ISLR | Chapter 8 Exercises

Marshall McQuillen

10/05/2018

Conceptual

1

(sketch on following page)

2

Considering the regression setting, knowing that an additive model takes the form given below, the equation can be expanded to illustrate how it would behave when using a decision tree (with one split) as the base model.

$$f(x) = \frac{\sum_{j=1}^{P} f_j(x)}{P} \tag{1}$$

$$f(x) = \frac{\beta_1 I(x_{1i} \leq t) + \beta_2 I(x_{2i} \leq t) + \beta_3 I(x_{3i} \leq t) + \cdots + \beta_p I(x_{pi} \leq t)}{P} \tag{2}$$

When each tree has depth one, the model will take the form $f_j(X_j) = \beta_j I(x_{ji} \leq t) + \beta_0$, where t is the value in the range of X_j that minimizes the RSS. Knowing this, it is clear that the boosted model $f(x)$ will take the final form below:

$$f(x) = \frac{\sum_{j=1}^{P} f_j(X_j)}{P} \quad where \quad f_j(X_j) = \beta_j I(x_{ji} \leq t)$$

When each tree is depth-one (meaning there is only one split), each tree will only consist of one split on one variable. Further, *recursive binary splitting* is a **greedy algorithm**, meaning that, "…at each step of the tree-building process, the best split is made at at that particular step, rather than looking ahead and picking a split that will lead to a better tree in some future step." - *ISLR p.307* This makes it clear that when one wants a tree consisting of one variable, the split will occur at the value that minimizes the RSS, misclassification error rate, gini index or cross entropy.

It is important to recognize the implications of this; once a variable has been "used", the value that dictates the split is **the** (not **a**) value that reduces disorder in the response the most. In order to further reduce disorder in the response (weighted residuals after the first tree), the tree will have to choose another variable, since using one it has already chosen won't lead to a further decrease in the response (it would simply split on the same value, leading to no change in the overall model).

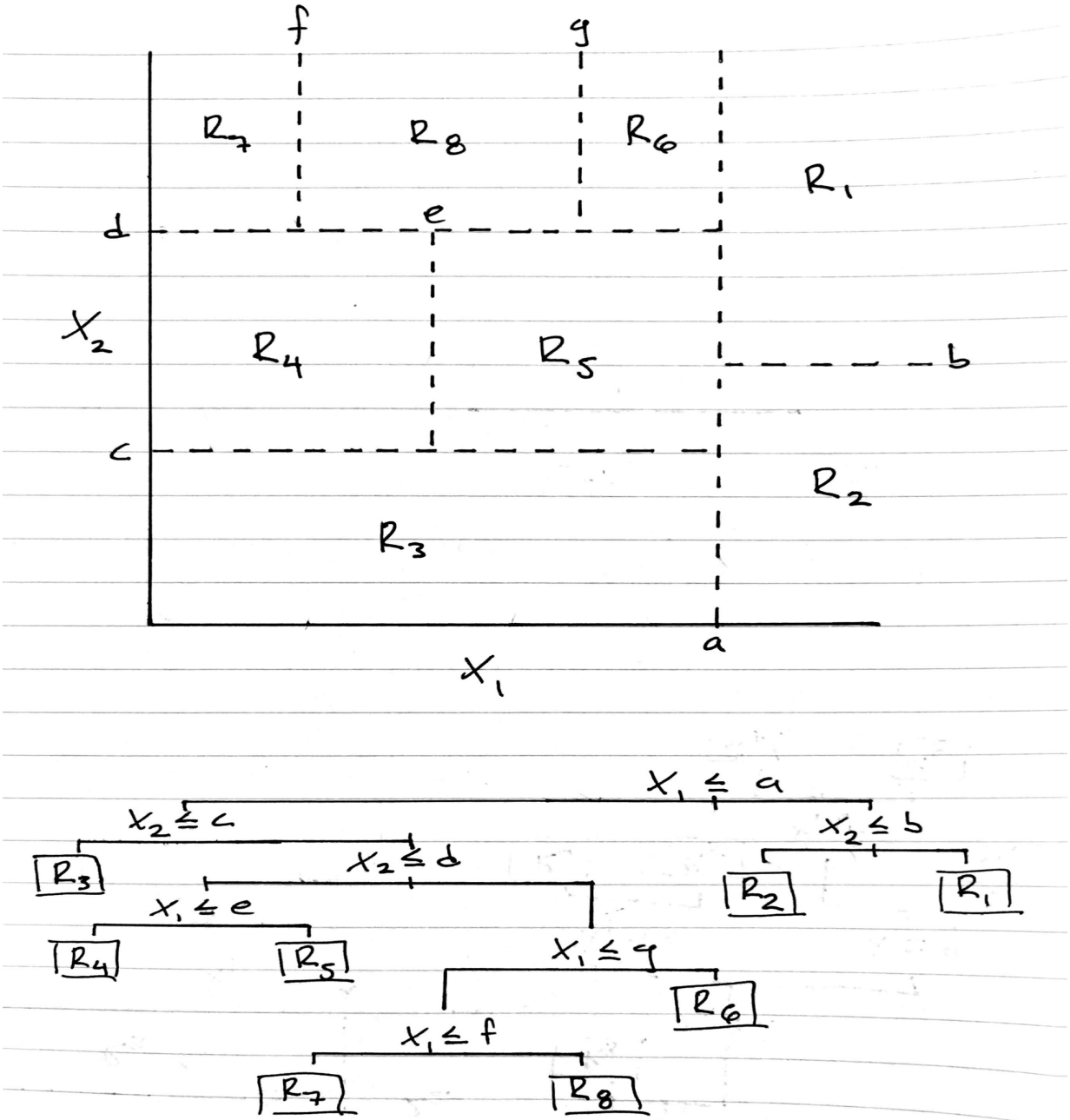

Figure 1: "Conceptual Exercise 1"