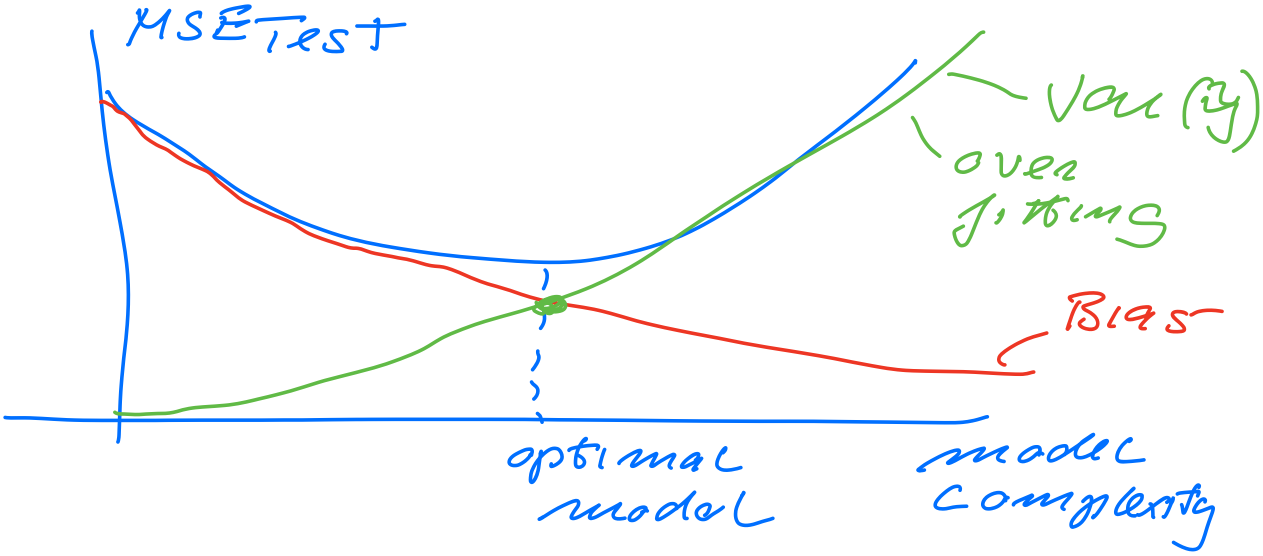
$$= \text{Bias} + \text{var}(\tilde{y}) + \sigma^2$$

$$y = \underbrace{f(x)}_{\tilde{y}} + \varepsilon$$

2.11 of Hastie et al



Bias-variance tradeoffs



Bootstrap + central limit theorem:

Bootstrap:

Domain $D = [z_0, z_1, \dots z_{n-1}]$

calculate $\mu = \frac{1}{n}\sum_{i=0}^{n-1} z_i$

Reshuffle data (randomly) with replacement

z_0 two times

$D' = [z_0' \; z_1' \; \underset{z_1'}{z_0}, \; z_{n-1}']$

repeat B-times.

assume you want to estimate μ. For each sample μ_b

$$\mu_B = \frac{1}{B} \sum_{b=1}^{B} \mu_b$$

$$\mu_b = \frac{1}{n} \sum_{i=0}^{n-1} z_i$$

suppose $z \sim N(\mu, \sigma^2)$

New distribution

$$\sim N\left(\mu, \frac{\sigma^2}{B}\right)$$