HW - week 11 R10946017 許鴻君 12/11

1. Computing the least square estimate of SVD.

Since $X = U\Lambda V^T$

$$\Rightarrow X^TX = (U\Lambda V^T)^T(U\Lambda V^T) = V\Lambda^T \cancel{U^T} \cancel{U} \Lambda V^T$$

$$= V\Lambda^T\Lambda V^T$$

$$(X^TX)^{-1} = (V\Lambda^T\Lambda V^T)^{-1} = (V^T)^{-1}(\Lambda^T\Lambda)^{-1}V^{-1}$$

$$= V\underbrace{(\Lambda^T\Lambda)^{-1}}_{=D}V^T \quad (V \text{ is orthogonal matrix}) \quad V^T = V^{-1}$$

Therefore, b is correct.

HW-week 11 R10946017 許鴻君 12/11

2. The ridge regression estimate

reformulate the given problem to a simple regression problem.

Let's consider $\|\vec{y}-X\beta\|^2+\lambda\|W\beta\|^2$ first,

$$\|\vec{y}-X\beta\|^2+\lambda\|W\beta\|^2=\vec{y}^T\vec{y}-2\beta^TX^T\vec{y}+\beta^TX^TX\beta+\lambda\beta^TW^TW\beta$$

combine

$$=\vec{y}^T\vec{y}-2\beta^TX^T\vec{y}+\beta^TX^TX\beta+\beta^T(\sqrt{\lambda}W)^T(\sqrt{\lambda}W)\beta$$

$$=\vec{y}^T\vec{y}-2\beta^TX^T\vec{y}+\beta^T[X^T,-\sqrt{\lambda}W^T]\begin{bmatrix}X\\-\sqrt{\lambda}W\end{bmatrix}\beta$$

$$=\vec{y}^T\vec{y}-2\beta^T[X^T,-\sqrt{\lambda}W^T]\begin{bmatrix}\vec{y}_{n\times1}\\\vec{0}_{p\times1}\end{bmatrix}+\beta^T[X^T,-\sqrt{\lambda}W^T]\begin{bmatrix}X\\-\sqrt{\lambda}W\end{bmatrix}\beta$$

$$=[\vec{y},\vec{0}]\underbrace{\begin{bmatrix}\vec{y}_{n\times1}\\\vec{0}_{p\times1}\end{bmatrix}}_{\text{Let }\vec{v}_{(n+p)\times1}}-2\beta^T[X^T,-\sqrt{\lambda}W^T]\begin{bmatrix}\vec{y}_{n\times1}\\\vec{0}_{p\times1}\end{bmatrix}+\beta^T[X^T,-\sqrt{\lambda}W^T]\underbrace{\begin{bmatrix}X\\-\sqrt{\lambda}W\end{bmatrix}}_{\text{Let }M_{(n+p)\times p}}\beta$$

$$=\vec{v}^T\vec{v}-2\beta^TM^T\vec{v}+\beta^TM^TM\beta=\|\vec{v}-M\beta\|^2$$

Therefore, we can simply the given problem to:

$$\hat{\beta}^{ridge}=\arg\min_{\beta}\left\{\frac{1}{2}\|\vec{v}-M\beta\|_2^2\right\}$$

with $\vec{v}_{(n+p)\times1}=\begin{bmatrix}\vec{y}_{n\times1}\\\vec{0}_{p\times1}\end{bmatrix}$ & $M_{(n+p)\times p}=\begin{bmatrix}X_{(n\times p)}\\-\sqrt{\lambda}W_{(p\times p)}\end{bmatrix}$