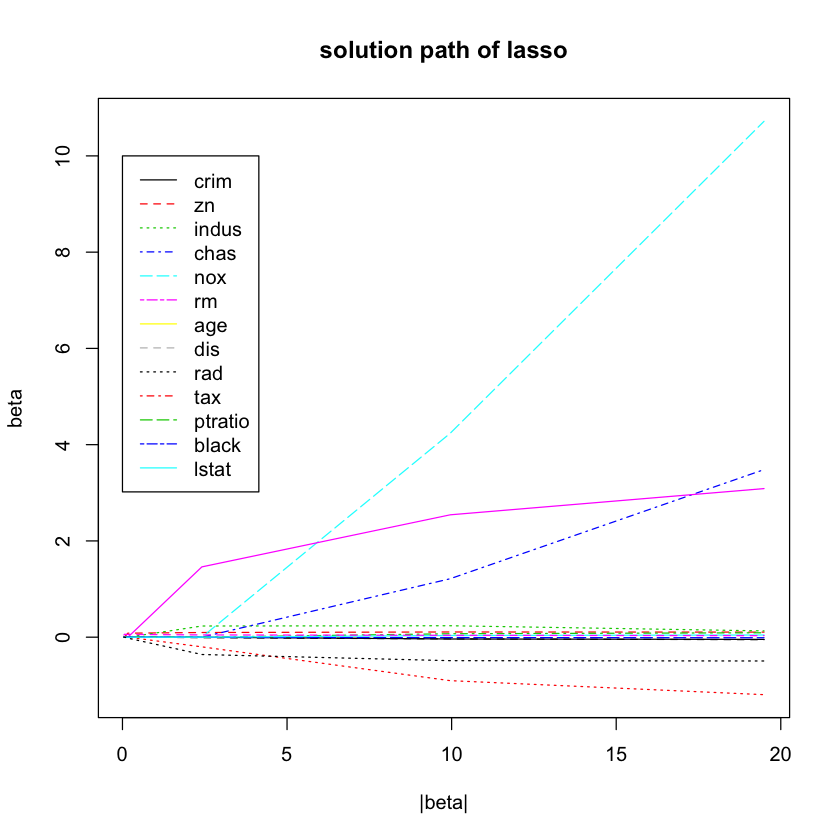
Solution path of Lasso:

The famous Boston Housing Dataset used here.



(1) Beta is initialed all zeros;

(2) As lambda decrease, the most related information will come in and so

It becomes nonzero;

(3) Beta l1 norm will becomes larger along the way as some of the

element becomes non-zero;

(4) Feature selected can be reflected along the way;

[R code]

library(Boston, package = "MASS")

X = Boston[, 1 : 13]

Y = Boston[, 14]

beta\_all = myLasso(X, Y, c(1000000, 100000, 10000, 1000, 100, 10))

matplot(t(t(matrix(rep(1, 13), nrow = 13)) %\*% abs(beta\_all)), t(beta\_all), type ="l", xlab = "|beta|", ylab = "beta", main = "solution path of lasso")

legend(0, 10, legend = colnames(Boston)[1:13], col = 1:13, lty = 1:13)