BIOSTAT 274 Spring 2021 Homework 1

Due 11:59 PM 04/21/2020, submit to CCLE

Remark. For **Theoretical Part**, you can summit a Scanned or Typed copy.

Theoretical Part

1. (*20 pt*) This problem illustrates the estimator property in the shrinkage methods. Let Y be a single observation, and regress Y on an intercept

$$Y = 1 \cdot \beta + \epsilon.$$

   (a) Using the formulation as shown in class, write down the optimization problem of general linear model, ridge regression and LASSO in estimating β respectively.

   (b) For fixed tuning parameter λ, solve for $\hat{\beta}$ (general linear model), $\hat{\beta}^R_\lambda$ (ridge regression) and $\hat{\beta}^L_\lambda$ (LASSO) respectively.

   Hint. For the LASSO problem, show that the objective function is convex even though not everywhere differentiable, hence any local minima is also a global minima. Argue that it suffices to solve for $\beta \geqslant 0$ and $\beta < 0$ respectively. Carefully discuss all possible situations (where the minima locates and what's the optimal value in both cases), and then pick the better one as $\hat{\beta}^L_\lambda$.

   (c) Represent $\hat{\beta}^R_\lambda$ and $\hat{\beta}^L_\lambda$ by $\hat{\beta}$ and create plots of them separately for $\lambda = 1, 5, 10$. What can you tell?

2. ([ISL] 6.5, *15 + 10 pt*) It is well-known that ridge regression tends to give similar coefficient values to correlated/collinear variables, whereas the LASSO may give quite different coefficient values to correlated/collinear variables. We will now explore this property in a very simple setting. Suppose that we have two observations (x_1, y_1) and (x_2, y_2), where $x_1 \neq 0$, $x_1 + x_2 = y_1 + y_2 = 0$. Consider the linear model y_i artificially regressed on (x_i, x_i) without intercept:

$$\begin{cases} y_1 = x_1\beta_1 + x_1\beta_2 + \epsilon_1 \\ y_2 = x_2\beta_1 + x_2\beta_2 + \epsilon_2 \end{cases}$$

   (a) Write out the ridge regression optimization problem in this setting.

   (b) Argue that the ridge coefficient estimates satisfy $\hat{\beta}^R_{\lambda,1} = \hat{\beta}^R_{\lambda,2}$.

   (c) Write out the LASSO optimization problem in this setting.

(d) (*Optional*) Argue that in this setting, the LASSO coefficients $\hat{\beta}^L_{\lambda,1}$ and $\hat{\beta}^L_{\lambda,2}$ are not unique. Describe these solutions.

Hint. Starting with an optimal coefficients $(\hat{\beta}^L_{\lambda,1}, \hat{\beta}^L_{\lambda,2})$, indicate that you can find another one. Use the relationship of a usual LASSO problem and its constrained version. Investigate into the relationship between the contour of the objective function and the constraint set in the constrained version.