Homework No. 04

Deadline – Sunday, April 14

1. Generate a random predictor X *from a normal distribution* of length $n = 100$, as well as a noise vector e of length $n = 100$ (**score = 5**). Generate a response vector Y of length $n = 100$ according to the model

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + e,$$

   where $\beta_0, \beta_1, \beta_2, \beta_3$ are constants of your choice (**score = 5**).

2. Fit the **Ridge** model to the simulated data, using $X, X^2, \ldots, X^{10}$ as predictors (**score =10**). Use cross-validation (10-fold) to select the optimal value of λ (**score =10**). Create plots of the cross-validation error as a function of λ (**score = 10**). Report the resulting coefficient estimates and discuss the results (**score = 10**).

3. Fit the **Lasso** model to the simulated data, using $X, X^2, \ldots, X^{10}$ as predictors (**score =10**). Use cross-validation (10-fold) to select the optimal value of λ (**score =10**). Create plots of the cross-validation error as a function of λ (**score = 10**). Report the resulting coefficient estimates and discuss the results (**score = 10**).

4. Compare the results (**score = 10**).