IEOR 165 – Course Project
Due Wednesday, May 6, 2020

Instructions:

The course project must be submitted on bCourses as a PDF file. You are allowed to consult and discuss with classmates and others, but each student must submit their own project writeup and code. The project will be graded on the basis of the quality of the modeling approach. You can use whichever software and libraries/packages you would like, and are not expected to implement statistical estimation algorithms yourself.

Project Tasks:

The authors of the following research paper:

Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, "Modeling wine preferences by data mining from physicochemical properties", *Decision Support Systems*, vol. 47, no. 4:547-553, 2009.

considered the problem of modeling wine preferences. Wine can be evaluated by experts who give a subjective score, and the question the authors of this paper considered was how to build a model that relates objective features of the wine (e.g., pH values) to its rated quality. For this project, we will use the data set available at:
http://courses.ieor.berkeley.edu/ieor165/homeworks/winequality-red.csv

Use the following methods to identify the coefficients of a linear model relating wine quality to different features of the wine: (1) ordinary least squares (OLS), (2) ridge regression (RR), (3) lasso regression, (4) elastic net. Make sure to include a constant (intercept) term in your model, and choose the tuning parameters using cross-validation. You may use any programming language you would like to. For your solutions, please include (i) plots of tuning parameters versus cross-validation error, (ii) tables of coefficients (labeled by the feature) computed by each method, (iii) the minimum cross-validation error for each method, and (iv) the source code used to generate the plots and coefficients. Some hints are below:

- a constant (intercept) term can be included in OLS by solving
$$\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}\end{bmatrix} = \arg\min_{\beta_0,\beta} \left\| Y - \begin{bmatrix}\mathbf{1}_n & X\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta\end{bmatrix}\right\|_2^2$$
- RR and lasso have one tuning parameter, while elastic net has two tuning parameters
- RR (with an intercept term) can be formulated as
$$\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}\end{bmatrix} = \arg\min_{\beta_0,\beta} \left\| \begin{bmatrix}Y\\ 0\end{bmatrix} - \begin{bmatrix}\mathbf{1}_n & X\\ 0 & \mu\cdot\mathbb{I}\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta\end{bmatrix}\right\|_2^2,$$
where μ is a tuning parameter.