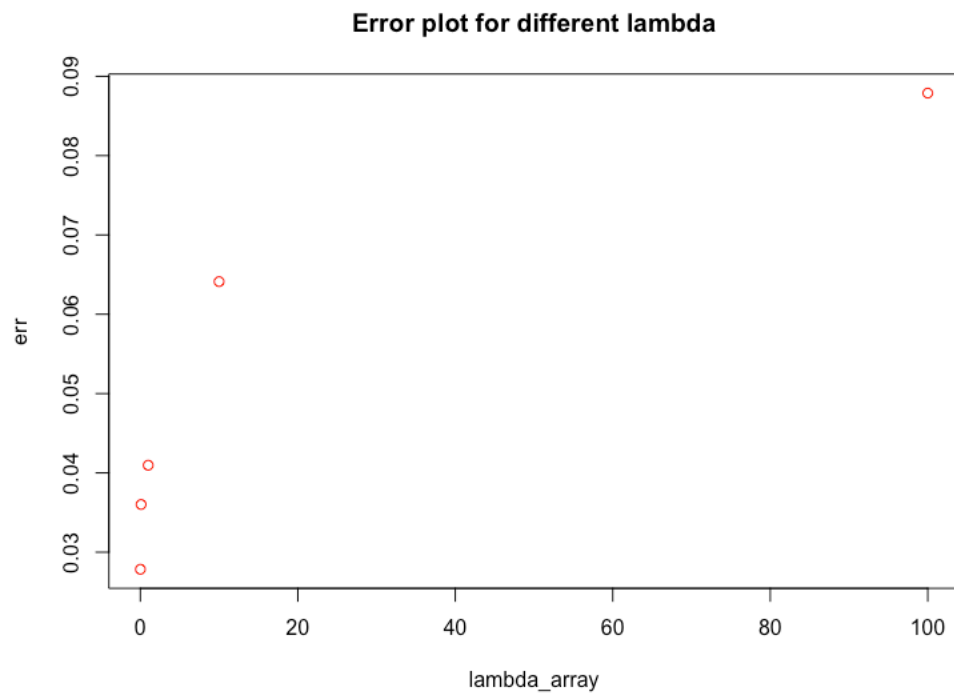# STATS 202A: Statistics Programming

Chaojie Feng

1. Ridge Regression and Linear Spline

Spline Test

Error plot for different lambda
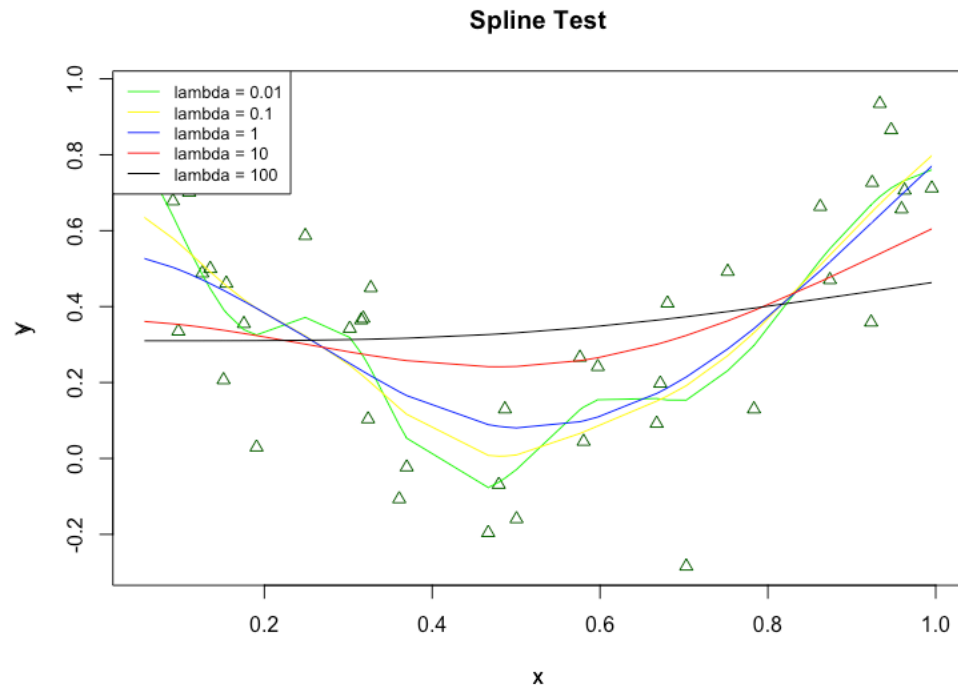
The 2-dimensional function I used to test is $y = 4(x - 0.5)^2 + \mathcal{N}(u_z, \sigma_z)$. From the spline test, we can clearly see the effect of overfitting and underfitting. As regularization is very small (lambda = 0.001), the spline curve tries to fit as many data points as possible to reduce training error so that it causes overfitting. As regularization term grows very large, the spline curve will not only throw away the noise information, but also throws away the core information from data, which causes underfitting.

2. Comparison between Lasso and epsilon-boosting solution path

Lasso solution path

Epsilon-boosting solution path
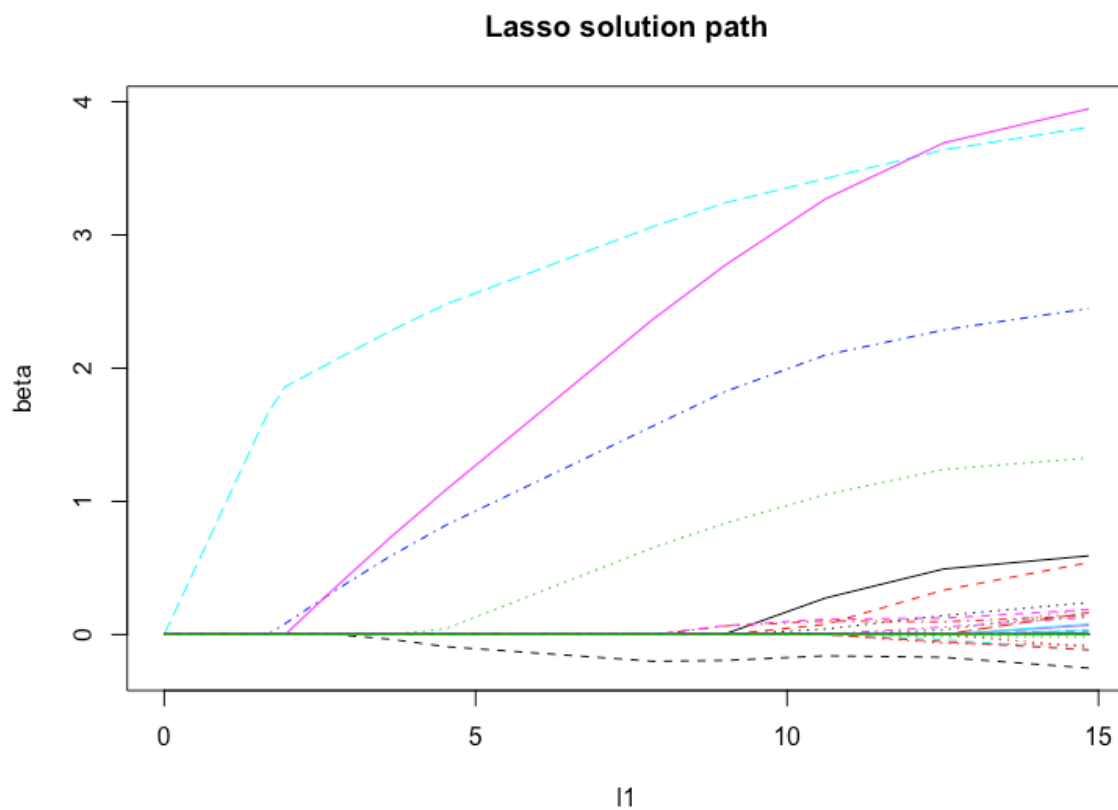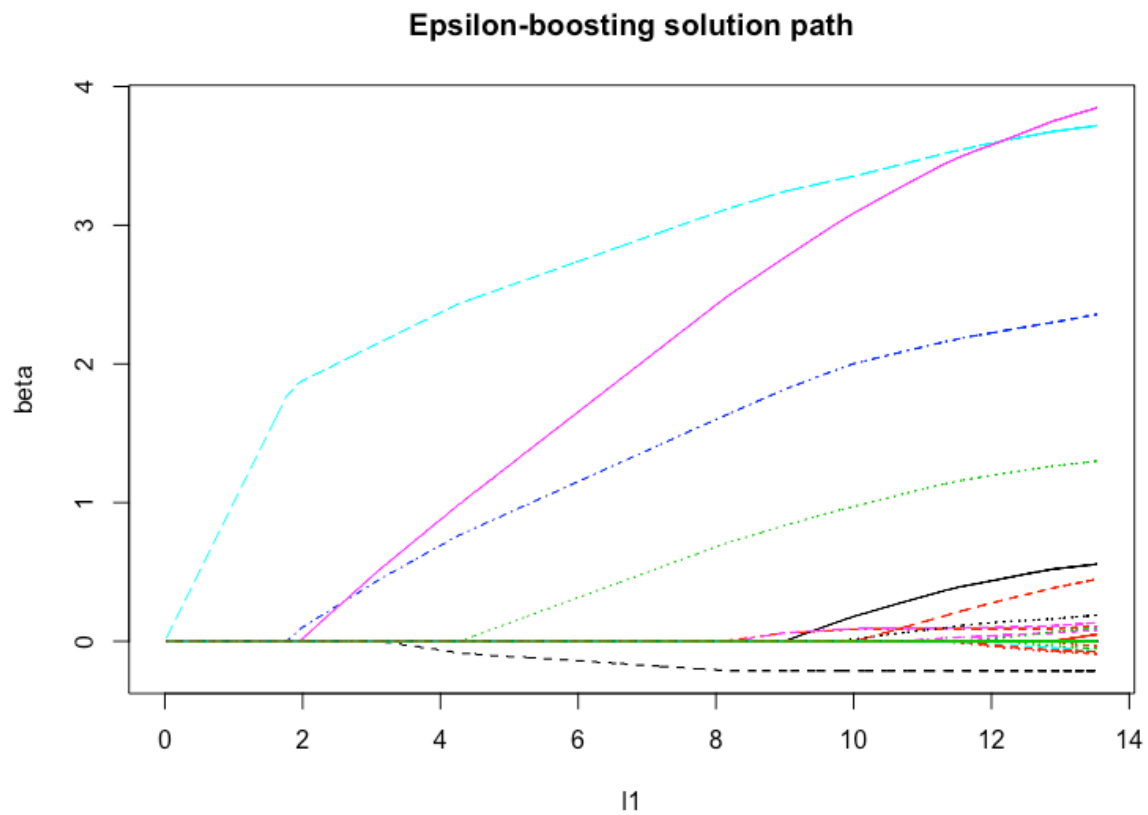
Since epsilon-boosting can be regarded as a transformation of Lasso, the solution paths are similar. The accuracy of epsilon-boosting depends on number of iteration steps. In practical, more than 1000 steps are required to make both epsilon boosting path similar to Lasso path.