Machine Learning Homework 01

Qiu Yihang

Mar.27, 2022

Proof. We use I_p to denote p-dimensional unit matrix, i.e. $I_p = \text{diag}(1, 1, ..., 1)_{p \times p}$. Since V is a $p \times p$ orthogonal matrix, we know $V^T V = V V^T = I_p$. Thus, we have

$$\begin{aligned} \left(\boldsymbol{Z}^{T}\boldsymbol{Z} + \lambda\boldsymbol{I}_{p}\right)^{-1}\boldsymbol{Z}^{T}\boldsymbol{y} &= \left(\left(\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^{T}\right)^{T}\left(\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^{T}\right) + \lambda\boldsymbol{V}^{T}\boldsymbol{V}\right)^{-1}\boldsymbol{Z}^{T}\boldsymbol{y} \\ &= \left(\boldsymbol{V}\boldsymbol{D}^{T}\boldsymbol{U}^{T}\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^{T} + \lambda\boldsymbol{V}^{T}\boldsymbol{I}_{p}\boldsymbol{V}\right)^{-1}\boldsymbol{Z}^{T}\boldsymbol{y} \\ &= \left(\boldsymbol{V}\boldsymbol{D}^{T}\boldsymbol{D}\boldsymbol{V}^{T} + \boldsymbol{V}^{T}\left(\lambda\boldsymbol{I}_{p}\right)\boldsymbol{V}\right)^{-1}\boldsymbol{Z}^{T}\boldsymbol{y} \\ &= \left(\boldsymbol{V}\left(\boldsymbol{D}^{2} + \lambda\boldsymbol{I}_{p}\right)\boldsymbol{V}^{T}\right)^{-1}\boldsymbol{Z}^{T}\boldsymbol{y} \end{aligned}$$

Meanwhile,

$$egin{aligned} \left(oldsymbol{V}\left(oldsymbol{D}^2+\lambdaoldsymbol{I}_p
ight)oldsymbol{V}^T
ight)oldsymbol{V} ext{diag}\left(rac{1}{d_j^2+\lambda}
ight)oldsymbol{V}^T
ight) \\ &=oldsymbol{V} ext{diag}\left(d_j^2+\lambda
ight)\left(oldsymbol{V}^Toldsymbol{V}
ight) ext{diag}\left(rac{1}{d_j^2+\lambda}
ight)oldsymbol{V}^T &=oldsymbol{V}oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbol{V}^T=oldsymbol{V}^T=oldsymbol{V}^T=oldsymbol{V}^T=oldsymbol{I}_poldsymbol{V}^T=oldsymbo$$

Therefore,

$$egin{aligned} \left(oldsymbol{Z}^Toldsymbol{Z} + \lambdaoldsymbol{I}_p
ight)^{-1}oldsymbol{Z}^Toldsymbol{y} &= oldsymbol{V} \left(oldsymbol{D}^2 + \lambdaoldsymbol{I}_p
ight)oldsymbol{V}^Toldsymbol{Z}^Toldsymbol{y} \ &= oldsymbol{V} ext{diag}\left(rac{1}{d_j^2 + \lambda}
ight)oldsymbol{V}^Toldsymbol{U}oldsymbol{U}^Toldsymbol{Y} \end{aligned} egin{aligned} &= oldsymbol{V} ext{diag}\left(rac{1}{d_j^2 + \lambda}
ight)oldsymbol{V}^Toldsymbol{V}oldsymbol{D}^Toldsymbol{U}^Toldsymbol{y} \end{aligned} \ &= oldsymbol{V} ext{diag}\left(rac{1}{d_j^2 + \lambda}
ight) ext{diag}\left(d_ioldsymbol{U}^Toldsymbol{y} \end{aligned} \ &= oldsymbol{V} ext{diag}\left(rac{d_j}{d_j^2 + \lambda}
ight)oldsymbol{U}^Toldsymbol{y} \end{aligned}$$

Qed.

1