An introduction to statistical learning - Ch6 - Ex11

Thalles Quinaglia Liduares

24/03/2023

  a. Try out some of the regression methods explored in this chapter, such as best subset selection, the
     lasso, ridge regression, and PCR. Present and discuss results for the approaches that you consider.

  i. Lasso Regression

```
# Load the Boston dataset
library(MASS)
data(Boston)
```

```
# Set up the predictor and response variables

# all predictors except the first (per capita crime rate)

X <- as.matrix(Boston[,-1])

# per capita crime rate

y <- Boston[,1]
```

```
# Split the data into training and test sets

set.seed(123)

train <- sample(nrow(X), nrow(X)/2)
X_train <- X[train,]
y_train <- y[train]
X_test <- X[-train,]
y_test <- y[-train]
```

```
# Fit a Lasso regression model with cross-validation

library(glmnet)
```

```
## Carregando pacotes exigidos: Matrix
```

```
## Loaded glmnet 4.1-6
```

```r
# alpha=1 specifies Lasso regression

cv_fit <- cv.glmnet(X_train, y_train, alpha=1)

# lambda that gives minimum mean cross-validated error

lambda_min <- cv_fit$lambda.min
```

```r
# Fit the final Lasso regression model using the selected lambda
lasso_fit <- glmnet(X_train, y_train, alpha=1, lambda=lambda_min)
```

```r
# Make predictions on the test set

y_pred <- predict(lasso_fit, newx=X_test)
```

```r
# Calculate the test error (root mean squared error)

test_error <- sqrt(mean((y_test - y_pred)^2))
```

```r
# Print the test error

cat("Test error:", test_error, "\n")
```

```
## Test error: 7.173717
```