

**Problem Set 8: Solutions****Part One: Hand-Written Exercise**

1.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n(\sum_{i=1}^n x_i^2 + \lambda) - (\sum_{i=1}^n x_i)^2}. \blacksquare$$

2. Let

$$Q := (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_R)'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_R) + \lambda \boldsymbol{\beta}_R' \boldsymbol{\beta}_R$$

The F.O.C. is

$$-2\mathbf{X}'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_R) + 2\lambda \boldsymbol{\beta}_R = \mathbf{0}$$

$$\Rightarrow (\mathbf{X}'\mathbf{X} + \lambda \mathbf{I})\boldsymbol{\beta}_R = \mathbf{X}'\mathbf{y}$$

$$\Rightarrow \hat{\boldsymbol{\beta}}_R = (\mathbf{X}'\mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}'\mathbf{y} = \left( \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i' + \lambda \mathbf{I} \right)^{-1} \left( \sum_{i=1}^n \mathbf{x}_i y_i \right). \blacksquare$$

3. (a) Best subset. Because best subset approach is to find the smallest training RSS model among all combinations given k predictors.
- (b) Not sure. It depends on the data testing the model.
- (c) True. For forward stepwise selection, once a predictor is included in the model, it will never be kicked out for the next step or thereafter.
- (d) True. Because it is to remove a predictor in the (k+1)-predictors model to become the k-predictors model.  $\blacksquare$