Notes on Non-Negative Garrote

Method Overview

Non-Negative Garrote is a regularization method that allows both shrinking and elimination of weights. It involves first calculating the OLS solution as a starting value for the coefficients and then using a constraint parameter to scale the the coefficients.

The original equation published by Leo Breiman in 1995:

$$\sum_k (y_n - \sum_k c_k \hat{\beta}_k x_{kn})^2$$

Constrained by

$$c_k \geq 0 \quad \sum_k c_k \leq s$$

Where $\hat{\beta}_k$ is the starting OLS estimate.

Our Data

Non-Negative Garrote can be viewed as a scaled version of the OLS estimate.

First, Non-Negative Garrote didn't perform much better than OLS. In all cases, the mean squared error recorded for both OLS and Non-Negative Garrote were negligibly different.

We ran OLS linear regression using a single feature as input and a single feature as output in order to determine the relationship between the two variables. What we found was that the data was very collinear and every feature strongly correlated. This makes sense because the four features have similar ranges. Open, close, high, low are all measures of stock price.

Also there is now a text file at acropolis/analysis/results/r_squared.txt. It contains matrices for the R^2 value for one feature as input and another feature as output. This effectively helps generalize how related each pair of feature

is. As the data shows, each feature is highly correlated with each other which explains why Non-Negative Garrote didn't produce less error than OLS.

I got the equations from a copy of the original paper posted here:

`http://www-personal.umich.edu/~jizhu/jizhu/wuke/Breiman-Technometrics95.pdf`

Also these two papers were helpful:

`http://link.springer.com/chapter/10.1007%2F978-3-642-25832-9_9`
`http://www2.isye.gatech.edu/statistics/papers/05-25.pdf`