Statistical Methods for Bioinformatics

Exercises

Exercises 6.8.5

5. It is well-known that ridge regression tends to give similar coefficient values to correlated variables, whereas the lasso may give quite different coefficient values to correlated variables. We will now explore this property in a very simple setting.

   Suppose that $n = 2$, $p = 2$, $x_{11} = x_{12}$, $x_{21} = x_{22}$. Furthermore, suppose that $y_1 + y_2 = 0$ and $x_{11} + x_{21} = 0$ and $x_{12} + x_{22} = 0$, so that the estimate for the intercept in a least squares, ridge regression, or lasso model is zero: $\hat{\beta}_0 = 0$.

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

   (b) Argue that in this setting, the ridge coefficient estimates satisfy $\hat{\beta}_1 = \hat{\beta}_2$.

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

   (d) Argue that in this setting, the lasso coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ are not unique—in other words, there are many possible solutions to the optimization problem in (c). Describe these solutions.

- Ordinary least squares with the intercept at 0:

$$RSS = \sum_{i=1}^{n}(y_i\text{-}\sum_{j=1}^{p}\beta_j x_{ij})^2$$

for 2 points: $RSS = (y_1\text{-}\beta_1 x_{11} - \beta_2 x_{12})^2 + (y_2\text{-}\beta_1 x_{21} - \beta_2 x_{22})^2$ with the equality conditions $x_{11} = x_{12} = x_1$ and $x_{21} = x_{22} = x_2$ and with $y1 = -y2$ and $x_1 = -x_2$

- Ridge regression minimizes:

$$RSS + \lambda(\beta_1^2 + \beta_2^2)$$

- Lasso minimizes:

$$RSS + \lambda(|\beta_1| + |\beta_2|)$$

- After simplification we have:

$(y - \beta_1 x - \beta_2 x)^2 + (-y + \beta_1 x + \beta_2 x)^2 + \lambda(\beta_1^2 + \beta_2^2)$
$2\beta_1^2 x^2 + 4\beta_1\beta_2 x^2 - 4\beta_1 xy + 2\beta_2^2 x^2 - 4\beta_2 xy + 2y^2$

- To minimize the expression, we take the first derivative of the function for the coefficients in turn, and solve for 0:

for β_1

$$-4xy + 4x^2\beta_1 + 4x^2\beta_2 + 2\beta_1\lambda = 0$$

for β_2

$$-4xy + 4x^2\beta_1 + 4x^2\beta_2 + 2\beta_2\lambda = 0$$

- To solve this system of equations multiply one of the two with -1 and add to the other (multiplication and addition scheme to solve a system of linear equations, Gaussian elimination) and find: $2\lambda\beta_1 - 2\lambda\beta_2 = 0$
- Hence both coefficients are the same.

Exercise 6.8.5d

- For Lasso we have:

$$(y - \beta_1 x - \beta_2 x)^2 + (-y + \beta_1 x + \beta_2 x)^2 + \lambda(|\beta_1| + |\beta_2|)$$

- The derivative for $\lambda |\beta_1|$ is $\frac{\lambda \beta_1}{|\beta_1|}$, for simplicity let's assume x and y are positive and that both coefficients are hence positive as well (or vice versa, it matters not for the solution). Then we end up when differentiating and solving to 0 for either coefficient:

$$\beta_1 + \beta_2 = \frac{y}{x} + \frac{\lambda}{4x^2}$$

- The system is underdefined and there are a set of solutions. Specifically, for any set of Y, X and λ values there will be the following set of solutions: $\beta_1 + \beta_2 = c$
- This describes the full edge of the lambda diamond solution space.