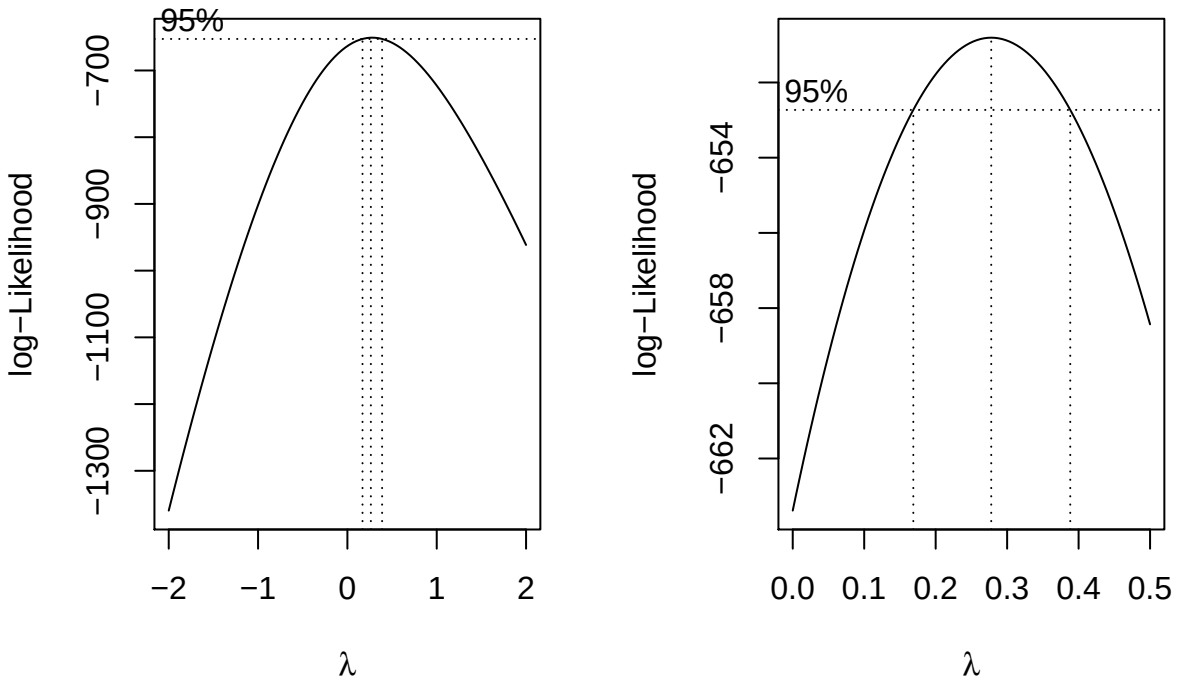
HW6 Solutions

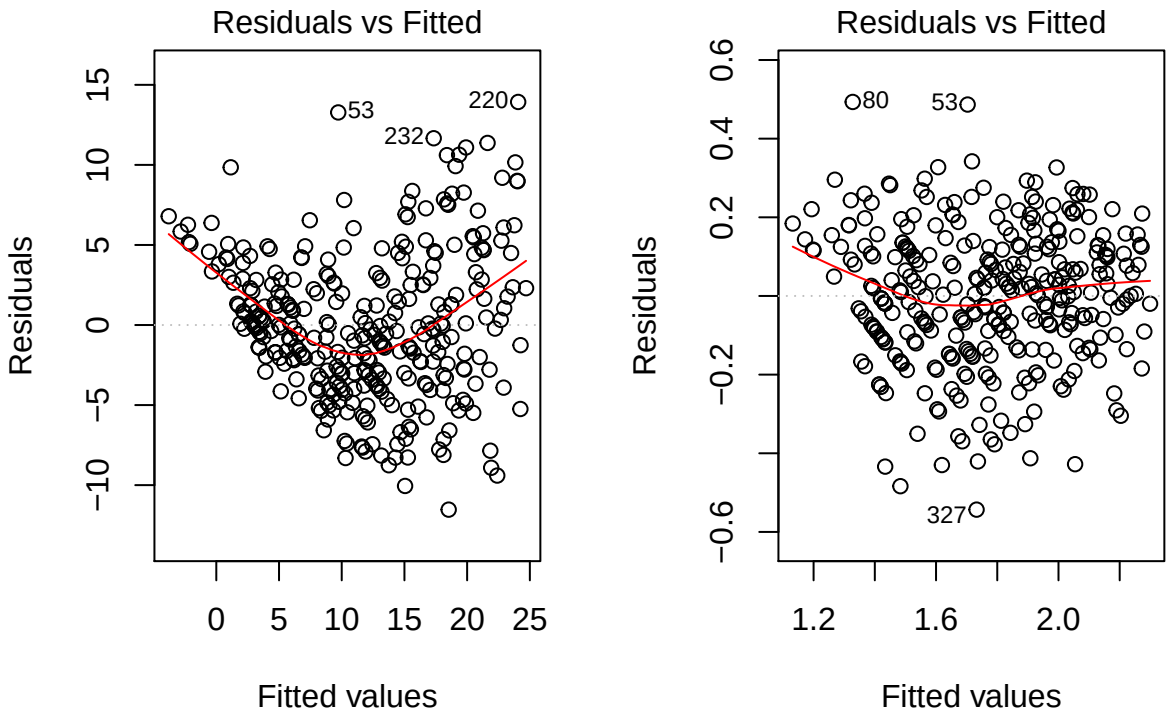
Problem 9.3

```
lmod <- lm(O3 ~ temp + humidity + ibh, data = ozone)
par(mfrow = c(1,2))
boxcox(lmod, plotit = TRUE)
boxcox(lmod, plotit = TRUE, lambda = seq(0,0.5, by = 0.01))
```



The log-Likelihood is maximized at $\lambda = 0.26 \approx \frac{1}{4}$.

```
lmod.transformed <- lm((O3^0.25) ~ temp + humidity + ibh, data = ozone)
par(mfrow = c(1,2))
plot(lmod, which = 1)
plot(lmod.transformed, which = 1)
```

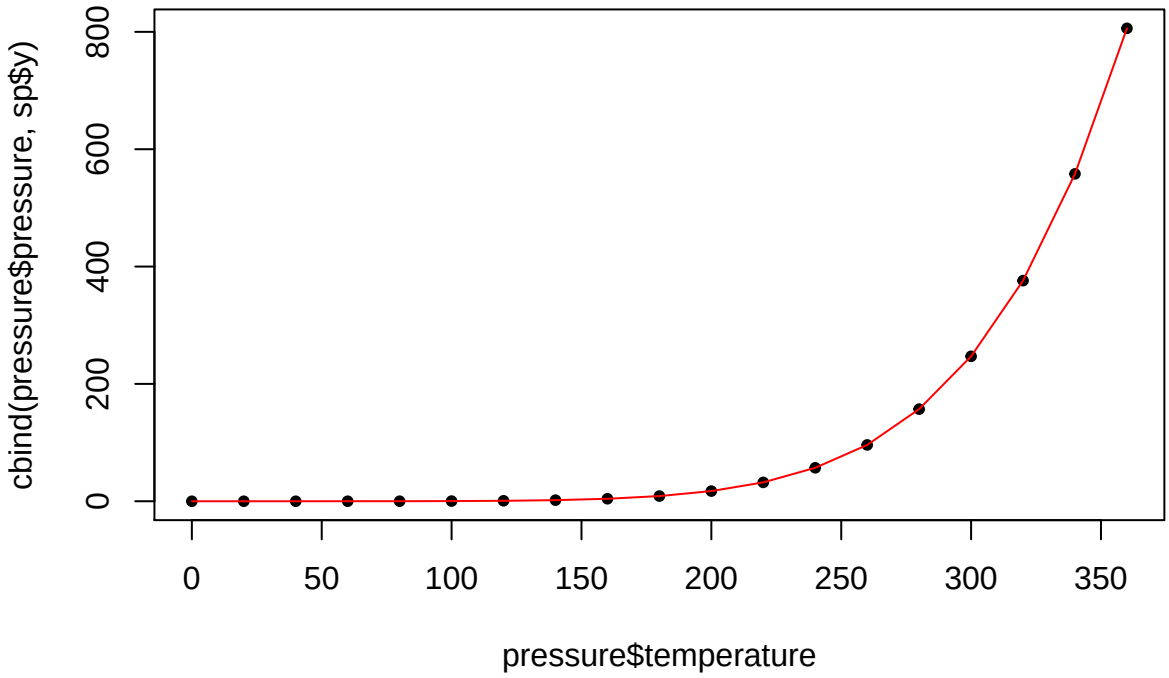


It can be seen that the residual-vs-fitted plot has greatly improved for the transformed model; more specifically, the non-constant variance and non-random structure issues present in the original residuals have been resolved (to some extent) by the transformation. Also, the R^2 increases from 0.68 to 0.72. (Note: Even if we had used the exact maximizer $\lambda = 0.26$, the results would essentially be the same. For instance the R^2 would still be 0.72.)

Problem 9.4

A smoothing spline gives a nice fit:

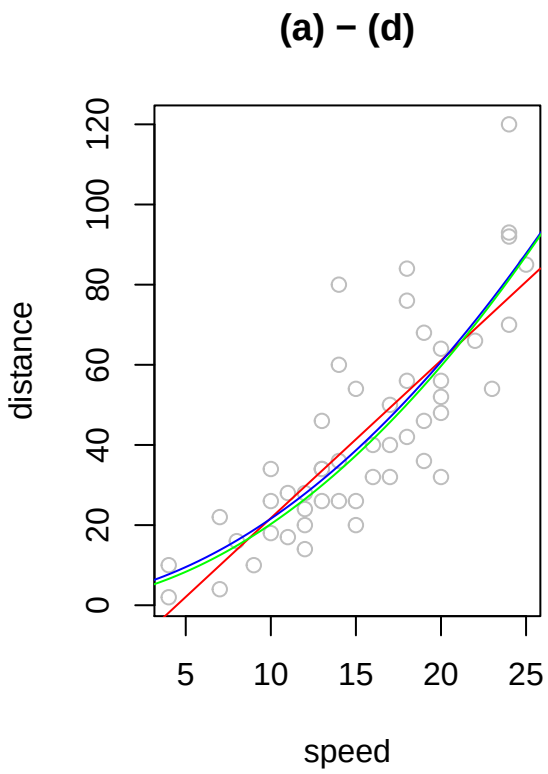
```
data(pressure)
sp <- with(pressure,smooth.spline(temperature, pressure))
matplot(pressure$temperature, cbind(pressure$pressure, sp$y), type = 'pl',lty=1, pch=20, col = c(1,2))
```

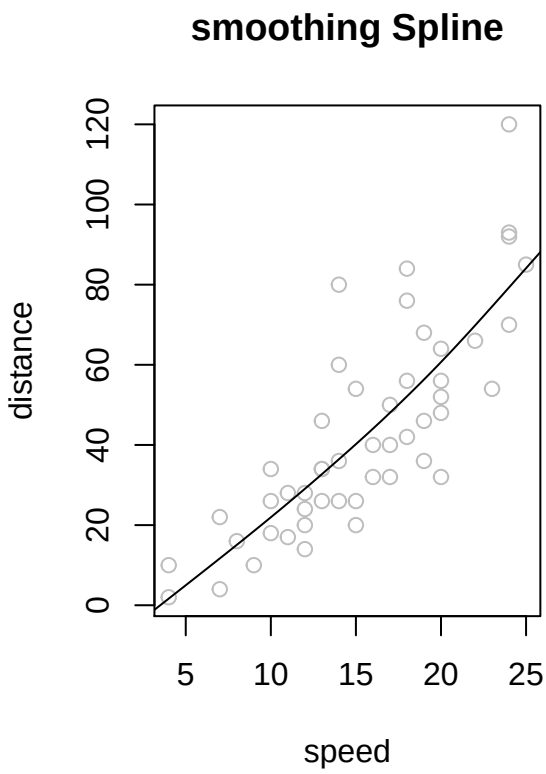


Problem 9.8

```
par(mfrow = c(1,2))
data(cars)
plot(cars$speed, cars$dist, col = 'gray', main = '(a) - (d)', ylab = 'distance', xlab = 'speed')
lmod = lm(dist~speed, data = cars)
abline(lmod, col = 'red')
qmod <- lm(dist ~ speed + I(speed^2), data = cars)
x = data.frame(seq(3,26, by = 0.1))
colnames(x) <- 'speed'
lines(cbind(x, predict(qmod, x)), col = 'blue')
sqrtmod <- lm(sqrt(dist) ~ speed, data = cars)
lines(x$speed, predict(sqrtmod, x)^2, col = 'green')

spmod <- smooth.spline(cars$speed, cars$dist)
result = predict(spmod, x$speed)
plot(cars$speed, cars$dist, col = 'gray',main = "smoothing Spline", ylab = 'distance', xlab = 'speed')
lines(result$x, result$y, type = 'l' )
```





The (default) smoothing spline is slightly curved but the curvature is smaller than the quadratic model. The R^2 for the spline is 0.663 which is slightly smaller than the one for the quadratic model. The R^2 for the linear model is 0.651.