

STAT 430: Basics of Statistical Learning

Quiz II - Review Questions

Exercise 1

Two models were fit using **ridge** regression. One with $\lambda = 0.1$ and the other with $\lambda = 0.002$. Which was fit using $\lambda = 0.1$?

Model A:

## (Intercept)	x1	x2	x3	x4	x5
## -0.7618	0.3647	0.1893	1.0853	0.7744	1.3136

Model B:

## (Intercept)	x1	x2	x3	x4	x5
## -0.5782	0.3322	0.1869	0.9741	0.6806	1.1916

Exercise 2

Two models were fit using the **lasso**. One with $\lambda = 0.1$ and the other with $\lambda = 0.002$. Which was fit using $\lambda = 0.1$?

Model A:

## (Intercept)	x1	x2	x3	x4	x5
## 0.2709	0.0000	0.0000	0.5547	0.2724	0.8676

Model B:

## (Intercept)	x1	x2	x3	x4	x5
## -0.7445	0.3581	0.1838	1.0772	0.7665	1.3075

Exercise 3

Two models were selected, one using backwards AIC and the other using backwards BIC. Which was found using backwards BIC?

Model A:

## (Intercept)	x5
## 0.5515	1.1272

Model B:

## (Intercept)	x3	x4	x5
## -0.4745	1.0845	0.7446	1.3418

Exercise 4

Suppose we estimate the regression coefficients in a linear regression model by minimizing

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \quad \text{subject to} \quad \sum_{j=1}^p \beta_j^2 \leq s.$$

for a particular value of s . Draw a plot that indicates how train and test RMSE are affected by the choice of s . Assume that simply minimizing RSS would result in a model that overfits. Also assume that a model with only the intercept would underfit.

Exercise 5

Continue with the setup from Exercise 4. Consider using this model to make a prediction. What happens to **variance** as s is increased from 0?

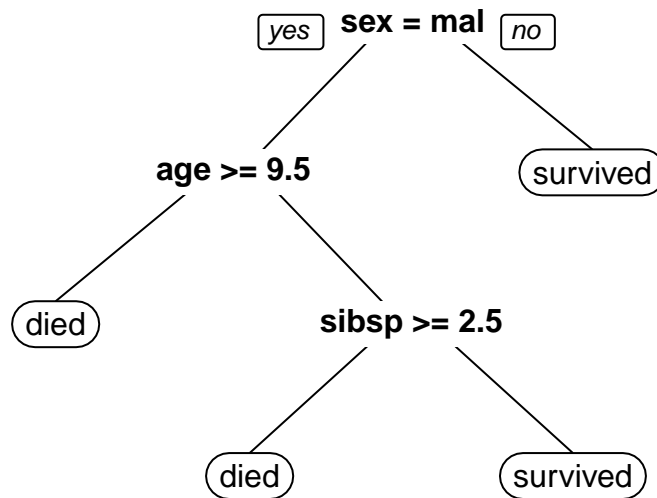
Exercise 6

Continue with the setup from Exercise 4. Consider using this model to make a prediction. What happens to **squared bias** as s is increased from 0?

Exercise 7

A tree model was fit to training data for a classification problem. Below is a sample of the data used, where `survived` is the response. We are attempting to predict the survival status of passengers of the Titanic:

pclass	survived	sex	age	sibsp	parch
1st	survived	female	29.0000	0	0
1st	survived	male	0.9167	1	2
1st	died	female	2.0000	1	2
1st	died	male	30.0000	1	2
1st	died	female	25.0000	1	2



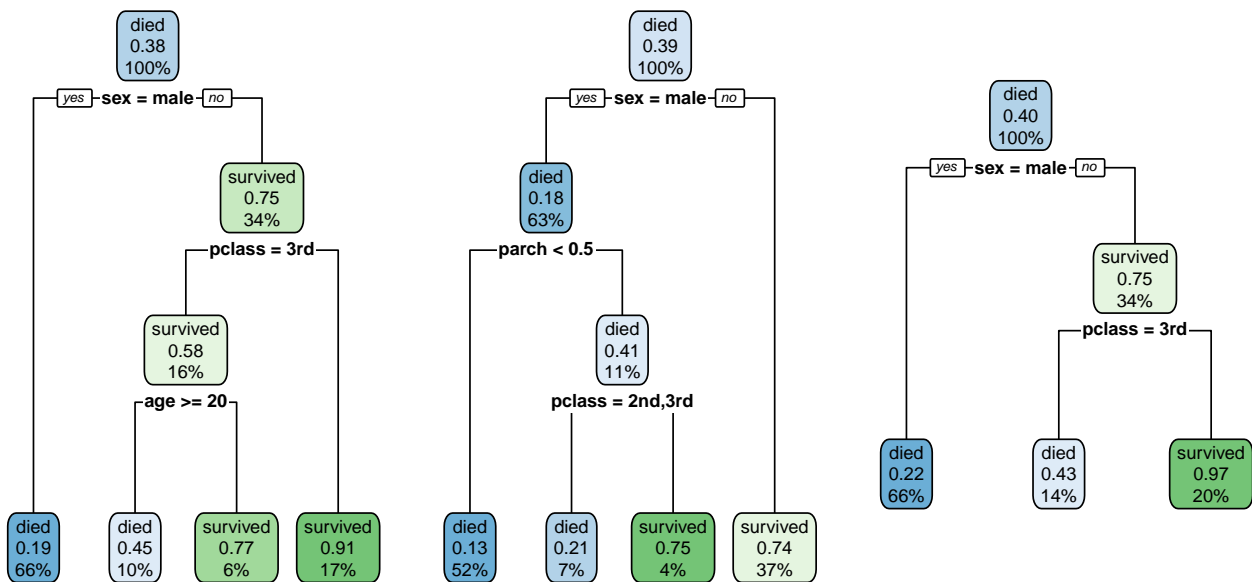
Predict the outcome of `survived` for the following three test passengers:

pclass	sex	age	sibsp	parch
1st	male	71	0	0
1st	female	48	1	0
2nd	male	3	1	1

Exercise 8

A bagged tree model (using a total of 3 trees) was fit to training data for a classification problem. Below is a sample of the data used, where `survived` is the response. We are attempting to predict the survival status of passengers of the Titanic:

pclass	survived	sex	age	sibsp	parch
1st	survived	female	29.0000	0	0
1st	survived	male	0.9167	1	2
1st	died	female	2.0000	1	2
1st	died	male	30.0000	1	2
1st	died	female	25.0000	1	2

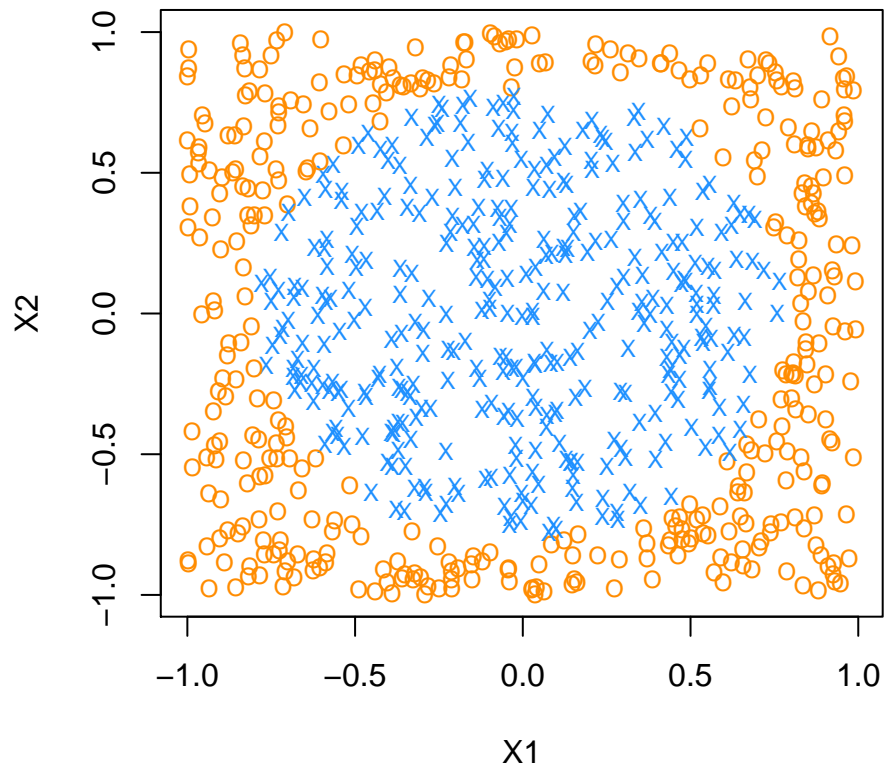


Predict the outcome of `survived`, and predicted probability of surviving for the following test passengers:

pclass	sex	age	sibsp	parch
1st	male	71	0	0
1st	female	48	1	0

Exercise 9

Consider fitting classification methods to the data in the plot below. For each part below, specify which of the two methods will perform better.



- (a) Additive Logistic Regression or k -Nearest Neighbors?
- (b) Additive Logistic Regression or a single Decision Tree?
- (c) Additive Logistic Regression or a Random Forest?
- (d) A single decision tree or a Random Forest?