Tutorial & Practical 3: Ridge Regression

Question 1

Given the model

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$ is full rank p and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$. Let $\hat{\boldsymbol{\beta}}$ be the estimate of $\boldsymbol{\beta}$ obtained by least square estimation and let $X = S\Sigma Q^\top$, where S is an $n \times n$ matrix, $\Sigma = \text{diag}(\sigma_1, ..., \sigma_r)$ is a diagonal matrix and Q is $p \times p$ matrix, be the singular value decomposition of X

1. Using the singular value decomposition of X, write the least square estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ as a sum of individual components b_i along $\mathbf{s}_i$, the columns vectors of S
2. Derive the expression of the ridge regression estimator and write its expression using b_i
3. Provide the form of the shrinkage functions of the singular values σ_i that characterize the ridge repression and principal component regression estimators
4. Derive the bias of the ridge regression estimator
5. Derive the variance covariance of the ridge regression estimator and compare the obtained variance with the variance covariance of the least square estimator
6. Provide the distribution of the ridge regression estimator
7. Derive the expression of the mean square error of the ridge regression estimator
8. Derive the expressions of the ridge regression estimator, its bias, variance, mean square error in the case of orthonormal design matrix X
9. Obtain the value of the regularization parameter that minimizes the mean square error in the case of orthonormal design matrix X
10. Derive the expression of the square norm of the ridge regression estimator and discuss its behavior when the regularization parameter is decreasing

Question 2

Let

$$\mathbf{y} = \mathbf{1}_n\beta + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$ is full rank p and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$.

1. Derive the degrees of freedom of the least square estimator
2. Provide its expression in the case of orthonormal design matrix X
3. Derive the degrees of freedom of the ridge regression estimator
4. Provide its expression in the case of orthonormal design matrix X