IEOR 290A – Homework 1
Due Wednesday, March 19, 2014 in class

In the paper: P. 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, the authors 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 homework, we will use the data set available at:
http://ieor.berkeley.edu/~aaswani/sp14_ieor290a/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) exterior derivative estimation (EDE) estimator. 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) 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
$$\begin{bmatrix}\hat{\beta}_0 \\ \hat{\beta}\end{bmatrix} = \arg\min_{\beta_0,\beta} \left\| Y - \begin{bmatrix}\mathbb{1}_n & X\end{bmatrix} \begin{bmatrix}\beta_0 \\ \beta\end{bmatrix} \right\|_2^2$$
- RR, lasso have one tuning parameter, while EDE has two tuning parameters
- RR (with an intercept term) can be formulated as
$$\begin{bmatrix}\hat{\beta}_0 \\ \hat{\beta}\end{bmatrix} = \arg\min_{\beta_0,\beta} \left\| \begin{bmatrix}Y \\ 0\end{bmatrix} - \begin{bmatrix}\mathbb{1}_n & X \\ 0 & \mu \cdot \mathbb{I}\end{bmatrix} \begin{bmatrix}\beta_0 \\ \beta\end{bmatrix} \right\|_2^2,$$
where μ is a tuning parameter.
- EDE (with an intercept term) can be formulated as
$$\begin{bmatrix}\hat{\beta}_0 \\ \hat{\beta}\end{bmatrix} = \arg\min_{\beta_0,\beta} \left\| \begin{bmatrix}Y \\ 0\end{bmatrix} - \begin{bmatrix}\mathbb{1}_n & X \\ 0 & \mu \cdot \Pi\end{bmatrix} \begin{bmatrix}\beta_0 \\ \beta\end{bmatrix} \right\|_2^2.$$
Here, μ is a tuning parameter and $\Pi = V_2 V_2'$, where the singular value decomposition (SVD) of X is
$$X = US\begin{bmatrix}V_1 & V_2\end{bmatrix}'$$
with $V_1 \in \mathbb{R}^{p\times d}$, $V_2 \in \mathbb{R}^{p\times(p-d)}$, and the singular values in S are listed in decreasing order.