Exercise 1b), week 36

Define U to be an orthogonal matrix $U \in \mathbb{R}^{n\times n}$

Define $V \in \mathbb{R}^{p\times p}$

$$UU^T = U^TU = V^TV = VV^T$$
$$= \mathbb{1}$$

Define $\Sigma \in \mathbb{R}^{n\times p}$

Example $n = 3$ $p = 2$

$$\Sigma = \begin{bmatrix} \sigma_0 & 0 \\ 0 & \sigma_1 \\ 0 & 0 \end{bmatrix}$$

$$U = \begin{bmatrix} u_{00} & u_{01} & u_{02} \\ u_{10} & u_{11} & u_{12} \\ u_{20} & u_{12} & u_{22} \end{bmatrix} = \begin{bmatrix} u_0 & u_1 & u_2 \end{bmatrix}$$

OLS

$$X = U \Sigma V^T \quad (SVD)$$

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

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

$$= U \Sigma V^T (X^T X)^{-1} V \Sigma^T U^T y$$

$$X^T X = V \Sigma^T \Sigma V^T \in \mathbb{R}^{2\times 2}$$

inverse of square and invertible matrices

$$(V \Sigma^T \Sigma V^T)^{-1} = (V^T)^{-1} (\Sigma^T \Sigma)^{-1} \times V^{-1}$$

$$(V^T)^{-1} = V \qquad V^{-1} = V^T$$

$$\Rightarrow$$

$$\tilde{y} = \underbrace{U \Sigma}_{3\times 2} \underbrace{V^T \cdot V}_{\substack{\mathbb{1} \\ 2\times 2}} \underbrace{(\Sigma^T \Sigma)^{-1}}_{2\times 2} \underbrace{V^T V}_{\substack{\mathbb{1} \\ 2\times 2}} \underbrace{\Sigma^T U^T y}_{2\times 3}$$

$$\tilde{y} = U \Sigma \frac{1}{\Sigma^T \Sigma} \Sigma^T U^T$$

$$(\Sigma^T \Sigma)^{-1} = \begin{bmatrix} \frac{1}{\sigma_0^2} & 0 \\ 0 & \frac{1}{\sigma_1^2} \end{bmatrix}$$

$$\Sigma^T U^T = \begin{bmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_1 & 0 \end{bmatrix} \begin{bmatrix} u_{00} & u_{10} & u_{20} \\ u_{01} & u_{11} & u_{21} \\ u_{02} & u_{12} & u_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_0 u_0^T \\ \sigma_1 u_1^T \end{bmatrix} \in \mathbb{R}^{2 \times 3}$$

u_2 gone?

$$U\Sigma = \begin{bmatrix} \sigma_0 u_0 & \sigma_1 u_1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

$$U\Sigma \; \frac{1}{\Sigma^T\Sigma} \Sigma^T U^T$$

$$= \begin{bmatrix} u_0 & u_1 \end{bmatrix} \begin{bmatrix} u_0^T \\ u_1^T \end{bmatrix}$$

Note outer product of two orthogonal vectors

$$u_0 u_1^T = 0 \, ?$$

Final product

$$\tilde{y} = \left(u_0 u_0^T + u_1 u_1^T \right) y$$

in general

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

only terms with $\sigma_j > 0$ survive ?

For Ridge, rewrite matrix inversion as

$$V\Sigma^T\Sigma V^T + \lambda \mathbb{I}$$

$$= V\Sigma^T\Sigma V^T + \lambda V V^T \mathbb{I}$$

$([V, I] = 0$ commuter$)$

$$= V \underbrace{\left(\Sigma^T\Sigma + \lambda I\right)}_{\text{invertible always}} V^T$$

proceed as with OLS and get

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