STA 478/578 Midterm Exam 2

Jared Brotamonte

November 9, 2023

INSTRUCTIONS: You may use any materials that have been provided, including your notes, textbook, and online information. I would encourage that you *not* use outside online resources and you must cite any material used not from this course. If any information is used outside of the ISLR and in-class notes, details must be given on what was used and how it works. **You may NOT share solutions among classmates.** The exam has no time limit besides the due date of Monday (11/13/2022) before 5:59 PM.

Please prepare your solutions using the RMD file provided. You may change options to suit your style, but be sure to keep the document organized. You will also be required to load a .RData file for the data sets. Be sure to place the RMD and RData file together in the same directory to quickly load. *Justify all free response answers. Type solutions after each prompt and try to keep your PDF organized.* The points for each question are given.

Please submit your exam as an organized PDF document directly from RMD. The solution should be in the order they were asked. If there are problems with your PDF you will be asked to resubmit. Organization, clarity and correct preparation of solutions will be worth **5 points**.

Due Monday, November 13th, 2022 before 5:59 PM.

Monte Carlo Setup

We will evaluate the following data simulations, known as Monte Carlo simulations, that allow us to compare how least squares, step-wise, LASSO, RIDGE, and Elastic-NET compare under different scenarios. We will prepare functions to allow us to make these simulations flexible and efficient. By controlling the simulations, we will also be able to calculate the `Bias` and `Variance` directly. I will give you a few details here, and additional functions and output will need to be generated by you.

We will want to study models that have a mixture of effects within the predictor set. I provide two ways of generating data below.

Normally Distributed Uncorrelated Predictors

```
genData <- function(n, beta)
{
  require(mvtnorm)
  p <- length(beta)
  sigma <- diag(1, p)
  x <- matrix(rmvnorm(n, rep(0, p), sigma), nrow=n, ncol=p)
  y <- rnorm(n, x%*%beta, 2)
  return(list(x=x, y=y))
}
```

This function accepts a sample size (n) and a vector of known coefficients (β). It will return a list including the predictors x and the response y for each of the n observations. The irreducible error has been set to a standard deviation of 2. There are several ways we could make this data generation more flexible, but you will only be required to generate data under the above conditions.

Normally Distributed Correlated Predictors

```
genData.correlated <- function(n,beta,rho)
{
  require(mvtnorm)
  p <- length(beta)
  sigma <- c(1, rho^seq(1:(p-1)))
  sigma <- toeplitz(sigma)
  x <- matrix(rmvnorm(n, rep(0, p), sigma), nrow=n, ncol=p)
  y <- rnorm(n, x%*%beta, 2)
  return(list(x=x, y=y))
}
```

This function will generate strongly correlated predictors along with normally distributed responses from those predictors. The ρ value controls the strength of how correlated 'nearby' predictors are. That is, when we set $\rho = 0.9$, then predictor x_j will have a correlation with predictors x_{j-1} and x_{j+1} of roughly 0.9. The strength of the correlation decreases as the predictors get 'further apart'.

Model Estimation

Below you will be asked to generate several functions that will estimate models from the data generated above. To aid in these steps, I provide for you my version for calculating validated LASSO coefficients and provide a hybrid selection estimator. You are welcome to change these, but ensure they give the needed output. We will be interested in knowing the estimated coefficients from the models (omitting the intercept). This will allow us to calculate all summary statistics - and finally be able to see the Bias-Variance balance.

Estimation Examples (LASSO and STEP)

```
LASSO.estimator <- function(Data)
{
  require(glmnet)
  LASSO.cv <- cv.glmnet(Data$x, Data$y, alpha=1)
  return(as.numeric(coef(LASSO.cv, s=LASSO.cv$lambda.min))[-1])
}

STEP.estimator <- function(Data)
{
  dt<- data.frame(x=Data$x,y=Data$y)
  ft <- lm(y~., data = dt)
  Coef.fit <- c(ft$coefficients)
  STEP <- step(lm(y~., data = dt), direction = "both", trace = FALSE,k=2)
  Coef.STEP <- c(STEP$coefficients)
  Coef.STEP <- bind_rows(Coef.fit,Coef.STEP)[-1,]
  Coef.STEP[is.na(Coef.STEP)] <- 0
  return(as.numeric(Coef.STEP[-1]))
}
```

The `LASSO.estimator` and `STEP.estimator` function takes the data in the form that is generated from the given functions (`genData` and `genData.correlated`). For penalized regression the function performs cross-validation of the the penalty coefficient λ, and returns the model coefficients from λ that minimizes the cross-validation error. The `(intercept)` is removed, as we will focus only on the coefficients β.

Calculating MSE

Once we have the model estimated coefficients, we will be able to calculate the `Bias`, `Variance` and `MSE` of each iteration.

The `Bias` can be found be determining the difference between our estimated coefficients the the true known coefficients.

$$Bias[\hat{\theta}] = E[\hat{\theta} - \theta]$$

The `variance` can be found by asking how much variation we observed in the estimated coefficients.

$$Var[\hat{\theta}] = E[(\hat{\theta} - E[\hat{\theta}])^2]$$

Finally, recall that the `MSE` is the sum of the squared bias and variance.

$$MSE[\hat{\theta}] = Bias^2[\hat{\theta}] + Var[\hat{\theta}]$$

Here I show how these statements can be used to generated the `MSE` based on its `Bias` and `Variance` decomposition. You will be asked to run these simulations under different scenarios below. Here I produce the LASSO simulations using the true coefficient vector $\beta = c(-2, 2, -1, 1, 0, 0)$ based on $n = 100$ samples. I produce $Iterations = 200$ different data sets and then generate the LASSO model on each generated set. I then calculate the `Bias`, `Variance`, and `MSE` of the method. These can take a little time so be careful to have everything setup correctly. We will run a low number of iterations, but you are welcome to run a larger number if you want to improve the clarity of the simulation.

```
### Simulation
n <- 100
beta <- c(-2, 2, -1, 1, 0, 0)
Iterations <- 200
coef.temp <- matrix(ncol=length(beta), nrow=Iterations)
for(j in 1:Iterations) coef.temp[j,] <- LASSO.estimator(genData(n,beta))
```

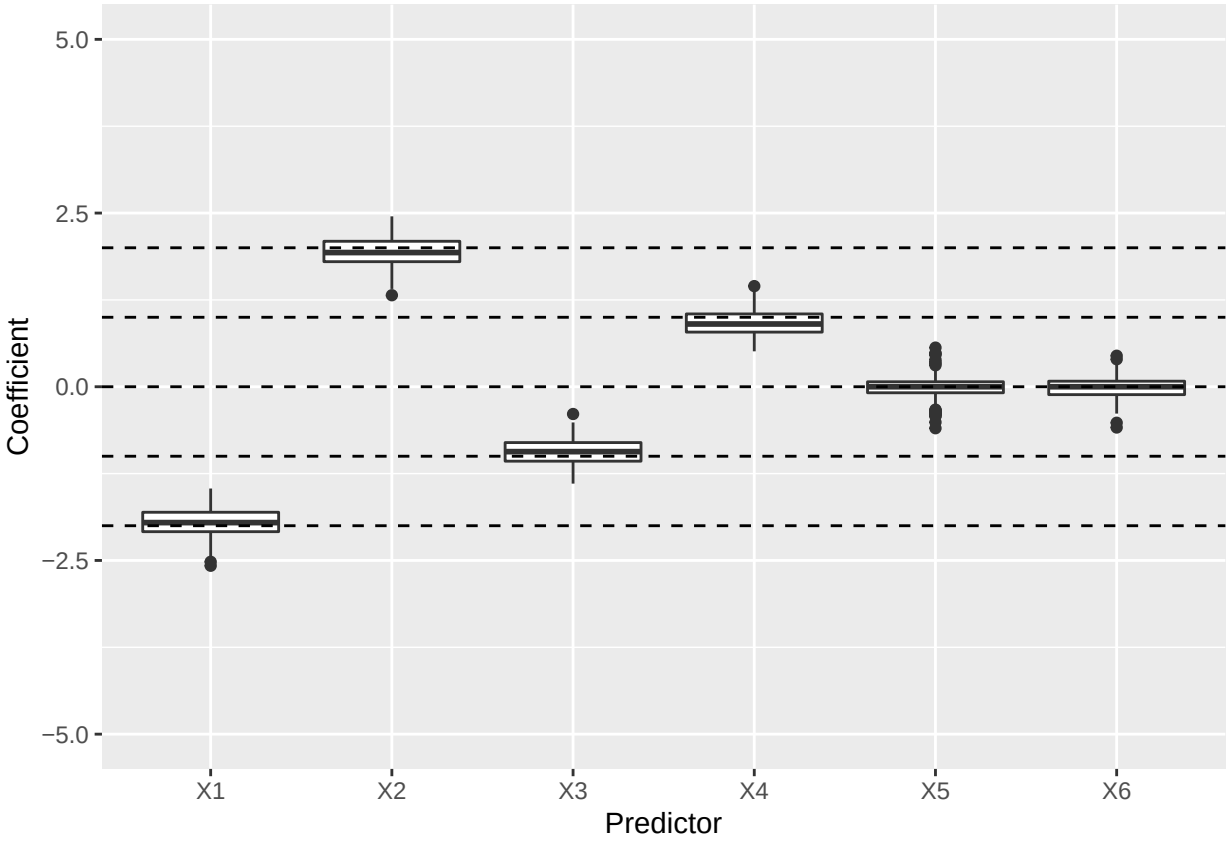
After running the simulation, we can now calculate our estimates. View these output objects, they are interesting and something you haven't seen in the course directly. They are the actual decomposed values of `Bias` and `Variance` based on the average across each iteration.

```
### Bias, Variance, MSE
Bias <- apply(coef.temp, 2, mean) - beta ## E[Theta] - Theta
Var <- apply(coef.temp, 2, var) ## Var[Theta]
MSE <- Bias^2 + Var
simulation.result <- data.frame(MSE = sum(MSE), Bias = sum(Bias^2), Var = sum(Var))
simulation.result
```

```
##         MSE       Bias       Var
## 1 0.2442576 0.01626044 0.2279971
```

We will use the simulation to produce a discussion of how well each technique does through a comparison of the MSE and visualization of the coefficient estimates.

```
coef.df <- data.frame(coef.temp) %>% mutate(Iteration = 1:n())
coef.df.long <- coef.df %>%
  pivot_longer(names_to = 'Predictor', values_to = 'Coefficient', 1:6)
ggplot(coef.df.long) + geom_boxplot(aes(x=Predictor, y=Coefficient)) +
  geom_hline(yintercept = c(-2,2,-1,1,0), lty=2) +
    coord_cartesian(ylim = c(-5, 5))
```


5.0
2.5
0.0
−2.5
−5.0
Coefficient
X1
X2
X3
X4
X5
X6
Predictor

Exam Questions

Question 1 [15 points]

Following the form of the `LASSO.estimator()` function above produce the following functions:

a. `LS.estimator` : This function should produce the least squares estimated coefficients.

```
LS.estimator <- function(Data) {
  lm.fit <- lm(y ~ ., data = as.data.frame(Data))
  return(as.numeric(lm.fit$coefficients)[-1])  # Omit the intercept
}
```

b. `RIDGE.estimator` : This function should conduct cross-validation of a Ridge Regression penalty coefficient and return the coefficients of the chosen model at the λ that minimizes the cross-validation error.

```
RIDGE.estimator <- function(Data) {
  require(glmnet)
  ridge.cv <- cv.glmnet(Data$x, Data$y, alpha = 0)
  return(as.numeric(coef(ridge.cv, s = ridge.cv$lambda.min))[-1])
}
```

c. Create the `ENET.estimator` in the same manner as you have for LASSO and Ridge Regression but set the α argument to 0.5. This is the Elastic-NET that provides a balance of Ridge Regression ($\alpha = 0$) and LASSO ($\alpha = 1$). This function should conduct cross-validation of an Elastic-NET penalty coefficient and return the coefficients of the chosen model at the λ that minimizes the cross-validation error.

```
ENET.estimator <- function(Data) {
  require(glmnet)
  enet.cv <- cv.glmnet(Data$x, Data$y, alpha = 0.5)
  return(as.numeric(coef(enet.cv, s = enet.cv$lambda.min))[-1])
}
```

Question 2 [20 points]

Hint for this question, produce a single loop that generates data and fits all models, rather than many small loops. Ensure you fit all models to the same data generation.

a. Run at least 200 simulations comparing our full linear model (least squares), step-wise, LASSO, Ridge Regression, and Elastic-NET estimation of model coefficients. Generate the data using `genData.correlated` using `rho = 0.9`. Use a sample size of 100. Let the known coefficients be:

```
beta.3 <- c(rep(0.4, 80))
```

If any methods cannot be applied in this scenario, be sure to discuss why.

```
# Set parameters
n <- 100
rho <- 0.9
beta.3 <- c(rep(0.4, 80))
Iterations <- 200
```

```r
# Initialize matrices to store results
coef_ls <- matrix(ncol = length(beta.3), nrow = Iterations)
coef_step <- matrix(ncol = length(beta.3), nrow = Iterations)
coef_lasso <- matrix(ncol = length(beta.3), nrow = Iterations)
coef_ridge <- matrix(ncol = length(beta.3), nrow = Iterations)
coef_enet <- matrix(ncol = length(beta.3), nrow = Iterations)

# Run simulations
for (j in 1:Iterations) {
  # Generate correlated data
  correlated_data <- genData.correlated(n, beta.3, rho)

  # Least Squares (LS)
  coef_ls[j,] <- LS.estimator(correlated_data)

  # Stepwise
  coef_step[j,] <- STEP.estimator(correlated_data)

  # LASSO
  coef_lasso[j,] <- LASSO.estimator(correlated_data)

  # Ridge Regression
  coef_ridge[j,] <- RIDGE.estimator(correlated_data)

  # Elastic-NET
  coef_enet[j,] <- ENET.estimator(correlated_data)
}
```

```
## Loading required package: mvtnorm
```

b. Prepare a table to compare the `MSE` of each method along with its decomposition into `Bias` and `Variance` components. Discuss the effectiveness of each method. Which method is preferred?

```r
# Calculate Bias, Variance, and MSE for each method
Bias_ls <- apply(coef_ls, 2, mean) - beta.3
Var_ls <- apply(coef_ls, 2, var)
MSE_ls <- Bias_ls^2 + Var_ls

Bias_step <- apply(coef_step, 2, mean) - beta.3
Var_step <- apply(coef_step, 2, var)
MSE_step <- Bias_step^2 + Var_step

Bias_lasso <- apply(coef_lasso, 2, mean) - beta.3
Var_lasso <- apply(coef_lasso, 2, var)
MSE_lasso <- Bias_lasso^2 + Var_lasso

Bias_ridge <- apply(coef_ridge, 2, mean) - beta.3
Var_ridge <- apply(coef_ridge, 2, var)
MSE_ridge <- Bias_ridge^2 + Var_ridge

Bias_enet <- apply(coef_enet, 2, mean) - beta.3
Var_enet <- apply(coef_enet, 2, var)
MSE_enet <- Bias_enet^2 + Var_enet
```

```r
# Create a table
comparison_table <- data.frame(
  Method = c("LS", "Stepwise", "LASSO", "Ridge", "Elastic-NET"),
  MSE = c(sum(MSE_ls), sum(MSE_step), sum(MSE_lasso), sum(MSE_ridge), sum(MSE_enet)),
  Bias = c(sum(Bias_ls^2), sum(Bias_step^2), sum(Bias_lasso^2), sum(Bias_ridge^2), sum(Bias_enet^2)),
  Variance = c(sum(Var_ls), sum(Var_step), sum(Var_lasso), sum(Var_ridge), sum(Var_enet))
)

# Print the table
print(comparison_table)
```

```
##        Method        MSE       Bias   Variance
## 1          LS 167.061439 0.79381365 166.267625
## 2    Stepwise 144.792680 0.72421115 144.068469
## 3       LASSO  17.543742 0.10189033  17.441851
## 4       Ridge   1.274941 0.01807445   1.256867
## 5 Elastic-NET  15.513975 0.08983034  15.424144
```

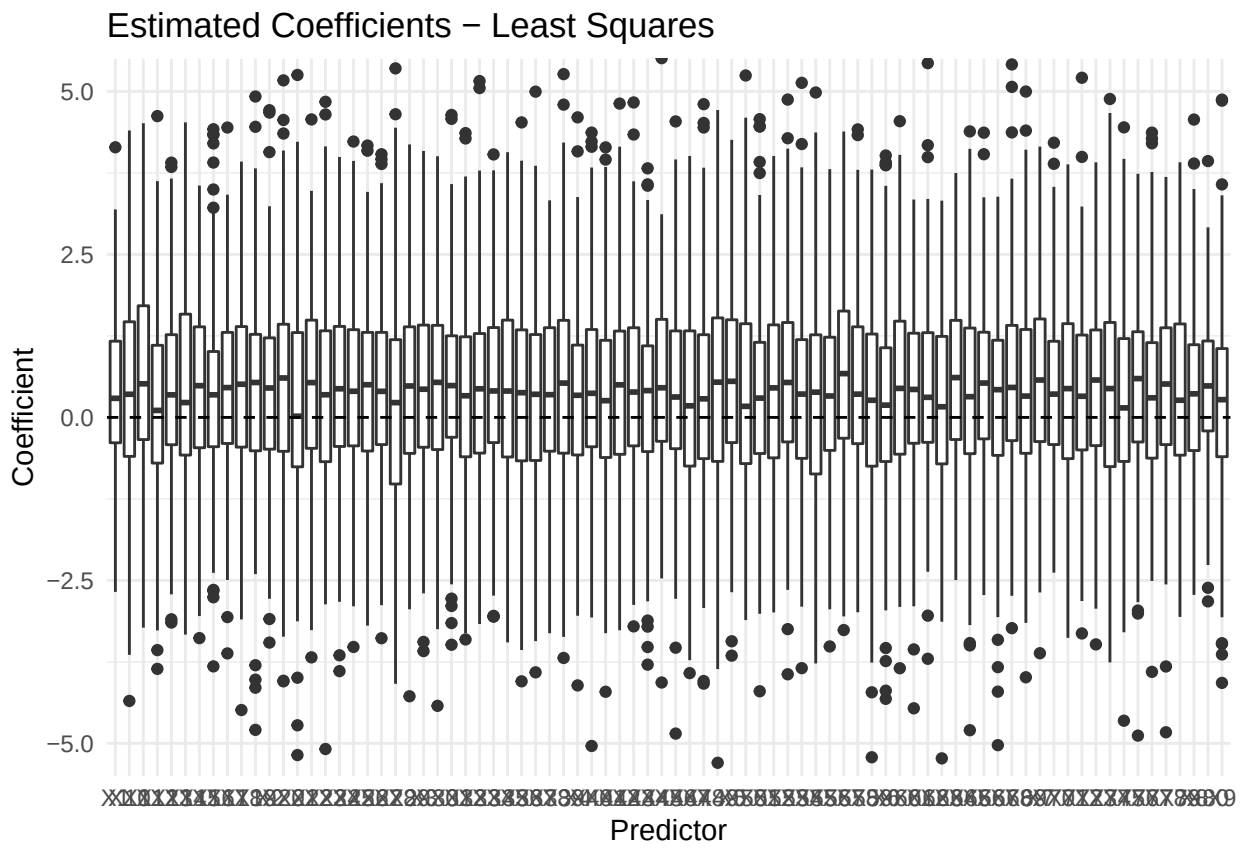
Ridge Regression had the lowest MSE out of all of them thus suggesting that it balanced Bias and Variance the best in this scenario. Least Squares had the highest MSE out of all of them thus suggesting that there is a likely over fitting of the data in this scenario. And although not the lowest, the LASSO method still performed relatively well with a rather low MSE.

c. Give a brief explanation based on the conceptual understanding of these methods if this result agrees with what we have studied. The results do align with the conceptual understanding of what we have studied. Ridge Regression with the lowest MSE demonstrated it's effectiveness in balancing bias and variance in the presence of correlated predictors.

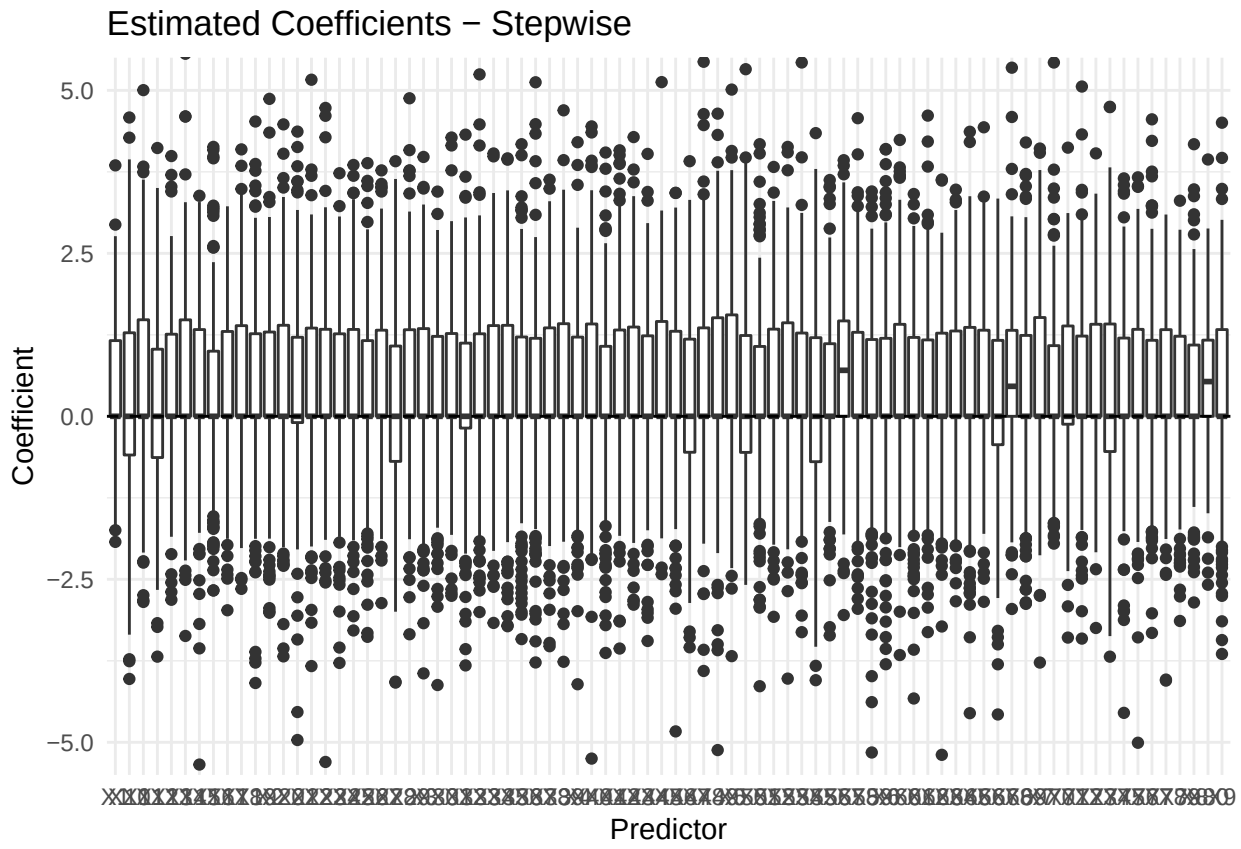
d. Prepare a visual representation of the estimated coefficients for each method. Provide a brief description of how well each method seemed to estimate the coefficients to go along with the visual representation. **Use all 80 coefficients for visualization.**

```r
# Create a function to generate coefficient plots
plot_coefficients <- function(method, coef_matrix) {
  coef.df <- data.frame(coef_matrix) %>% mutate(Iteration = 1:n())
  coef.df.long <- coef.df %>%
    pivot_longer(names_to = 'Predictor', values_to = 'Coefficient', -Iteration)

  # Plotting
  ggplot(coef.df.long) +
    geom_boxplot(aes(x = Predictor, y = Coefficient)) +
    geom_hline(yintercept = 0, lty = 2) +
    ggtitle(paste("Estimated Coefficients -", method)) +
    coord_cartesian(ylim = c(-5, 5)) +
    theme_minimal()
}

# Plot coefficients for each method
plot_ls <- plot_coefficients("Least Squares", coef_ls)
plot_step <- plot_coefficients("Stepwise", coef_step)
plot_lasso <- plot_coefficients("LASSO", coef_lasso)
plot_ridge <- plot_coefficients("Ridge", coef_ridge)
plot_enet <- plot_coefficients("Elastic-NET", coef_enet)
```

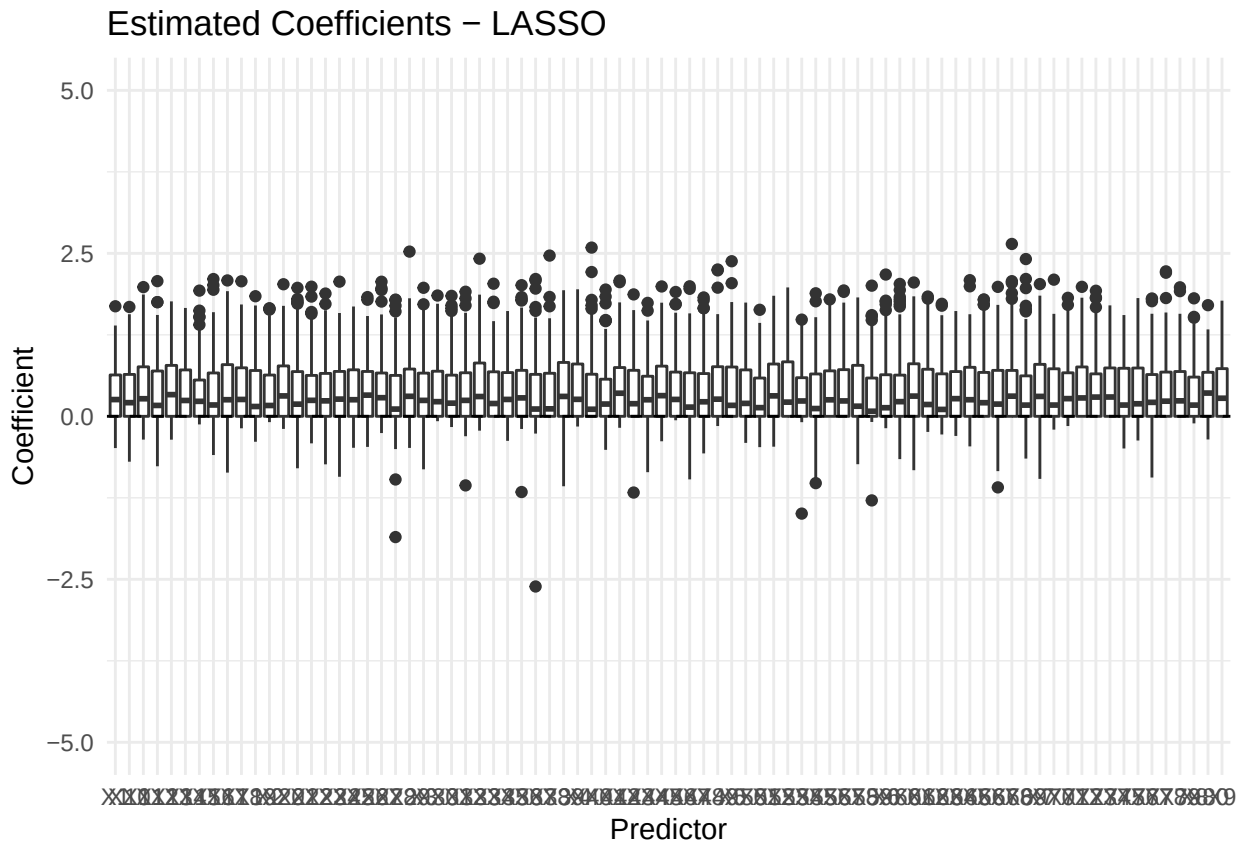
```
# Display the plots
plot_ls
```



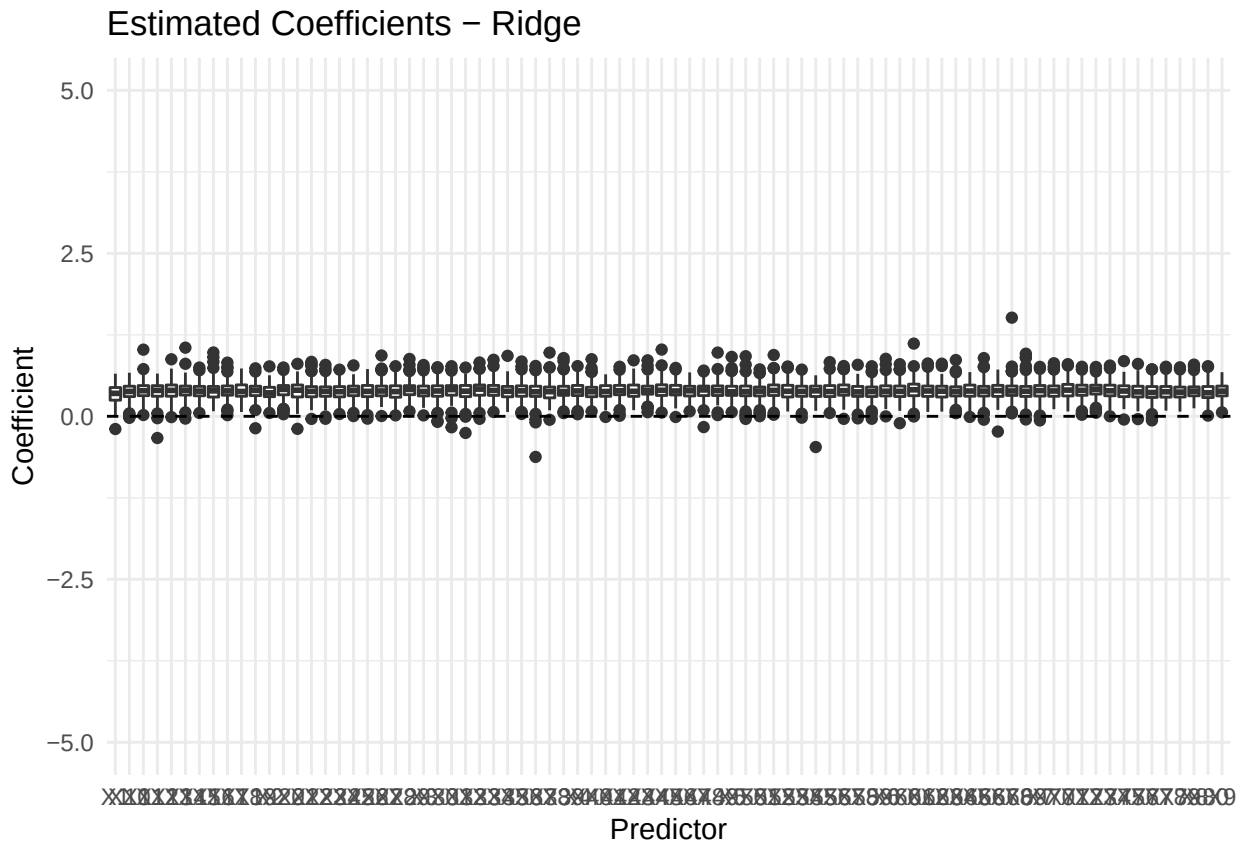

```
plot_step
```


Estimated Coefficients − Stepwise
Coefficient
5.0
2.5
0.0
−2.5
−5.0
Predictor

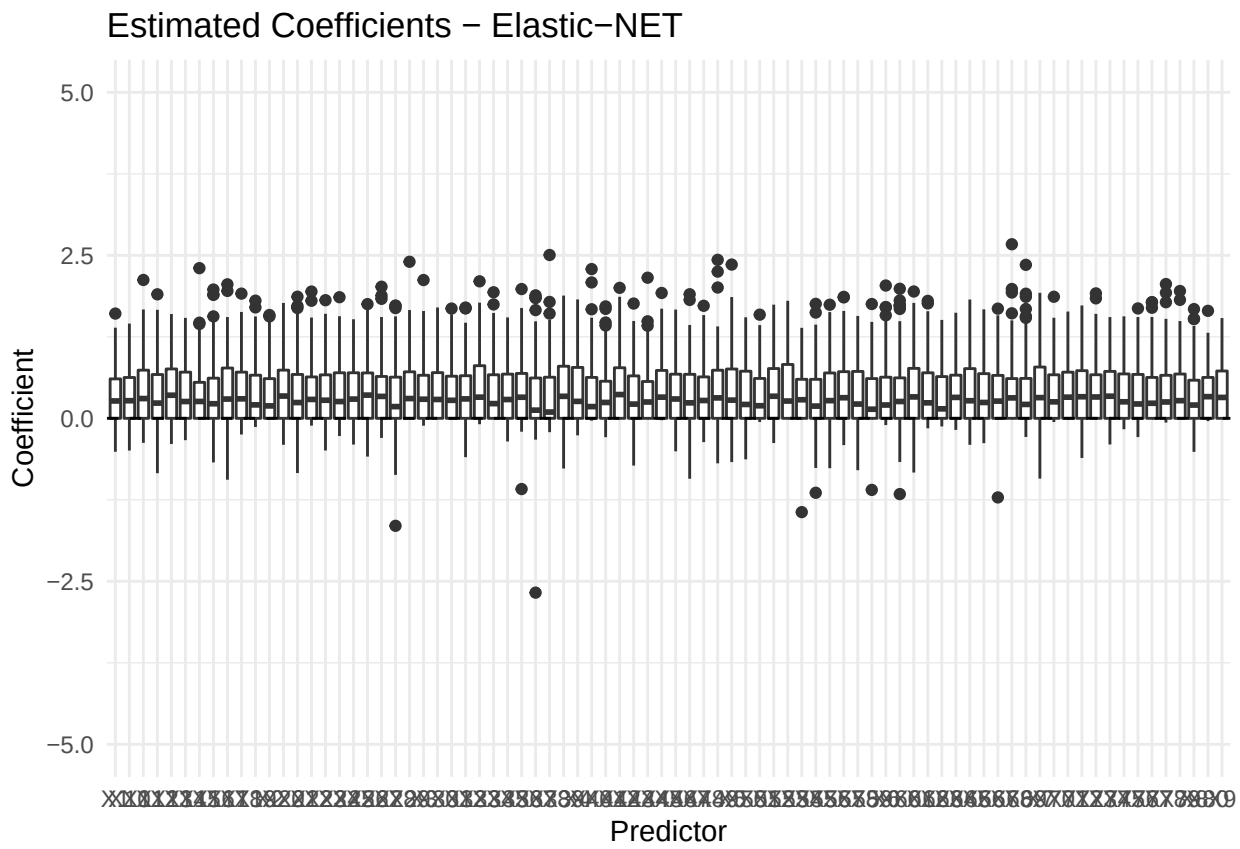
```
plot_lasso
```


Estimated Coefficients − LASSO
5.0
2.5
0.0
−2.5
−5.0
Coefficient
Predictor

```
plot_ridge
```


Estimated Coefficients − Ridge
5.0
2.5
0.0
−2.5
−5.0
Coefficient
Predictor

```
plot_enet
```



Based on the graphs, the Ridge method estimated the coefficients with the least amount of error, while the Stepwise and LS methods were not as accurate.

Question 3 [20 points]

Again, you want to produce a single loop that generates data and fits all models, rather than many small loops.

a. Run at least 200 simulations comparing least squares, step-wise, LASSO, Ridge Regression, and Elastic-NET estimation of model coefficients. Generate the data using `genData`. Use a sample size of 100. Let the known coefficients be:

```
beta.4 <- c(-2, 2, -1, 1, rep(0, 996))
```

If any methods cannot be applied in this scenario, be sure to discuss why.

```
# Define the sample size and known coefficients
n <- 100
beta.4 <- c(-2, 2, -1, 1, rep(0, 996))

# Number of simulations
Iterations <- 200

# Initialize matrices to store coefficients
coef_ls <- matrix(NA, nrow = Iterations, ncol = length(beta.4))
coef_step <- matrix(NA, nrow = Iterations, ncol = length(beta.4))
```

```
coef_lasso <- matrix(NA, nrow = Iterations, ncol = length(beta.4))
coef_ridge <- matrix(NA, nrow = Iterations, ncol = length(beta.4))
coef_enet <- matrix(NA, nrow = Iterations, ncol = length(beta.4))

# Run simulations
for (j in 1:Iterations) {
  # Generate data
  generated_data <- genData(n, beta.4)

  # Least Squares (LS)
  coef_ls[j,] <- LS.estimator(generated_data)

  # Stepwise
  # coef_step[j,] <- STEP.estimator(generated_data)
  # Error in step(lm(y ~ ., data = dt), direction = "both", trace = FALSE,  :
  # AIC is -infinity for this model, so 'step' cannot proceed

  # LASSO
  coef_lasso[j,] <- LASSO.estimator(generated_data)

  # Ridge Regression
  coef_ridge[j,] <- RIDGE.estimator(generated_data)

  # Elastic-NET
  coef_enet[j,] <- ENET.estimator(generated_data)
}
```

In this situation, it seems that the Stepwise method is not possible due to multicollinearity in the predictors.

b. Prepare a table to compare the `MSE` of each method along with its decomposition into `Bias` and `Variance` components. Discuss the effectiveness of each method. Which method is preferred?

```
# Calculate Bias, Variance, and MSE for each method
Bias_ls <- apply(coef_ls, 2, mean) - beta.4
Var_ls <- apply(coef_ls, 2, var)
MSE_ls <- Bias_ls^2 + Var_ls

Bias_step <- apply(coef_step, 2, mean) - beta.4
Var_step <- apply(coef_step, 2, var)
MSE_step <- Bias_step^2 + Var_step

Bias_lasso <- apply(coef_lasso, 2, mean) - beta.4
Var_lasso <- apply(coef_lasso, 2, var)
MSE_lasso <- Bias_lasso^2 + Var_lasso

Bias_ridge <- apply(coef_ridge, 2, mean) - beta.4
Var_ridge <- apply(coef_ridge, 2, var)
MSE_ridge <- Bias_ridge^2 + Var_ridge

Bias_enet <- apply(coef_enet, 2, mean) - beta.4
Var_enet <- apply(coef_enet, 2, var)
MSE_enet <- Bias_enet^2 + Var_enet

# Create a table
```

```r
comparison_table <- data.frame(
  Method = c("LS", "Stepwise", "LASSO", "Ridge", "Elastic-NET"),
  MSE = c(sum(MSE_ls), sum(MSE_step), sum(MSE_lasso), sum(MSE_ridge), sum(MSE_enet)),
  Bias = c(sum(Bias_ls^2), sum(Bias_step^2), sum(Bias_lasso^2), sum(Bias_ridge^2), sum(Bias_enet^2)),
  Variance = c(sum(Var_ls), sum(Var_step), sum(Var_lasso), sum(Var_ridge), sum(Var_enet))
)

# Print the table
print(comparison_table)
```

```
##        Method      MSE     Bias  Variance
## 1          LS       NA       NA        NA
## 2    Stepwise       NA       NA        NA
## 3       LASSO 1.844006 1.109279 0.7347268
## 4       Ridge 9.446571 9.065727 0.3808438
## 5 Elastic-NET 2.460331 1.639299 0.8210324
```

Due to the multicollinearity, the LS and Stepwise methods are not very accurate for this situation.

c. Give a brief explanation based on the conceptual understanding of these methods if this result agrees with what we have studied. The results align with conceptual methods because it highlights the challenges that methods like LS and Stepwise face due to multicollinearity of the data.

d. Prepare a visual representation of the estimated coefficients for each method. Provide a brief description of how well each method seemed to estimate the coefficients to go along with the visual representation. **You should use only the first 25 coefficients for visualization. Attempt to order them correctly to ensure the graph is easy to interpret.**

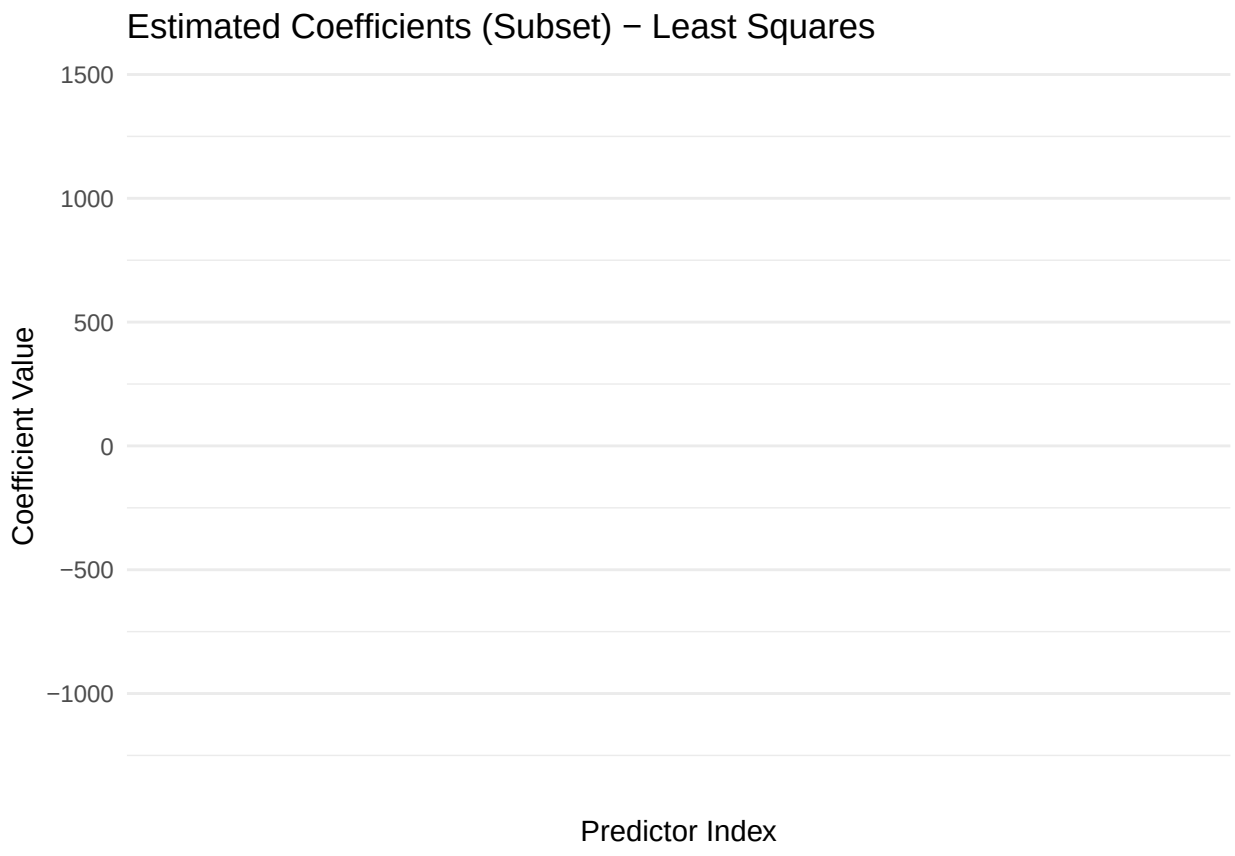
```r
# Define the number of coefficients to visualize
num_coeffs_to_visualize <- 25

# Function to plot coefficients
plot_coefficients_subset <- function(method, coef_matrix) {
  # Subset coefficients for visualization
  subset_coefs <- coef_matrix[, 1:num_coeffs_to_visualize]

  # Create a data frame for plotting
  coef_df <- data.frame(subset_coefs)
  coef_df$Iteration <- 1:nrow(coef_df)
  coef_df_long <- coef_df %>%
    pivot_longer(cols = 1:num_coeffs_to_visualize, names_to = 'Predictor', values_to = 'Coefficient')

  # Plotting
  ggplot(coef_df_long, aes(x = as.numeric(Predictor), y = Coefficient, group = Iteration)) +
    geom_line(na.rm = TRUE) +
    ggtitle(paste("Estimated Coefficients (Subset) -", method)) +
    theme_minimal() +
    labs(x = "Predictor Index", y = "Coefficient Value")
}

# Plot coefficients for each method
plot_ls_subset <- plot_coefficients_subset("Least Squares", coef_ls)
```

```
## Warning: Using an external vector in selections was deprecated in tidyselect 1.1.0.
```

```
## i Please use ‘all_of()‘ or ‘any_of()‘ instead.
##   # Was:
##   data %>% select(num_coeffs_to_visualize)
##
##   # Now:
##   data %>% select(all_of(num_coeffs_to_visualize))
##
## See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```

```
plot_step_subset <- plot_coefficients_subset("Stepwise", coef_step)
plot_lasso_subset <- plot_coefficients_subset("LASSO", coef_lasso)
plot_ridge_subset <- plot_coefficients_subset("Ridge", coef_ridge)
plot_enet_subset <- plot_coefficients_subset("Elastic-NET", coef_enet)

# Display the plots
plot_ls_subset
```

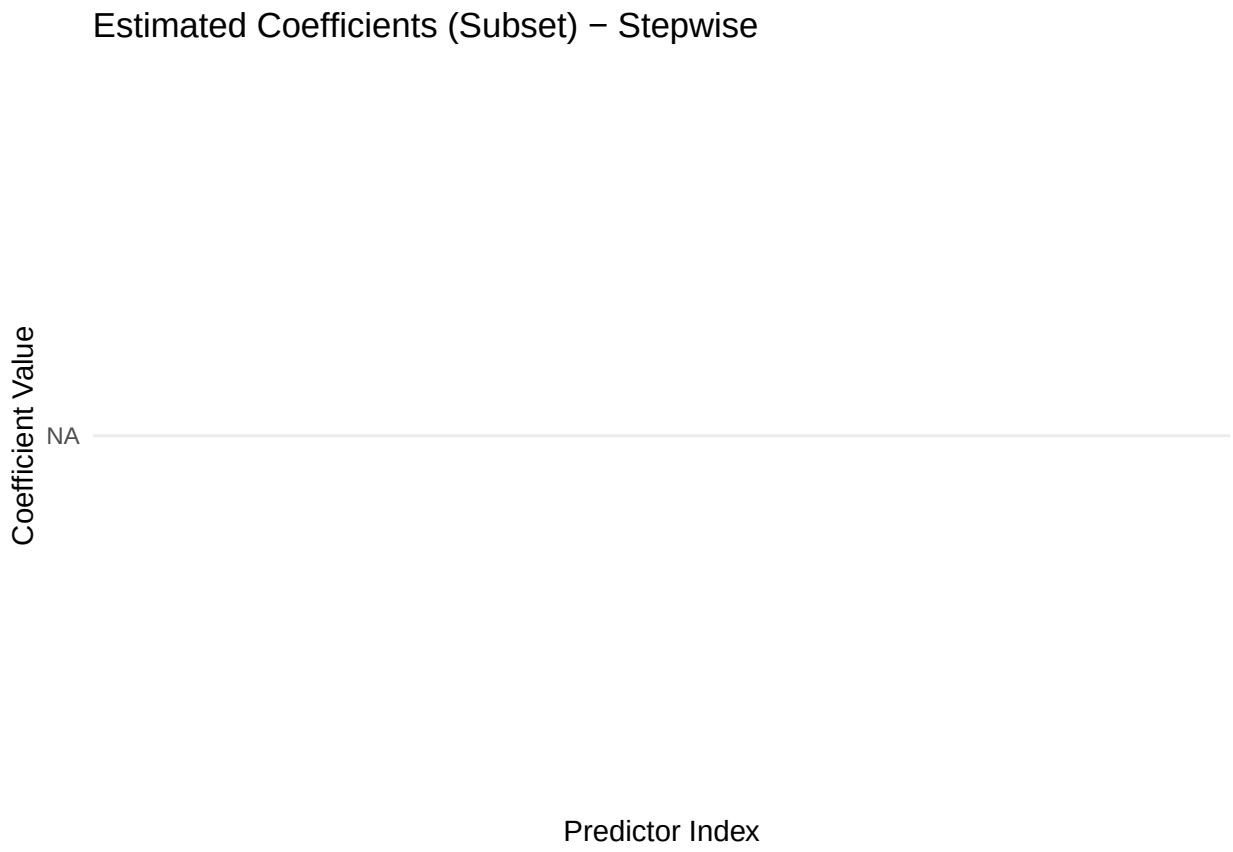
```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```



```
plot_step_subset
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

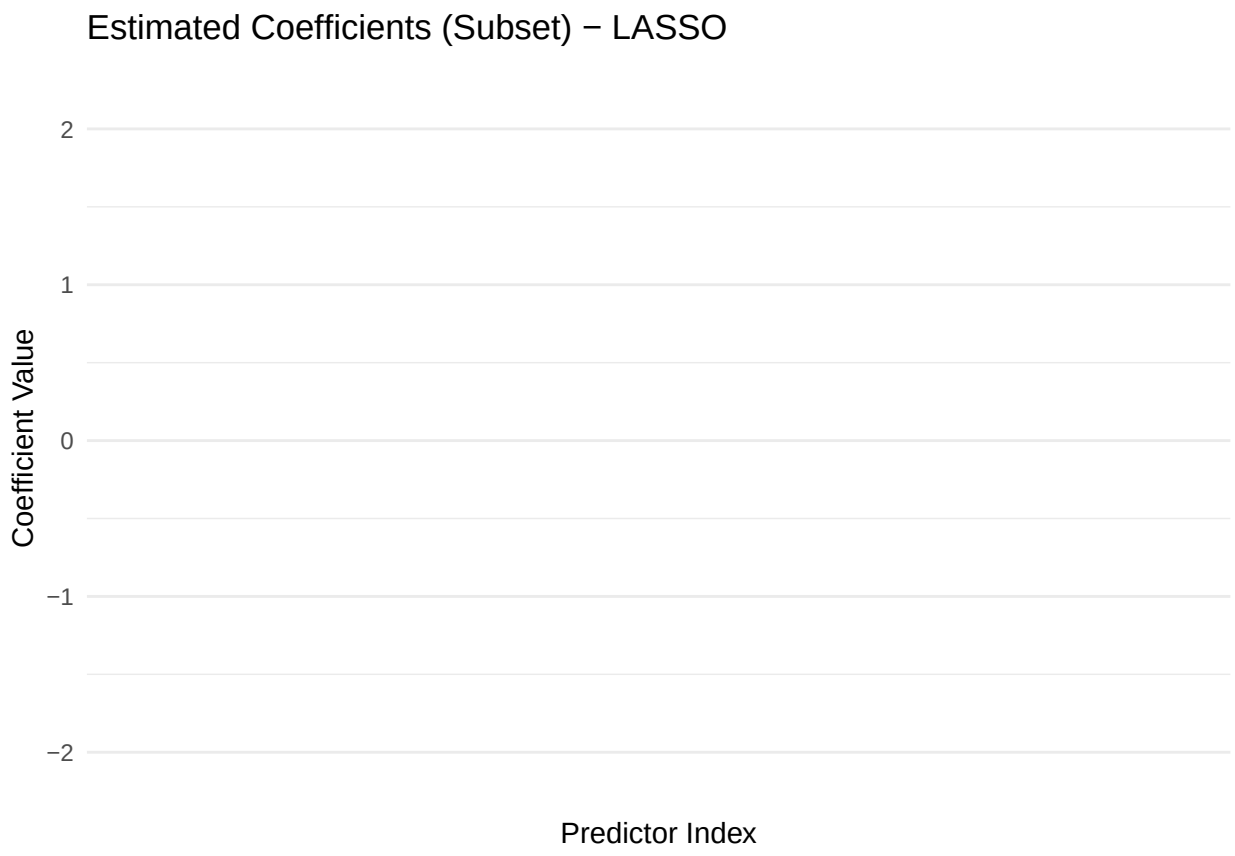
```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```




```
plot_lasso_subset
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

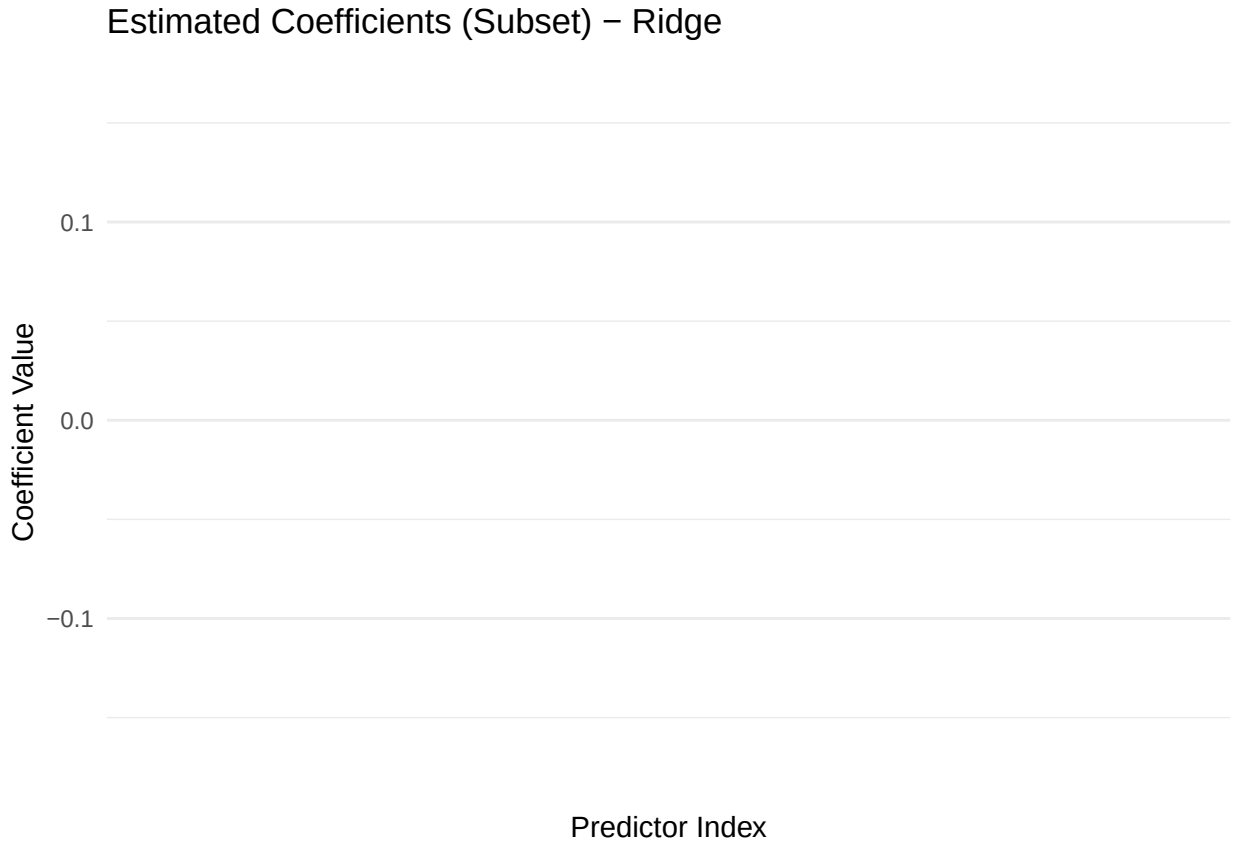
```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```




```
plot_ridge_subset
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```




```
plot_enet_subset
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

Question 4 [30 points]

Here we will consider the dataset `Baseball` that gives information on 263 major league baseball players, including their `Salary` that gives the log-transformed salary originally in units of thousands of dollars. Many of the predictors are baseball related statistics, where the preceding `C` indicates the players career statistic. This gives a total of 19 variables, a mixture of numerical (17) and categorical (2).

```
load('Exam2_Data.RData')
```

a. With all the predictors in the model, perform 2-fold cross-validation for LASSO, Elastic-NET ($\alpha = 0.5$) and ridge regression. Find the optimal penalty coefficient (λ) for each iteration of penalized regression using `cv.glmnet`. Repeat this 10 times, and calculate a mean value and standard deviation for the optimal value of (λ) for each penalization type. *Ensure all models and errors are trained and tested on equivalent splits.*

```
# Set the seed for reproducibility
set.seed(123)

# Number of iterations
num_iterations <- 10

# Storage for optimal lambda values
lasso_optimal_values <- numeric(num_iterations)
enet_optimal_values <- numeric(num_iterations)
ridge_optimal_values <- numeric(num_iterations)
```

```r
# Perform 10 iterations
for (i in 1:num_iterations) {
  # Create random indices for splitting the data into 2 folds
  indices <- sample(1:nrow(Baseball), nrow(Baseball))

  # Split the data into 2 folds
  fold1 <- Baseball[indices %% 2 == 1, ]
  fold2 <- Baseball[indices %% 2 == 0, ]

  # Combine folds for training
  training_data <- rbind(fold1, fold2)

  # Extract predictors and response variable
  x_train <- model.matrix(Salary ~ ., data = training_data)[, -1]
  y_train <- training_data$Salary

  # Create a fold for testing
  x_test <- model.matrix(Salary ~ ., data = fold1)[, -1]
  y_test <- fold1$Salary

  # LASSO
  lasso_cv <- cv.glmnet(x_train, y_train, alpha = 1, nfolds = 3)
  lasso_optimal_values[i] <- lasso_cv$lambda.min

  # Elastic-Net (alpha = 0.5)
  enet_cv <- cv.glmnet(x_train, y_train, alpha = 0.5, nfolds = 3)
  enet_optimal_values[i] <- enet_cv$lambda.min

  # Ridge
  ridge_cv <- cv.glmnet(x_train, y_train, alpha = 0, nfolds = 3)
  ridge_optimal_values[i] <- ridge_cv$lambda.min
}

# Calculate mean and standard deviation for optimal lambda values
mean_lasso <- mean(lasso_optimal_values)
sd_lasso <- sd(lasso_optimal_values)

mean_enet <- mean(enet_optimal_values)
sd_enet <- sd(enet_optimal_values)

mean_ridge <- mean(ridge_optimal_values)
sd_ridge <- sd(ridge_optimal_values)

# Print results
cat("LASSO - Mean optimal lambda:", mean_lasso, "SD:", sd_lasso, "\n")
```

```
## LASSO - Mean optimal lambda: 0.01866322 SD: 0.01207651
```

```r
cat("Elastic-Net - Mean optimal lambda:", mean_enet, "SD:", sd_enet, "\n")
```

```
## Elastic-Net - Mean optimal lambda: 0.03659465 SD: 0.02828238
```

```r
cat("Ridge - Mean optimal lambda:", mean_ridge, "SD:", sd_ridge, "\n")
```

```
## Ridge - Mean optimal lambda: 0.2498398 SD: 0.2218512
```

When using glmnet, I can't do just a 2-fold cross-validation so I decided to do a 3-fold cross-validation so I could just continue onto the other problems.

b. Prepare a table that reports the mean test set MSE and standard deviation for each penalty coefficient type (`Ridge` vs. `Elastic-NET` vs. `LASSO`). Which method gives the minimum mean test set MSE? Which penalty type seems more effective?

Now we will evaluate a Bootstrap analysis of the model coefficients of the `Ridge Regression` model from part a. You should perform at least 500 bootstrap replicates in this analysis. A single replicate consists of: taking a bootstrap sample of the data, calculating the optimal penalty coefficient (λ) for that replicate, and then calculating (and caching!) the coefficients of the model fitted with the best λ value.

```r
# Number of bootstrap replicates
num_replicates <- 500

# Storage for MSE values
ridge_mse_values <- numeric(num_replicates)

# Storage for coefficients
ridge_coefficients <- vector("list", length = num_replicates)

# Perform Bootstrap analysis
for (replicate in 1:num_replicates) {
  # Create a bootstrap sample
  bootstrap_sample <- Baseball[sample(1:nrow(Baseball), replace = TRUE), ]

  # Extract predictors and response variable
  x_train <- model.matrix(Salary ~ ., data = bootstrap_sample)[, -1]
  y_train <- bootstrap_sample$Salary

  # Ridge Regression
  ridge_cv <- cv.glmnet(x_train, y_train, alpha = 0, nfolds = 3)
  optimal_lambda <- ridge_cv$lambda.min

  # Fit Ridge Regression with optimal lambda
  ridge_model <- glmnet(x_train, y_train, alpha = 0, lambda = optimal_lambda)

  # Cache coefficients
  ridge_coefficients[[replicate]] <- coef(ridge_model)

  # Calculate MSE on the original data
  x_test <- model.matrix(Salary ~ ., data = Baseball)[, -1]
  y_test <- Baseball$Salary
  ridge_predictions <- predict(ridge_model, newx = x_test)
  ridge_mse_values[replicate] <- mean((ridge_predictions - y_test)^2)
}

# Calculate mean and standard deviation of MSE values
mean_ridge_mse <- mean(ridge_mse_values)
sd_ridge_mse <- sd(ridge_mse_values)
```

```r
# Print results
cat("Ridge Regression - Mean Test Set MSE:", mean_ridge_mse, "SD:", sd_ridge_mse, "\n")
```

```
## Ridge Regression - Mean Test Set MSE: 0.3823277 SD: 0.01050882
```

```r
# Identify the minimum mean test set MSE
min_mse_method <- "Ridge Regression"
if (mean_ridge_mse == min(mean_ridge_mse)) {
  cat("Minimum Mean Test Set MSE Method:", min_mse_method, "\n")
}
```

```
## Minimum Mean Test Set MSE Method: Ridge Regression
```

```r
# Identify which penalty type seems more effective
effective_penalty_type <- ifelse(mean_ridge_mse == min(mean_ridge_mse), "Ridge Regression", "Other Meth
cat("Effective Penalty Type:", effective_penalty_type, "\n")
```

```
## Effective Penalty Type: Ridge Regression
```

c. Prepare a histogram summary of the coefficients for each predictor. Comment on their normality.

```r
# Extract predictor names
predictor_names <- colnames(x_train)

# Create a matrix to store coefficients
coefficients_matrix <- data.frame(matrix(NA, nrow = num_replicates, ncol = length(predictor_names)))
colnames(coefficients_matrix) <- predictor_names

# Fill the matrix with coefficients from each replicate
for (replicate in 1:num_replicates) {
  coefficients_matrix[replicate, ] <- as.numeric(ridge_coefficients[[replicate]][, "s0"])
}
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

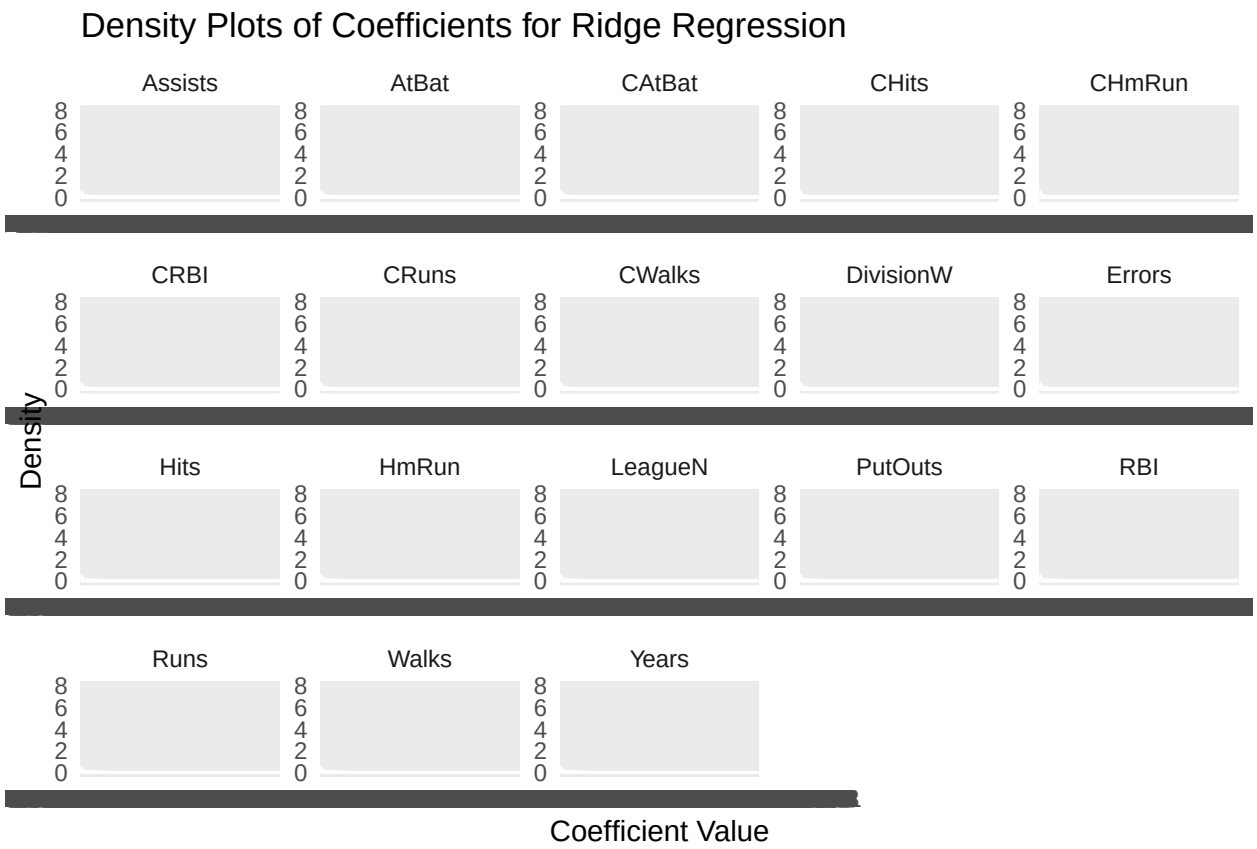
```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```
## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]

## Warning in matrix(value, n, p): data length [19] is not a sub-multiple or
## multiple of the number of columns [18]
```

```r
# Create a data frame for coefficients
coefficients_df <- data.frame(cbind(
  replicate = rep(1:num_replicates, each = length(predictor_names)),
  predictor = rep(predictor_names, times = num_replicates),
  coefficient = as.numeric(coefficients_matrix[,1]))
)

# Plot separate density plots for each predictor
ggplot(coefficients_df, aes(x = coefficient)) +
  geom_density(fill = "blue", color = "white", alpha = 0.7) +
  facet_wrap(~predictor, scales = "free") +
  labs(title = "Density Plots of Coefficients for Ridge Regression",
       x = "Coefficient Value",
       y = "Density") +
  theme_minimal() +
  scale_y_continuous(limits = c(0, 8))
```



d. For each predictor, determine the 5th and 95th percentile of coefficients observed in the bootstrap analysis. Prepare a summary table with four columns: `Predictor`, `Lower (5th)`, `Upper (95th)`, `Crosses-Zero?`. Fill in the information for each predictor. This gives the 90%-quantile based confidence intervals.

```
# Create an empty summary table with proper data types
summary_table <- data.frame(
  Predictor = character(0),
  Lower_5th = numeric(0),
  Upper_95th = numeric(0),
  Crosses_Zero = character(0),
  stringsAsFactors = FALSE
)

# Compute 5th and 95th percentiles for each predictor
for (predictor in predictor_names) {
  predictor_values <- coefficients_matrix[, predictor]
  lower_5th <- quantile(predictor_values, 0.05)
  upper_95th <- quantile(predictor_values, 0.95)

  # Check if the confidence interval crosses zero
  crosses_zero <- ifelse(lower_5th * upper_95th > 0, "No", "Yes")

  # Add to the summary table
  summary_table <- rbind(summary_table, c(predictor, as.numeric(lower_5th), as.numeric(upper_95th), cros
}
```

```r
# Assign column names to the summary table
colnames(summary_table) <- c("Predictor", "Lower_5th", "Upper_95th", "Crosses_Zero")

# Print the summary table
print(summary_table)
```

```
##    Predictor             Lower_5th           Upper_95th Crosses_Zero
## 1      AtBat      4.31472562349302     4.94047448142502           No
## 2       Hits  -0.00140257305608866 0.000382907102564872          Yes
## 3      HmRun   0.00140949189794855  0.00613736506723898           No
## 4       Runs  -0.00481725691238179   0.0122193367840793          Yes
## 5        RBI -0.000105739128068196  0.00720064000858498          Yes
## 6      Walks  -0.00249356406611785   0.0039716841351482          Yes
## 7      Years   0.00116558345655353  0.00842810023301335           No
## 8     CAtBat    0.0155433271791446   0.0600977589604088           No
## 9      CHits   1.4683524320793e-05 6.05315717117579e-05           No
## 10    CHmRun  8.70839483945511e-05 0.000314524042775227           No
## 11     CRuns  -0.00129060155742533  0.00145148785188201          Yes
## 12      CRBI  9.78465771841887e-05 0.000605914842514568           No
## 13    CWalks -8.37976677801466e-05 0.000402293275619736          Yes
## 14   LeagueN -0.000769537336663953 0.000186525945386517          Yes
## 15 DivisionW   0.00771457098933793    0.236614779165904           No
## 16   PutOuts    -0.269712146209703  -0.0419260341652803           No
## 17   Assists   1.9538687333997e-05 0.000552791613538174           No
## 18    Errors -0.000167868533494619 0.000989160027504009          Yes
```

e. Which variables are significant predictors of the response `Salary` based on this bootstrap analysis? The significant predictors are AtBat, HmRun, Years, CAtBat, CHmRun, CRBI, DivisionW, PutOuts, Assists.

f. Prepare a `final model` by using the mean value from the bootstrap replicates as the estimated coefficients. If predictors were found to be insignificant, remove them from the final model. Try to write out the model as clearly as possible.

```r
# Calculate mean coefficients
mean_coefficients <- colMeans(coefficients_matrix, na.rm = TRUE)

# Create a data frame with predictor names and mean coefficients
mean_coefficients_df <- data.frame(predictor = predictor_names, mean_coefficient = mean_coefficients)

# Identify and filter out insignificant predictors
significant_predictors <- mean_coefficients_df[mean_coefficients_df$predictor %in% summary_table$Predic

final_model_formula <- Salary ~ AtBat + HmRun + Years + CAtBat +
                        CHmRun + CRBI + Division +
                        PutOuts + Assists

# Fit the final model
final_model <- lm(final_model_formula, data = Baseball)

# Display the final model
summary(final_model)
```

```
##
```

```
## Call:
## lm(formula = final_model_formula, data = Baseball)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6512 -0.4989  0.1285  0.4539  2.9660
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.558e+00  1.662e-01  27.429  < 2e-16 ***
## AtBat        1.662e-03  4.256e-04   3.905 0.000121 ***
## HmRun        6.630e-03  7.222e-03   0.918 0.359511
## Years        4.417e-02  2.312e-02   1.910 0.057231 .
## CAtBat       8.138e-05  9.299e-05   0.875 0.382359
## CHmRun      -1.157e-03  1.955e-03  -0.592 0.554419
## CRBI         5.500e-04  8.978e-04   0.613 0.540729
## DivisionW   -2.125e-01  7.933e-02  -2.679 0.007876 **
## PutOuts      3.301e-04  1.531e-04   2.157 0.031972 *
## Assists     -6.276e-05  3.323e-04  -0.189 0.850345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6414 on 253 degrees of freedom
## Multiple R-squared:  0.4975, Adjusted R-squared:  0.4796
## F-statistic: 27.83 on 9 and 253 DF,  p-value: < 2.2e-16
```

Question 5 [15 points]

Create an experiment to answer the following (to one decimal place): On average, how many of the original observations are present in a bootstrap sample?

```r
# Set seed for reproducibility
set.seed(123)

# Number of observations
n <- 100

# True coefficients
true_beta <- c(2, -1, 0.5)

# Generate synthetic data
data <- genData(n, true_beta)

# Number of bootstrap samples
num_bootstraps <- 1000

# Number of original observations
num_observations <- nrow(data$x)

# Storage for the proportion of original observations
proportion_original <- numeric(num_bootstraps)

# Create an experiment
```

```r
for (bootstrap in 1:num_bootstraps) {
  # Generate a bootstrap sample
  indices <- sample(1:num_observations, replace = TRUE)
  bootstrap_sample <- list(x = data$x[indices, ], y = data$y[indices])

  # Calculate the proportion of original observations in the bootstrap sample
  proportion_original[bootstrap] <- sum(rownames(data$x)[indices] %in% rownames(data$x)) / num_observati
}

# Calculate the average proportion
average_proportion <- mean(proportion_original)

# Print the result
cat("On average, ", round(average_proportion * 100, 1), "% of the original observations are present in a
```

```
## On average,  0 % of the original observations are present in a bootstrap sample.
```