University of Washington

Stat / Biostat 527 Nonparametric Regression and Classification

Homework 2

Issued: Thursday, April 11, 2013 **Due:** Thursday, April 18, 2013

The data "Prostate.csv" were collected on n = 97 men before radical prostatectomy. We take as response, Y, the log of prostate specific antigen (PSA); PSA is a concentration and is measured in ng/ml. We want to examine $\log(PSA)$ as a function of age and seven other histological and morphometric covariates by ridge regression and Lasso.

Problem 2.1

Ridge regression has closed form solutions, as provided in Equation (10.23) of "Bayesian and Frequentist Regression Methods". Write your own code to get the ridge regression parameter estimates for the PSA data. Specifically, set the smoothing parameter λ from 0 to 10000 by a 100 difference. For each given λ, store the estimated regression parameters and corresponding effective degrees of freedom.

(a) Plot the parameter estimates associated with the eight covariates versus $\mathbf{log(\boldsymbol{\lambda})}$. Comment on their relationship.

(b) Plot the parameter estimates associated with the eight covariates versus the **effective degrees of freedom**. Comment on their relationship.

Hint: To avoid penalizing intercept, you could first center y and all covariates. The intercept $\hat{\beta}_0$ can be estimated at the end by $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}_1 - \cdots - \hat{\beta}_p\bar{x}_p$. In this problem, you only need to report the eight parameters associated with the eight covariates. After centering x and y, the model doesn't need to include intercept, since the fitted line, plane or hyperplane should pass the mean of x and mean of y, which are zero after centering.
You may also want to scale the covariates to have unit variance in order to get parameters of similar magnitude.

Problem 2.2

Write your own code to implement the "shooting algorithm" for Lasso with the PSA data.

Hint: Again to avoid penalizing intercept and to get parameters of similar magnitude, you need to standardize all covariates and center y.

(a) Run the procedure for 1000 iterations for λ from 0.01 to 200 by a 1 difference (one iteration means randomly choosing a coordinate j and update β_j). Treat the parameter estimates after 1000 iterations as the final parameter estimates.

For each given λ, store the final parameter estimates and its corresponding shrinkage factor. The shrinkage factor is defined as $\sum_{j=1}^{p} |\beta_j^{Lasso}| / \sum_{j=1}^{p} |\beta_j^{LS}|$. When the shrinkage factor is 1, the lasso estimates are the same as the least squares estimates, which corresponds to zero penalization on parameters, i.e. $\lambda = 0$. When the shrinkage factor is 0, the lasso estimates are all zero, which corresponds to $\lambda = \infty$.

(i) Plot the final parameter estimates associated with the eight covariates versus $\mathbf{log(\boldsymbol{\lambda})}$ and describe their relationship. Compare your results with the one obtained in Problem 2.1 (a).

(ii) Plot the final parameter estimates associated with the eight covariates versus the **shrinkage factor** and describe their relationship.

Hint: You could check your results with outputs from built-in packages. Please see the notes of Recitation 1 for using built-in packages in `R`.

(b) Suggest a convergence criterion. Run your shooting algorithm again with your convergence criterion.

(i) Plot the final parameter estimates associated with the eight covariates versus the **shrinkage factor**. Compare your results with the one obtained in Problem 2.2 (a) part (ii).

(ii) Plot the number of iterations it takes to converge versus the **shrinkage factor**.