Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

1. Fuel price as a function of time[see figure 1]

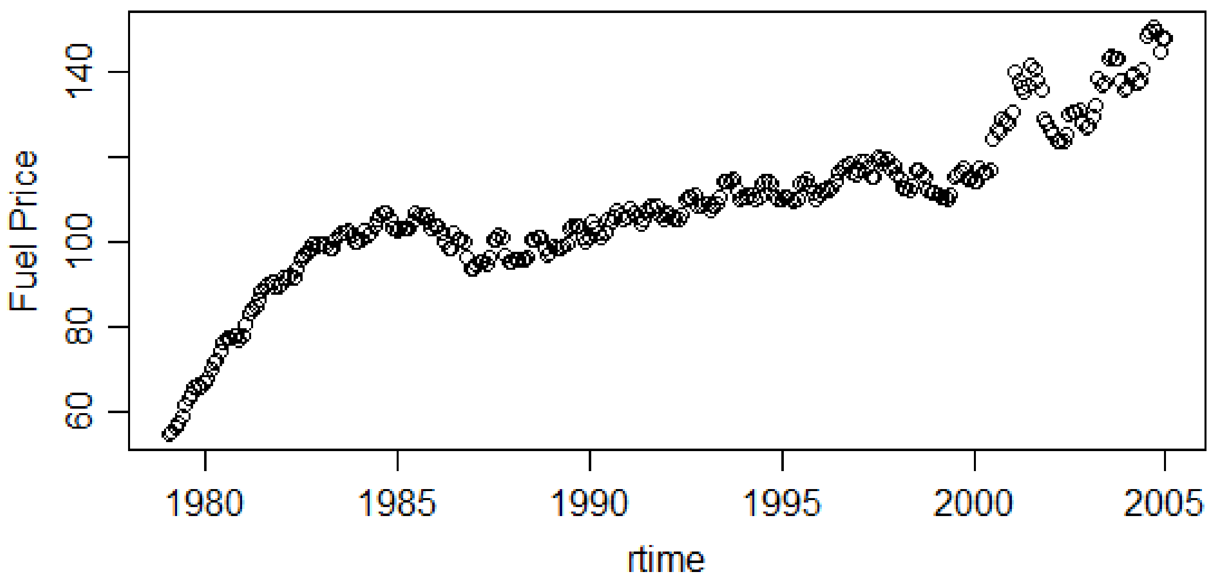


Figure 1 fuel price as a function of time

2. It is reasonable to estimate the mean value and the standard deviation. Because the data is roughly normal distributed, as the normal qq plot [see figure 2] demonstrates that the there is a high density in the middle (the mean), the farther away from the mean, the less data are distributed.

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

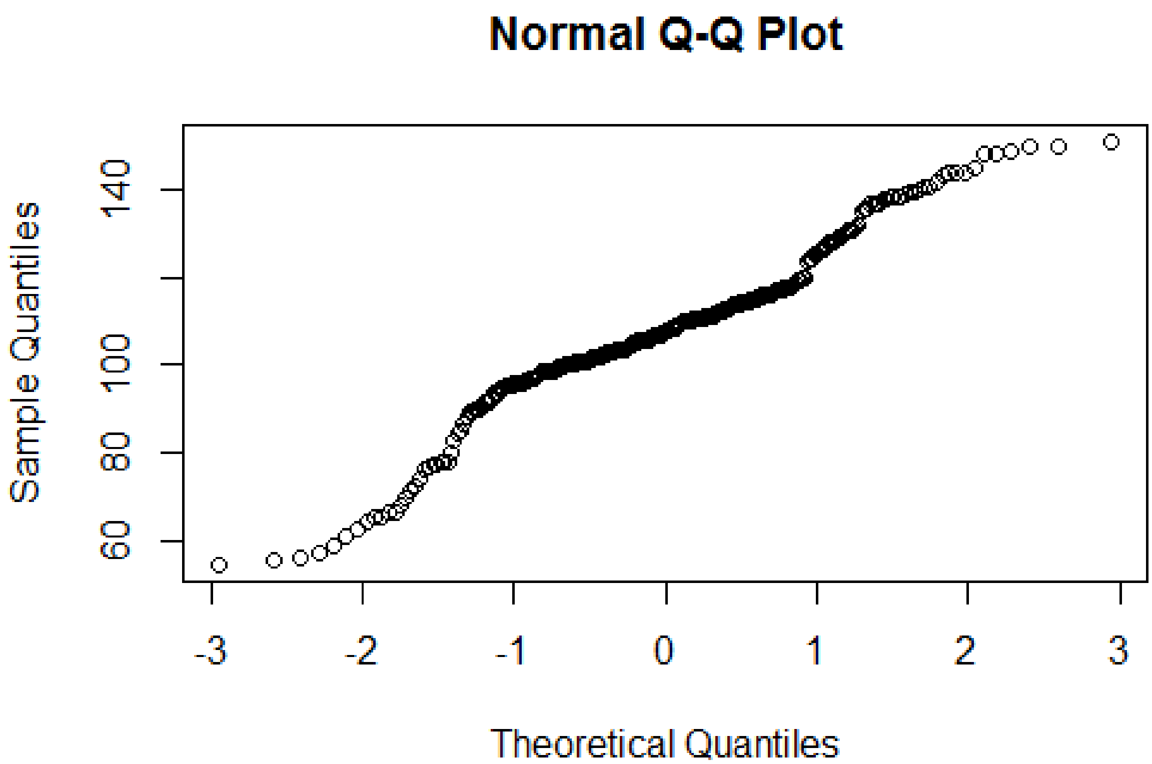


Figure 2 Normal QQ plot

3. GLM

The estimated parameters are $\hat{\theta}_0 = -4.198$, $\hat{\theta}_1 = 2.162$. It is hardly to completely trust this GLM. It is roughly linear, but the period such as before 1985 and after 2000 have considerable residuals.

Q3: OK.

- Missing measure of uncertainty for the parameter estimates.

+ True that residuals are not behaving the same.

Ran Wang
S111503

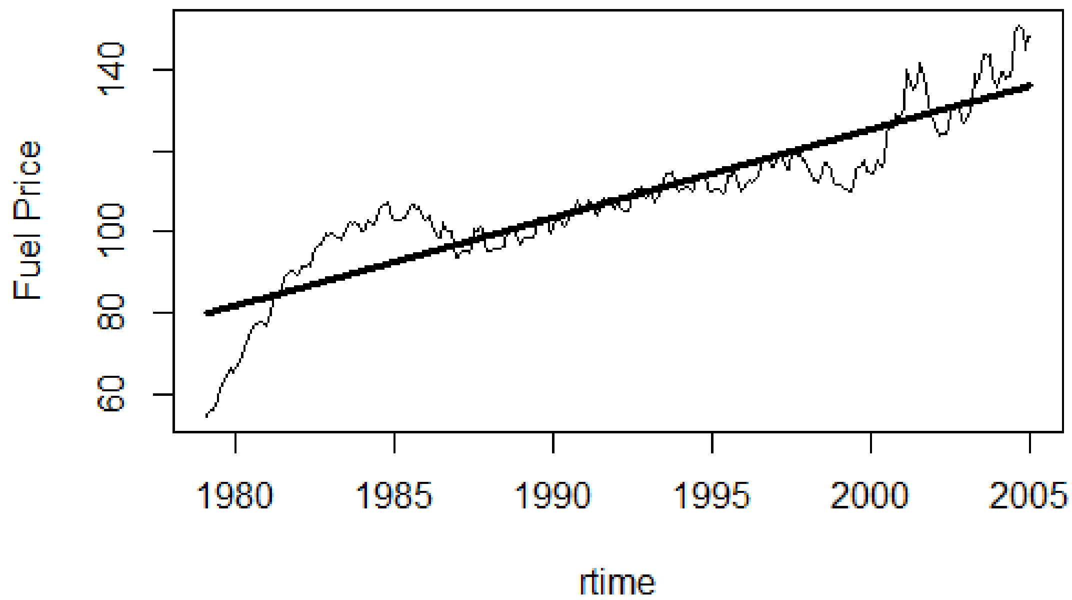


Figure 3 GLM Fitted line

4. Simple exponential smoothing
   The idea of simple exponential smoothing is to reduce the impact of the old observation in an exponential manner (p,51.) The formula used for computing it is from the curriculum book
   $$S_N = (1 - \lambda)Y_N + \lambda S_{N-1}$$
   Where the first observation Y1 can be used as initial value S1 (p,52)
   As the Figure 4 shows the red line smoothed out the details of the observation. Figure 4 illustrates that when the smoothing constant a=0.1, the latest observations do not influence the prediction so much, as the red line counts much more on the older observations, so that it appears very smoothing curve.

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

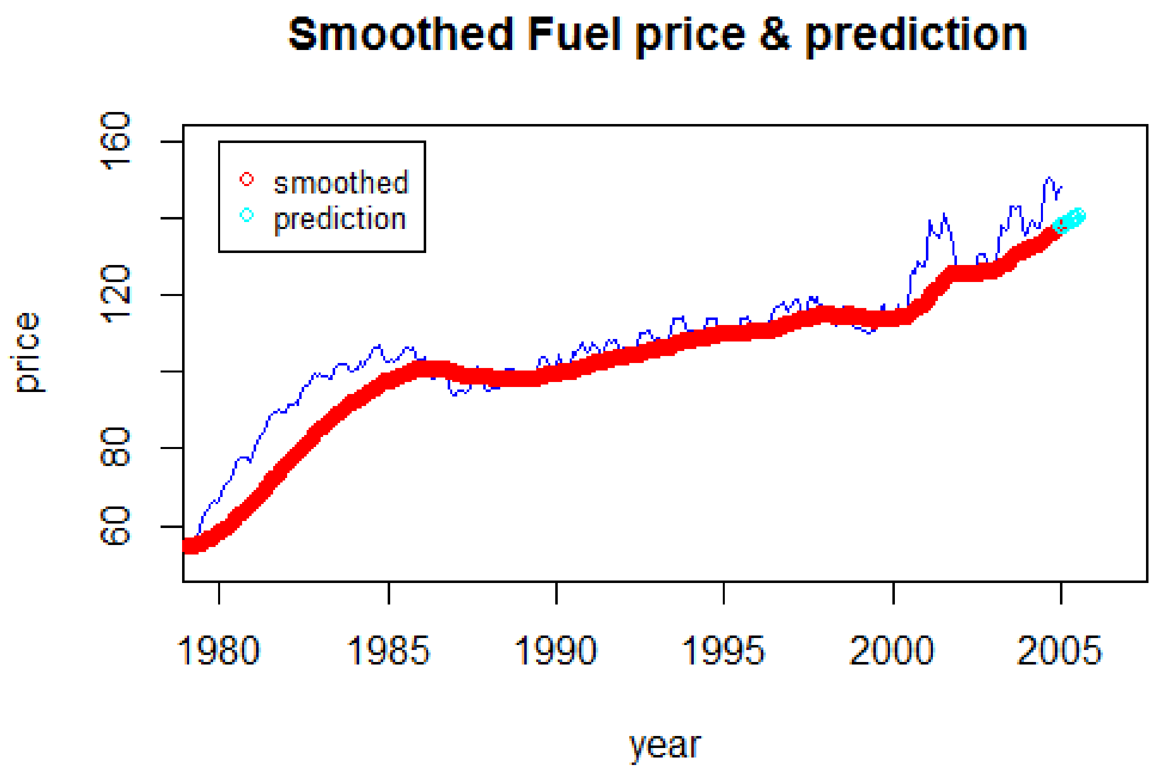


Figure 4 Simple exponential smoothing a=0.1

5. Local trend model

Q4: Mostly fine.
- Predictions are wrong. Should be constant

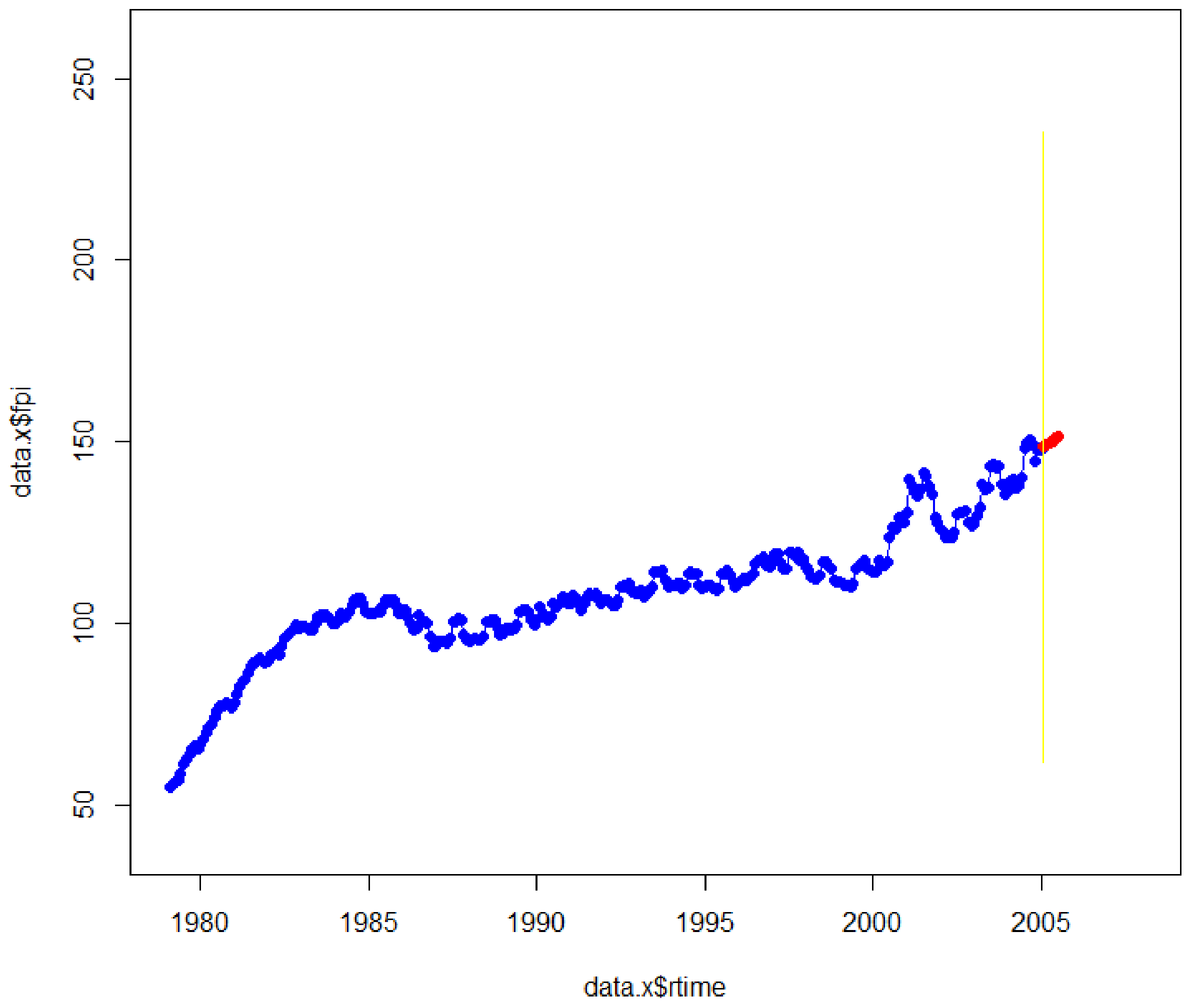
Q5: Mostly fine
+ Estimates are fine
- missing uncertainties of estimates
+ Fine to compare with global model.

Q6: Fine - could have found it more precisely.
This is only considering the last observation - one-step prediction errors are generally better.

Q7: Fine observation - hard to see from the graphs ...
- missing reflection on which of the previous models you find is best.

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index



Estimation

The prediction of local trend model use the same equation as in global trend model, but the estimation procedure differs. So by using this equation:

$$\hat{Y}_{N+l|N} = f^T(l)\hat{\theta}_N$$

The prediction for the first half of 2005 given the observations at N is computed as following:

[1,] 148.6245
[2,] 149.1982
[3,] 149.7719
[4,] 150.3456
[5,] 150.9193
[6,] 151.4931

Local:

$$\hat{\theta} = (x^T \Sigma^{-1}\, x)^{-1}\ \Sigma^{-1}\, x^T\, Y$$

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

$$F_N = \sum_{j=0}^{N-1} \lambda^j f(-j) f^T(-j)$$
$$h_N = \sum_{j=0}^{N-1} \lambda^j f(-j)\, Y_{N-j}$$
$$\widehat{\theta_N} = F_N^{-1}\, h_N$$

Global:
$\hat{\theta} = (x^T\, x)^{-1}\ \ x^T\, Y$

The order j starts from 0, which makes λ^j =1, the corresponding f(0) is the latest observation that has the biggest weight. As j goes up, λ^j will eventually close to 0, and the corresponding observations become older. That is why local trend model has a trend mostly based on the recent observations.

It is easier to see the tendency of the two different trend models by extending them [see figure 5] . The local prediction follows mostly the resent observations, whereas the global follows the trend in general across the whole time period.

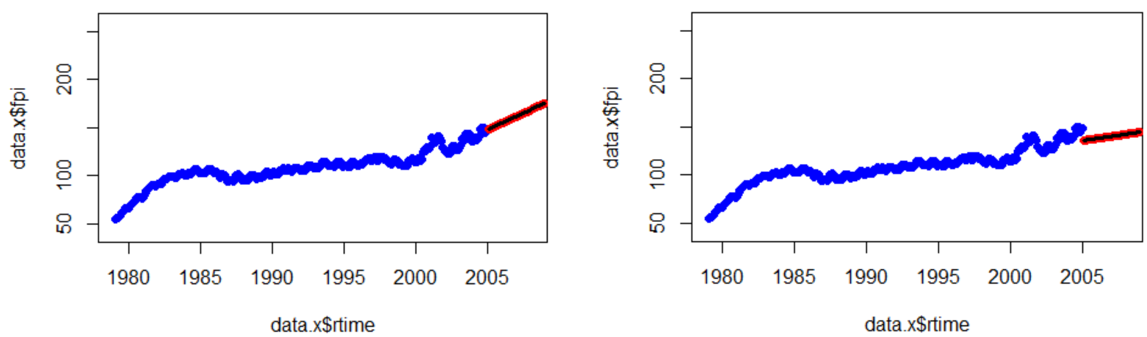


Figure 5 Extended local trend prediction (left) vs. extended global trend prediction

6. Optimal value of forgetting factor.
   Even though Brown(1963) proposed a range for λ^p is between 0.7 and 0.95, where p is the number of the total amount of observations, is it somehow not precise nor confident enough. Therefore, it is rather ideal to choose a forgetting factor by following some methods, thus a value is considered o ptimal if it brings the minimum variance of the prediction error. The variance can be gained by us ing this equation(p57. 3.102)

$$Var[e_N(l)] = \sigma^2[1 + f^T(l)\, F_N^{-1}\, f(l)]$$

There are six lambda values tested, which give six corresponding variances:

λ=0.95	**λ**=0.90	**λ**=0.80	**λ**=0.75	**λ**=0.70	**λ**=0.60
v=59.9081	V=60.0847	60.5522	60.85681	61.2206	62.1797

Obviously, the minimum variance of the prediction error happens when **λ**=0.95, so **λ**=0.95 is the optimal forgetting factor.

7. The interesting finding is that this data follows a seasonal pattern (see figure 5), the 3 quarter always goes up and slows down at the end of a year.

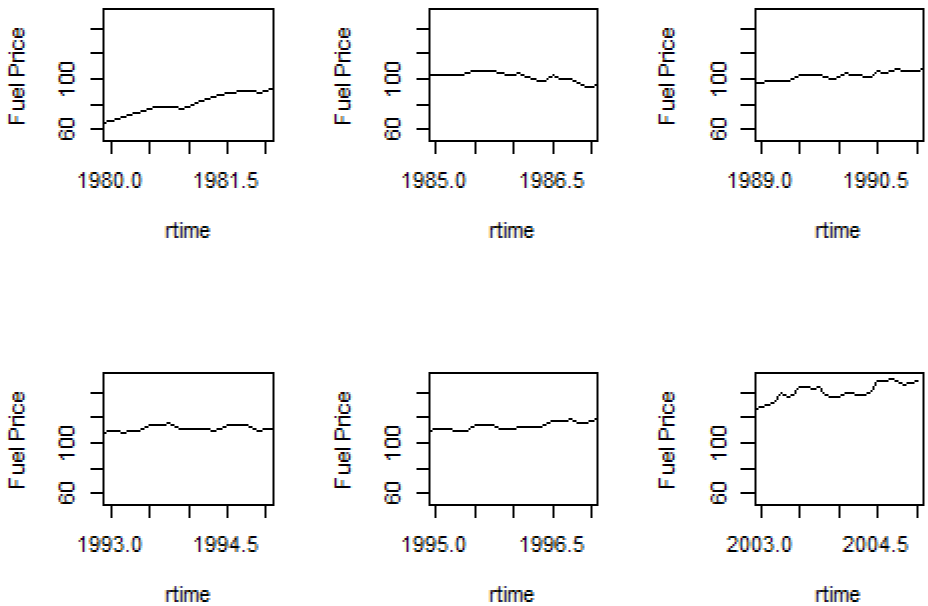


Figure 6 Seasonal Pattern

Appendix

A General linear model

```
setwd("C:\\Users\\RAN\\OneDrive\\Documents\\R")

data.x <- read.table("fuel.txt",header=TRUE)

analysis.x <- lm(fpi~rtime,data=data.x)

summary(analysis.x)

plot(c(1979.083,2008),c(50,250),type="n",main="Fitted line")

plot(data.x$rtime,analysis.x$residuals,ylab='residual',xlab='rtime',xlim=c(1979.083,2008.000),ylim=c(-20,20) )

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l")

lines(data.x$rtime,analysis.x$fitted.values,type="l",lwd=3,xlim=c(1979.083,2008.000),ylim=c(-20,20))
```

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

#simple exponential smoothing

```
n=312

j = 0:(n-1)

#Forgetting factor

l=0.95

#normalizaion constant

c=(1-l)/(1-l^n)

#yhat313 = c* sum((l^j) %*% rev(data.x$fpi))

#smoothing the original data

s=data.x$fpi[1]


for(i in 2:n) {s[i] = (1-l) %*% data.x$fpi[i] + l %*% s[i-1]}

#plot original graph

plot(c(1980,2006.5),c(50,160),type="n",xlab="year", ylab="price",main="Smoothed Fuel price & predic-
tion ")

lines(data.x$fpi ~ data.x$rtime, data= data.x, xlim=c(1979.083,2005.000), ylim=c(40,160),type="l",col=4,
pch=19)

#plot smoothed graph

lines(s ~ data.x$rtime, data= data.x, xlim=c(1979.083,2005.000),  ylim=c(40,160),type="o",col=2,
pch=19 )

#smoothing predictions

for(j in 313:318){s[j] = (1-l) %*% data.x$fpi[312] + l %*% s[j-1]}


o=1:6

#data.x$rtime[313:318] = 2005+o/12

#newrow = cbind(2005,1,2005+1/12,s[313])

#rbind(data.x,newrow)

newrow = c(2005+o/12)

data.frame(rtime = 2005+o/12)

leg.txt <- c("smoothed","prediction")
```

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

```
lines(c(2005,newrow),s[312:318],col=5,type="o")

legend(x=1980,y=160,c("smoothed","prediction"),cex=.8, col=c("red","cyan"),pch=c(1,1))
```

#Local trend model

```
rm(list=ls())

setwd("C:\\Users\\RAN\\OneDrive\\Documents\\R")

data.x <- read.table("fuel.txt",header=TRUE)


n=312

j = 0:(n-1)

l=0.9

rtime<-cbind(1,(-n+1):0)

#Estimation

F <- t(rtime) %*% solve(diag(1/l^rev(j), n, n)) %*% rtime

h <- t(rtime) %*% solve(diag(1/l^rev(j), n, n)) %*% data.x$fpi

theta <- solve(F) %*% h

#the std error of the prediction

sigma <- sqrt( sum((data.x$fpi - rtime%*%theta)^2)/(n-2) )


plot(data.x$rtime,data.x$fpi, ylim=c(40,160),type="o",col=4, pch=19)

plot(data.x$fpi ~ data.x$rtime, data= data.x, xlim=c(1979.083,2008.000), ylim=c(40,260),type="o",col=4,
pch=19 )

lines( rtime%*%theta, col=3)


## Prediction

f <- function(j) return( rbind(1,j) )

j=1:48

#thereom 3.
```

Ran Wang
S111503

Exercise 1. Forecasting the Fuel Price Index

```
hatY <- t(f(j)) %*% theta

points(2005+j/12, hatY, col=2, pch=19,)


Ve <- sigma^2 * (1 + t(f(j)) %*% solve(F) %*% f(j))

sqrt(Ve)

(interval <- hatY + sqrt(Ve) * qt(0.9,n-2) * c(-1,1))


lines(2005+j/12,hatY,type="l",lwd=3)
```

 # Seasonal trend plot

```
par(mfrow=c(1,1))

#seasonal trend

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l")

par(mfrow=c(2,3))

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l",xlim=c(1980,1982))

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l",xlim=c(1985,1987))

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l",xlim=c(1989,1991))

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l",xlim=c(1993,1995))

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l",xlim=c(1995,1997))

plot(data.x$rtime,data.x$fpi,xlab='rtime',ylab='Fuel Price', type="l",xlim=c(2003,2005))
```