Homework V Data Mining II

Allen Rahrooh

March 4, 2020

Problem 8.4.2

It is mentioned in Section 8.2.3 that boosting using depth-one trees (or stumps) leads to an additive model: that is, a model of the form

$$f(X) = \sum_{j=1}^{p} f_j(X_j)$$

Explain why this is the case. You can begin with (8.12) in Algorithm 8.2.

Beginning with the boost algorithmn. Setting $\hat{f}(x) = 0$ and $r_i = y_i \forall i$ we get:

$$\hat{f}^1(x) = c_1 I(x_1 < t_1) + c_1^{'} = \frac{1}{\lambda} f_1(x_1)$$

which gives $\hat{f}(x) = \lambda \hat{f}^1(x)$ and $r_i = y_i - \lambda \hat{f}^1(x_i) \forall i$

Then we have:

$$\hat{f}^2(x) = c_2 I(x_2 < t_2) + c_2^{'} = \frac{1}{\lambda} f_2(x_2)$$

We then maximize the fit to the residuals which gives $\rightarrow \hat{f}(x) = \lambda \hat{f}^1(x) + \lambda \hat{f}^2(x)$ and $r_i = y_i - \lambda \hat{f}^1(x_i) - \lambda \hat{f}^2(x_i) \forall i$

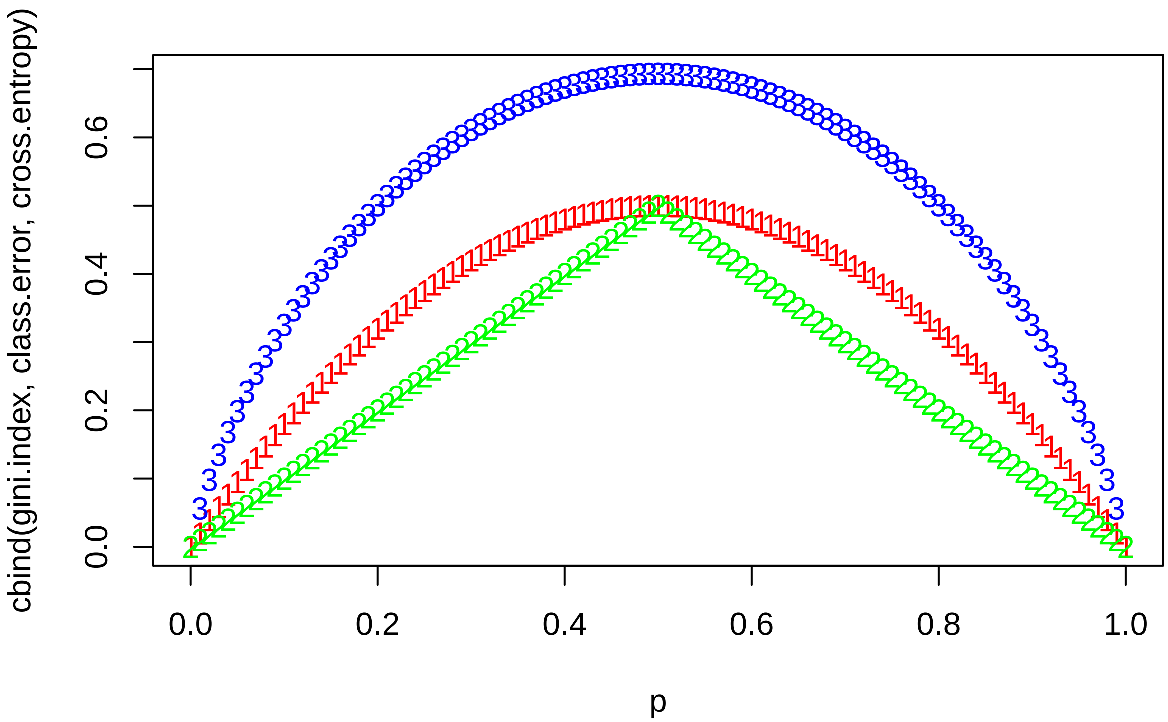
This results in:

$$\hat{f}(x) = \sum_{j=1}^{p} f_j(x_j)$$

Problem 8.4.3

Consider the Gini index, classification error, and cross-entropy in a simple classification setting with two classes. Create a single plot that displays each of these quantities as a function of $\hat{p}_{m1}$. The x-axis should display $\hat{p}_{m1}$, ranging from 0 to 1, and the y-axis should display the value of the Gini index, classification error, and entropy.

```r
p <- seq(0, 1, 0.01)
gini.index <- 2 * p * (1 - p)
class.error <- 1 - pmax(p, 1 - p)
cross.entropy <- - (p * log(p) + (1 - p) * log(1 - p))
matplot(p, cbind(gini.index, class.error, cross.entropy), col = c("red", "green", "blue"))
```



Problem 8.4.4

This question relates to the plots in Figure 8.12.

(a) Sketch the tree corresponding to the partition of the predictor space illustrated in the left-hand panel of Figure 8.12. The numbers inside the boxes indicate the mean of Y within each region.

If $X_1 \geq 1$ then 5, else if $X_2 \geq 1$ then 15, else if $X_1 < 0$ then 3, else if $X_2 < 0$ then 10, else 0.

```r
library(knitr)
include_graphics("hw5_allen_rahrooh.pdf")
```

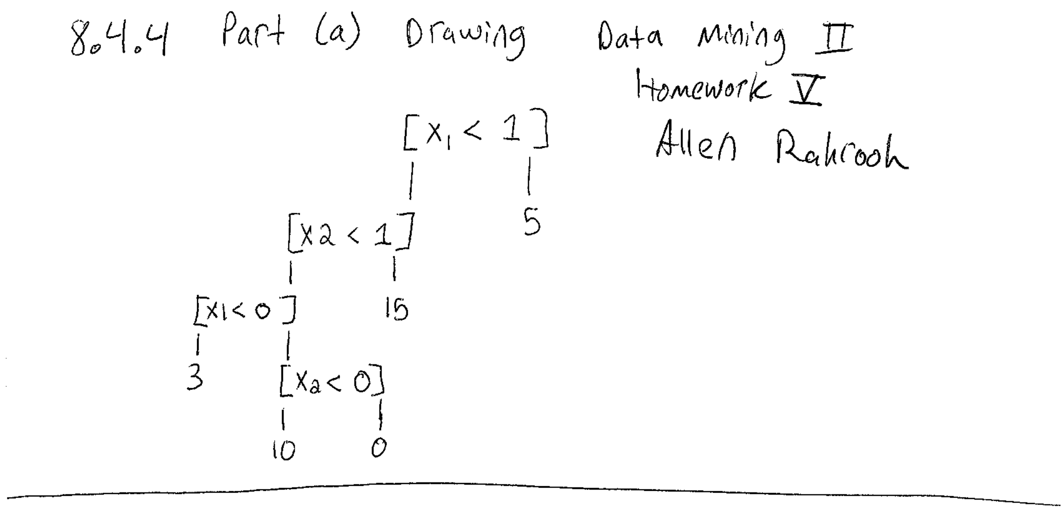


Figure 1: Sketch of Part (a)

(b) Create a diagram similar to the left-hand panel of Figure 8.12, using the tree illustrated in the right-hand panel of the same figure. You should divide up the predictor space into the correct regions, and indicate the mean for each region.

```
par(xpd = NA)
plot(NA, NA, type = "n", xlim = c(-2, 2), ylim = c(-3, 3),
     xlab = "X1", ylab = "X2", main = "Plot for 8.4.4")
# X2 < 1
lines(x = c(-2, 2), y = c(1, 1))

# X1 < 1 with X2 < 1
lines(x = c(1, 1), y = c(-3, 1))
text(x = (-2 + 1)/2, y = -1, labels = c(-1.8))
text(x = 1.5, y = -1, labels = c(0.63))

# X2 < 2 with X2 >= 1
lines(x = c(-2, 2), y = c(2, 2))
text(x = 0, y = 2.5, labels = c(2.49))

# X1 < 0 with X2<2 and X2>=1
lines(x = c(0, 0), y = c(1, 2))
text(x = -1, y = 1.5, labels = c(-1.06))
text(x = 1, y = 1.5, labels = c(0.21))
```