MATH 569 MIDTERM EXAM

Part II: Take-home exam

3/10, Friday 10:00am - 3/11, Saturday 10:00am

You must complete the work all by yourself without assistance from other persons within or outside of this class.

Submission instruction:

- For problem 3, submit the source code, the report with data and plots, and any comments you feel necessary to understand your code. Put all files in one folder and upload the zipped file to Blackboard.
- The submission link will disappear 15 minutes after the deadline. We will not accept submissions via email or other means.

Problem 1 (10 pts) Derive equation (5.55) in the book, using equation (5.54) as the starting point.

Problem 2 (10 pts) Let $\hat{\boldsymbol{\beta}}$ be the least square estimate of parameters $\boldsymbol{\beta}$ in the linear regression model. Consider estimation of any linear combination of the parameters $\theta = \boldsymbol{\alpha}^T\boldsymbol{\beta}$. Prove that the MSE from the least square estimator $\boldsymbol{\alpha}^T\hat{\boldsymbol{\beta}}$ is no greater than the MSE from any unbiased estimator $\tilde{\theta} = \boldsymbol{c}^T\boldsymbol{y}$. Show your full derivation.

Problem 3 (30 pts) Consider the prostate cancer data. Implement the following two methods to fit the data in the training set.

- Method I, ridge regression
- Method II, generalized ridge regression with regularization using reproducing Gaussian kernel: $K(\boldsymbol{x}, \boldsymbol{y}) = e^{-\nu\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}$, where $\nu = 1$.

The dataset has eight features: (*lcavol, lweight, age, lbph, svi, lcp, gleason, pgg45*). The last column (*lpsa*) is the response.

Note that for method I, you need to center and scale the features (see book page 64), but for method II, you do not need to.

(1) Report the mean squared error (formula: $\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2$) for the testing dataset at $\lambda = \{0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$ for both methods.

λ	0	10	20	30	40	50	60	70	80	90	100
method I											
method II											

(2) Plot the regression coefficients ($\hat{\beta}$ for method I and $\hat{\alpha}$ for method II) versus the effective degrees of freedom (as shown in Figure 3.8 of the book). Note that in order to fully observe the shrinkage effect, you need $\lambda = \{0, \ldots, 10000\}$ for method I, and $\lambda = \{0, \ldots, 100\}$ for method II.

You are required to implement method I and method II using the equations learned from this class. No credit will be given if you use library functions to substitute them.