

1. We have two absolutely identical preliminary standardized regressors x and x . The dependent variable y is centered.

In the ridge regression one minimizes the loss function

$$\text{loss}(\hat{\beta}) = (y - \hat{y})^T (y - \hat{y}) + \lambda \hat{\beta}^T \hat{\beta}, \quad \hat{y} = \hat{\beta}_1 x + \hat{\beta}_2 x.$$

- (a) Find the optimal $\hat{\beta}_1$ and $\hat{\beta}_2$ for fixed λ .
 - (b) What happens to the estimates when $\lambda \rightarrow \infty$?
 - (c) What happens to the sum $\hat{\beta}_1 + \hat{\beta}_2$ when $\lambda \rightarrow 0$?
2. Consider the model $y = X\beta + u$ where β is non-random, $\mathbb{E}(u | X) = 0$, the matrix X of size $n \times k$ has rank $X = k$, but $\text{Var}(u | X) = \sigma^2 W$ with $W \neq I$. Let $\hat{\beta}$ be the standard OLS estimator of β .
 - (a) Find $\mathbb{E}(\hat{\beta} | X)$, $\mathbb{E}(\hat{\beta})$.
 - (b) Find $\text{Var}(\hat{\beta} | X)$.
 - (c) How do you think, will the standard confidence interval for β be valid in this case?
 - (d) Find $\text{Cov}(y, \hat{\beta} | X)$.
 3. Consider the matrix

$$X = \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}.$$

- (a) Find the matrix $X^T X$ and diagonalize it.
- (b) Find the SVD of X .
- (c) Find the best approximation to X with rank equal to 1.

Remark: in the principal component analysis the variables in the matrix X should be standardized. If you can't do this by bare hands, feel free to use python, but provide code!

4. The columns of X are standardized. You know the SVD of the matrix $X = UDV^T$. The diagonal elements of D are positive and ordered from highest to lowest, $d_{11} > d_{22} > \dots > 0$.

Let's maximize $\|Xw\|^2$ by choosing an optimal vector w subject to $\|w\|^2 = 1$.

- (d) Write the Lagrangian function for this problem.
- (e) Find the first order conditions. Differential is your friend!
- (f) Find the optimal w in terms of columns of V .

Hint: one may interpret the FOC in terms of eigenvalues and eigenvectors!