FYS-STK4155 Sept 9

SVD: $X = U \Sigma V^T$

$$U U^T = U^T U = \mathbb{1}$$

$$U \in \mathbb{R}^{n \times n} \qquad X \in \mathbb{R}^{n \times p}$$

$$n \geq p$$

$$V V^T = V^T V = \mathbb{1}$$

$$X^T X = V^T \Sigma^T \Sigma V$$

$$\Sigma = \begin{bmatrix} \sigma_0 & & & 0 \\ & \sigma_1 & & \\ & & \ddots & \\ 0 & & & \sigma_{p-1} \\ & & & 0 \end{bmatrix}$$

$$\tilde{y}_{OLS} = X \hat{\beta} = X (X^T X)^{-1} X^T y$$

$$= U U^T y$$

$$= \left(\sum_{i=0}^{p-1} u_i u_i^T \right) y$$

$$\tilde{y}_{OLS} = \left[(U \Sigma V^T)(V \Sigma^T \Sigma V^T)^{-1} \times V \Sigma^T U^T \right] y$$

$$\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}$$

$$\Sigma^T \Sigma = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

$$U \in \mathbb{R}^{3 \times 3} \qquad V \in \mathbb{R}^{2 \times 2}$$

$$\Big\{ \underbrace{U}_{3 \times 3} \underbrace{\Sigma}_{3 \times 2} \underbrace{\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{bmatrix}}_{2 \times 2} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{2 \times 3} \underbrace{U^T}_{3 \times 3} \Big\} y$$

$$U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T$$

$$U = \left[u_0 \; u_1 \; u_2 \right]$$

$$\Rightarrow \left(u_0 u_0^T + u_1 u_1^T \right) y$$

$$\tilde{y}_{OLS} = \left(\sum_{j=0}^{p-1} u_j u_j^T \right) y$$

$$\tilde{y}_{Ridge} = \left(\sum_{j=0}^{p-1} u_j u_j^T \frac{\sigma_j^2}{\sigma_j^2 + \lambda} \right) y$$

Example

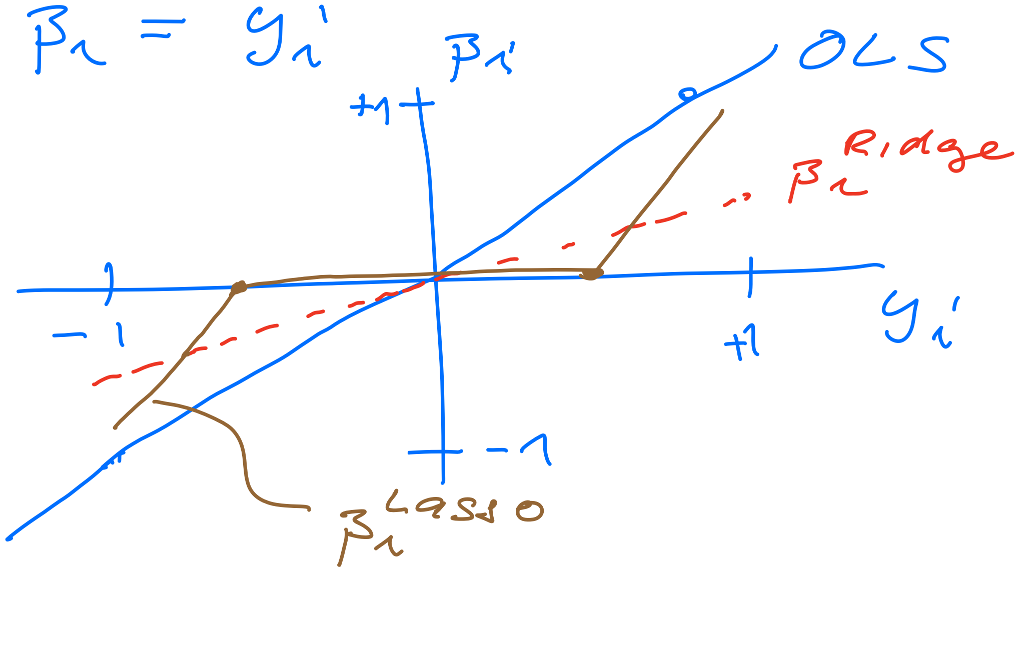
$\beta_i = \tilde{y}_i \quad , \quad n = p$

OLS $\qquad X = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{bmatrix}$

$$C(\beta) = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \beta_i)^2$$

$$\frac{\partial C}{\partial \beta_j} = 0 \Rightarrow -\frac{2}{n}(y_i - \beta_i) = 0 \Rightarrow$$

$$\hat{\beta}_i = y_i$$



Ridge

$$C(\beta) = \frac{1}{n}\sum_{i=0}^{n-1}(y_i-\beta_i)^2 + \lambda\sum_{i=0}^{n-1}\beta_i^2$$

$$\frac{\partial C}{\partial \beta_j} = 0 \Rightarrow$$

$$-\frac{2}{n}(y_i-\beta_i)+2\lambda\beta_i = 0$$

$$-(y_i-\beta_i) + \tilde{\lambda}\beta_i = 0$$

$$\tilde{\lambda} \Rightarrow \lambda$$

$$\beta_i^{Ridge} = \frac{y_i}{1+\lambda}$$

Lasso

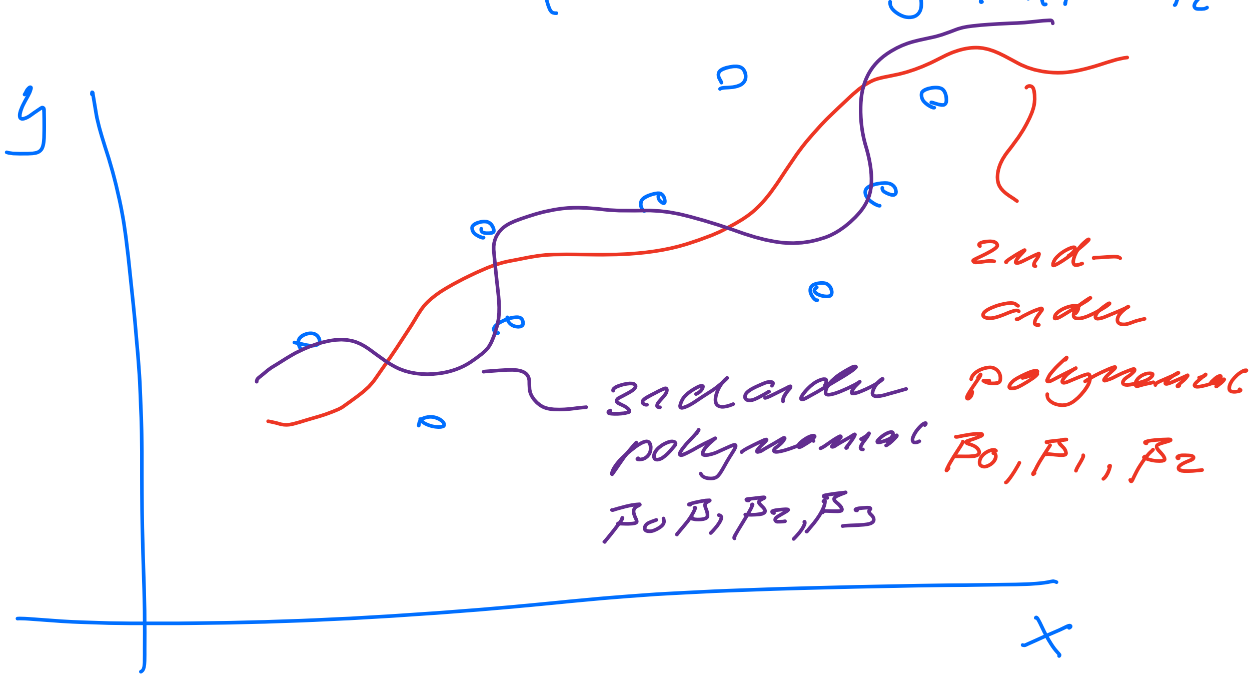
$$C(\beta) = \frac{1}{n}\sum_{i=0}^{n-1}(y_i-\beta_i)^2 + \lambda\sum_{i=0}^{n-1}\sqrt{\beta_i^2}$$

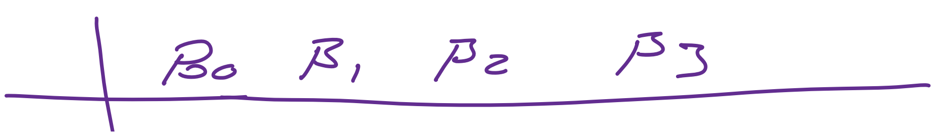
$$\frac{d|\beta_i|}{d\beta_i} = \frac{\beta_i}{|\beta_i|} \qquad |\beta_i| = \sqrt{\beta_i^2}$$

$$= \begin{cases} +1 & \beta_i > 0 \\ (0 & \beta_i = 0 \,) \\ -1 & \beta_i < 0 \end{cases}$$

$$\frac{\partial C}{\partial \beta_i} = 0 = -\frac{2}{n}(y_i - \beta_i) + \lambda \frac{\beta_i}{|\beta_i|} = 0$$

$$\hat{\beta}_i^{\,Lasso} = \begin{cases} y_i - \lambda/2 & \text{if } y_i > \lambda/2 \\ y_i + \lambda/2 & \text{if } y_i < -1/2 \\ 0 & \text{if } |y_i| \le \lambda/2 \end{cases}$$





2	1.0	0.5	0.01	0
3	0.9	-0.3	2	3
3	1.0	0.45	0.02	3.1

$$\mathrm{var}(\hat{\beta}_{OLS}) = \sigma^2 (X^T X)^{-1}$$

$$\mathrm{var}(\hat{\beta}_{OLS})_j = \sigma^2 (X^T X)^{-1}_{jj}$$

$$\mathrm{std} = \sqrt{\mathrm{var}(\hat{\beta}_{OLS})} = \Delta\beta$$

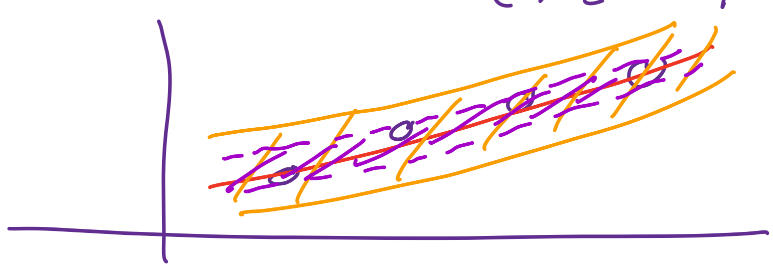
2nd order polynomial

$\beta_0 \pm \Delta\beta_0$

$\beta_1 \pm \Delta\beta_1$

$\beta_2 \pm \Delta\beta_2$

$$\tilde{y}_i = (\beta_0 \pm \Delta\beta_0) + (\beta_1 \pm \Delta\beta_1) x_i + (\beta_2 \pm \Delta\beta_2) x_i^2$$



with Ridge Regression

$var(\hat{\beta}_{Ridge}(\lambda)) < var(\hat{\beta}_{OLS})$