STAT 542: Homework 1

Due: Feb. 10 midnight on Canvas

Please make sure that your solutions are readable and the file size is reasonable. Typing the answers is highly encouraged.

Problem 1.

Suppose that the true observation model is given by

$$Y = X\beta + \epsilon \tag{1}$$

where $X \in \mathbb{R}^{n\times p}$, and ϵ satisfies $\mathbb{E}[\epsilon] = 0$ and $\mathbb{E}[\epsilon\epsilon^\top] = \sigma^2 I$. Further assume that the $X_1, X_2 \in \mathbb{R}^n$ are the two columns of X, $\|X_1\|_2 = \|X_2\|_2 = 1$, and the inner product $\langle X_1, X_2\rangle = r$. Denote by

$$\hat{\beta} := (X^\top X)^{-1} X^\top Y \tag{2}$$

and OLS estimator using the full model, and

$$\hat{\beta}^{\mathsf{r}} := (X_1^\top X_1)^{-1} X_1^\top Y \tag{3}$$

the OLS estimator using the reduced model.

- [1pts] Suppose that we are only interested in estimating the first coordinate, β_1. Compute $\mathbb{E}[\hat{\beta}_1]$ and $\mathrm{var}(\hat{\beta}_1)$ (express the answers using β, σ and r).
- [2pts] Compute $\mathbb{E}[\hat{\beta}_1^{\mathsf{r}}]$ and $\mathrm{var}(\hat{\beta}_1^{\mathsf{r}})$.
- [2pts] Use the bias-variance tradeoffs to compute the mean square errors of $\hat{\beta}_1$ and $\hat{\beta}_1^{\mathsf{r}}$. Show that the reduced model has a smaller mean square error when

$$\frac{|\beta_2|}{\sigma} < \frac{1}{\sqrt{1-r^2}}. \tag{4}$$

Problem 2.

[1pts] Use R or Python to perform the following experiment: you pick arbitrary numbers $\rho \in (0,1)$ and $r > 0$ satisfying

$$\frac{6\rho}{1+\rho^2} > \frac{1}{r} + 2r. \tag{5}$$

Set $X = \begin{pmatrix} 1 & \rho r \\ \rho & r \end{pmatrix}$ and $Y = X \begin{pmatrix} -1 \\ 2 \end{pmatrix}$. For any $\lambda > 0$, define

$$\hat{\beta}_\lambda := \arg\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{2} \|Y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\}. \tag{6}$$

Plot the coefficients of $\hat{\beta}_\lambda$ as a function of $\|\hat{\beta}_\lambda\|_1$, and check if $\|\hat{\beta}_\lambda\|_0$ is not a monotonic function of the ℓ_1 norm.

Hint: `lasso.R` *in Canvas contains most of the ingredients of the code.*

Problem 3.

[1pts] Repeat the experiment in Problem 2, but choose ρ and r such that the inequality sign in (5) is reversed. Check if $\|\hat{\beta}_\lambda\|_0$ is monotonic in $\|\hat{\beta}_\lambda\|_1$ this time.

Problem 4.

[2 bonus pts.] verify that (5) is equivalent to the condition that "Line PT has a smaller slope (in absolute value) than PO" in `Notes Lasso.pdf`.