

Question_2

Akash Gupta

2023-12-07

Question 2

First we will fit the Linear regression model.

```
library(faraway)
library(MASS)

# Load the dataset
data(savings)
# View the structure of the dataset
str(savings)

## 'data.frame':    50 obs. of  5 variables:
## $ sr      : num  11.43 12.07 13.17 5.75 12.88 ...
## $ pop15: num  29.4 23.3 23.8 41.9 42.2 ...
## $ pop75: num  2.87 4.41 4.43 1.67 0.83 2.85 1.34 0.67 1.06 1.14 ...
## $ dpi     : num  2330 1508 2108 189 728 ...
## $ ddpi    : num  2.87 3.93 3.82 0.22 4.56 2.43 2.67 6.51 3.08 2.8 ...

# (1) Fit a linear regression model
model <- lm(sr ~ pop15 + pop75 + dpi + ddpi, data = savings)
model

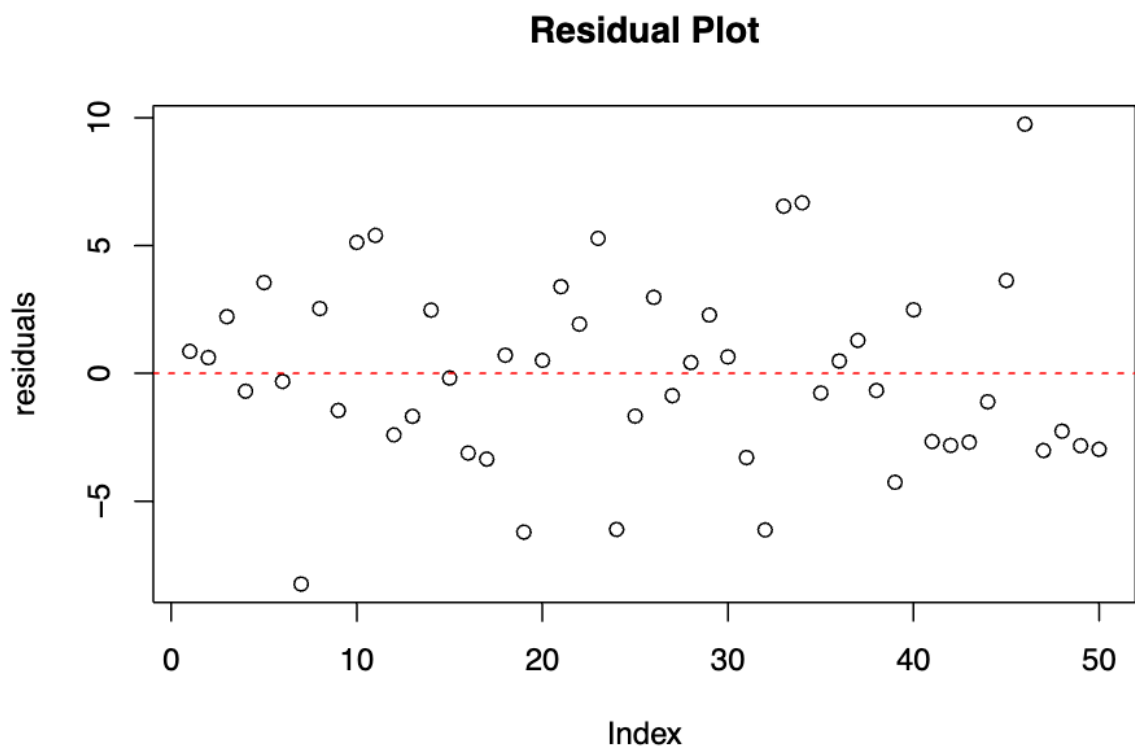
##
## Call:
## lm(formula = sr ~ pop15 + pop75 + dpi + ddpi, data = savings)
##
## Coefficients:
## (Intercept)      pop15      pop75        dpi        ddpi
##  28.5660865   -0.4611931   -1.6914977   -0.0003369    0.4096949

summary(model)

##
## Call:
## lm(formula = sr ~ pop15 + pop75 + dpi + ddpi, data = savings)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -8.2422 -2.6857 -0.2488  2.4280  9.7509
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 28.5660865  7.3545161   3.884 0.000334 ***
## pop15      -0.4611931  0.1446422  -3.189 0.002603 **
## pop75      -1.6914977  1.0835989  -1.561 0.125530
## dpi        -0.0003369  0.0009311  -0.362 0.719173
## ddpi        0.4096949  0.1961971   2.088 0.042471 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.803 on 45 degrees of freedom
## Multiple R-squared:  0.3385, Adjusted R-squared:  0.2797
## F-statistic: 5.756 on 4 and 45 DF,  p-value: 0.0007904
```

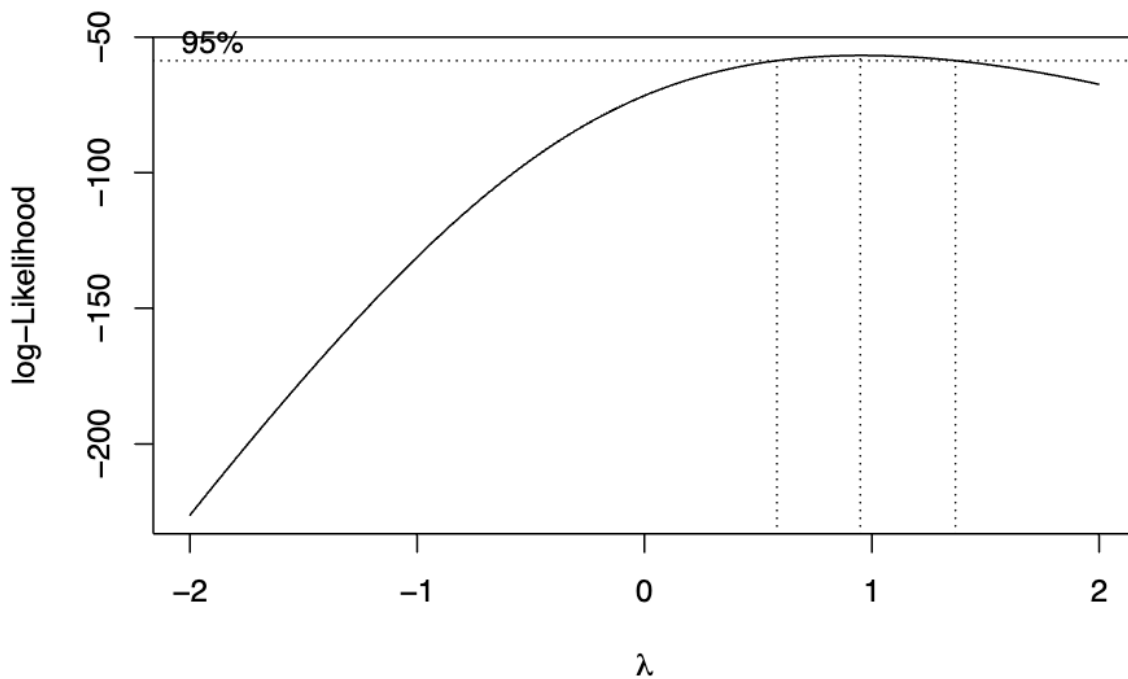
```
#Examining the residuals
residuals <- residuals(model)
plot(residuals)
abline(h = 0, col = "red", lty = 2)
title("Residual Plot")
```



residuals are as follows: Min -8.2422
 1Q -2.6857
 Median -0.2488
 3Q 2.4280
 Max 9.7509

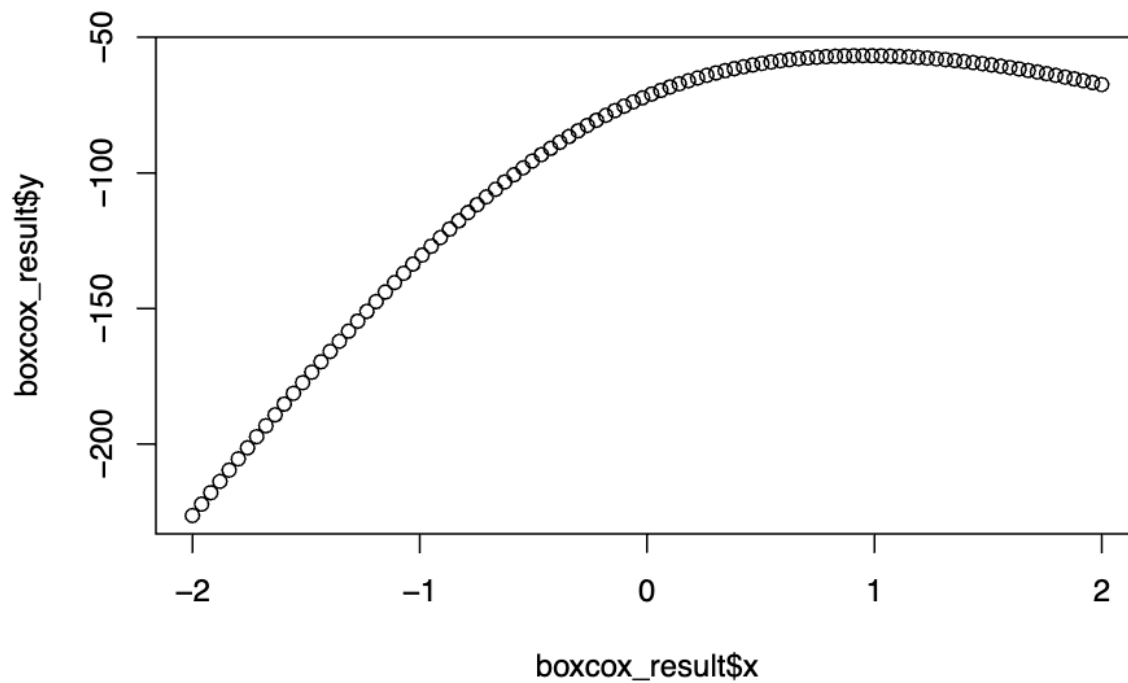
The

Including Plots



Optimal lambda: 0.9494949

Box-Cox Transformation



The optimal lambda comes out to be: 0.9494949

##Part(c): The main purpose of the Box-Cox plot is to help us identify the value of lambda that maximizes the log-likelihood. The optimal lambda corresponds to the peak of the plot, and it indicates the best power

transformation for stabilizing the variance and improving the fit of the linear regression model. The plot provides a visual representation of how the log-likelihood changes across different lambda values. Examining the shape of the curve can give you insights into the nature of the transformation and its impact on the model's goodness of fit.