Homework Assignment 3

Brian Detweiler

January 31, 2017

1. For the cosine model on Page 34-35, it can be shown that the variance of the estimate for the trend in January can be given by Equation (3.4.6) on Page 38:

Answer:

$$Var(\hat{\mu}_1) = Var(\hat{\beta}_0) + Var(\hat{\beta}_1)\left[cos\left(\frac{2\pi}{12}\right)^2\right] + Var(\hat{\beta}_2)\left[sin\left(\frac{2\pi}{12}\right)\right]^2$$

Given the fact that

$$\hat{\beta}_0 = \frac{1}{n}\sum_{t=1}^{n} Y_t$$
$$\hat{\beta}_1 = \frac{2}{n}\sum_{t=1}^{n}\left[cos\left(\frac{2\pi}{12}t\right)Y_t\right]$$
$$\hat{\beta}_2 = \frac{2}{n}\sum_{t=1}^{n}\left[sin\left(\frac{2\pi}{12}t\right)Y_t\right]$$

and

$$Y_t = \mu_t + X_t$$
$$X_t \text{ is white noise with mean 0 and variance } \sigma^2$$

Show that $Var(\hat{\mu}_1) = \frac{3\sigma^2}{n}$. (Hint: $\sum_{t=1}^{n}\left[cos\left(\frac{2\pi}{12}t\right)\right]^2 = \left[sin\left(\frac{2\pi}{12}t\right)\right]^2 = \frac{n}{2}$.)

Answer:

Expanding the betas in the variance formula, and making note that $Var(\overline{Y}) = \frac{\gamma_0}{n}$ and $\gamma_0 = \sigma^2$, we get

$$
\begin{aligned}
Var(\hat{\mu}_1) &= Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) + Var\left(\frac{2}{n}\sum_{t=1}^{n}\left[cos(\frac{2\pi}{12}t)Y_t\right]\right)\left[cos(\frac{2\pi}{12})\right]^2 + Var\left(\frac{2}{n}\sum_{t=1}^{n}\left[sin(\frac{2\pi}{12}t)Y_t\right]\right)\left[sin(\frac{2\pi}{12})\right]^2 \\
&= Var(\overline{Y}) + 4Var\left(\frac{1}{n}\sum_{t=1}^{n} cos(\frac{2\pi}{12}t)\sum_{t=1}^{n} Y_t\right)\left[cos(\frac{2\pi}{12})\right]^2 + 4Var\left(\frac{1}{n}\sum_{t=1}^{n} sin(\frac{2\pi}{12}t)\sum_{t=1}^{n} Y_t\right)\left[sin(\frac{2\pi}{12})\right]^2 \\
&= Var(\overline{Y}) + 4\left[\sum_{t=1}^{n} cos(\frac{2\pi}{12}t)\right]^2 Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right)\left[cos(\frac{2\pi}{12})\right]^2 + 4\left[\sum_{t=1}^{n} sin(\frac{2\pi}{12}t)\right]^2 Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right)\left[sin(\frac{2\pi}{12})\right]^2 \\
&= Var(\overline{Y}) + 4\frac{n}{2}Var(\overline{Y})\left[cos(\frac{2\pi}{12})\right]^2 + 4\frac{n}{2}Var(\overline{Y})\left[sin(\frac{2\pi}{12})\right]^2 \\
&= Var(\overline{Y}) + 2nVar(\overline{Y})\left[\left(cos(\frac{2\pi}{12})\right)^2 + \left(sin(\frac{2\pi}{12})\right)^2\right] \text{ (note the trig identity } (sin(t))^2 + (cos(t))^2 = 1) \\
&= Var(\overline{Y}) + 2nVar(\overline{Y}) \\
&= \frac{\gamma_0}{n} + 2n\frac{\gamma_0}{n} \\
&= \frac{\gamma_0}{n}(1 + 2n) \\
&= \frac{\sigma^2}{n}(1 + 2n)
\end{aligned}
$$

$\blacksquare$

2. Let μ be a constant, and e_t be a white noise process with mean zero and variance σ_e^2. Consider the following three stochastic processes of Y_t. For each of the three processes find ρ_k for $k > 0$. Furthermore, for each Y_t, find $Var(\overline{Y})$ where $\overline{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$.

(a) $Y_t = \mu + e_t$

Answer:

For the autocorrelation function, $\rho_k = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}}$, we first find the Covariance.

$$\begin{aligned} Cov(Y_t, Y_{t-k}) &= Cov(\mu + e_t, \mu + e_{t-k}) \\ &= Cov(e_t, e_{t-k}) \\ &= \begin{cases} \sigma_e^2 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Now, we need the variance.

$$\begin{aligned} Var(Y_t) = Var(Y_{t-k}) = Var(\mu + e_t) &= Var(\mu + e_{t-k}) \text{ due to stationality} \\ &= Var(\mu) + Var(e_t) \\ &= 0 + \sigma_e^2 \\ &= \sigma_e^2 \end{aligned}$$

Now, putting these together, we can find the autocorrelation function,

$$\begin{aligned} \rho_k &= \frac{Cov(e_t, e_{t-k})}{\sqrt{Var(e_t)Var(e_{t-k})}} \\ &= \begin{cases} 1 & \text{if } t = k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Finding the variance of $\overline{Y}$,

$$
\begin{aligned}
Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) &= Var\left(\frac{1}{n}[\mu + e_1 + \mu + e_2 + \ldots + \mu + e_n]\right) \\
&= Var\left(\frac{1}{n}[n\mu + e_1 + e_2 + \ldots + e_n]\right) \\
&= \frac{1}{n^2}\left[Var(n\mu + e_1 + e_2 + \ldots + e_n)\right] \\
&= \frac{1}{n^2}\left[Var(n\mu) + Var(e_1) + Var(e_2) + \ldots + Var(e_n)\right] \\
&= \frac{1}{n^2}\left[0 + \sigma_e^2 + \sigma_e^2 + \ldots + \sigma_e^2\right] \\
&= \frac{1}{n^2}\left[\sigma_e^2 + \sigma_e^2 + \ldots + \sigma_e^2\right] \\
&= \frac{n\sigma_e^2}{n^2} \\
&= \frac{\sigma_e^2}{n}
\end{aligned}
$$

(b) $Y_t = \mu + e_t - e_{t-1}$

Answer:

For the autocorrelation function, $\rho_k = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}}$, we first find the Covariance.

$$
\begin{aligned}
Cov(Y_t, Y_{t-k}) &= Cov(\mu + e_t - e_{t-1}, \mu + e_{t-k} - e_{t-k-1}) \\
&= Cov(e_t - e_{t-1}, e_{t-k} - e_{t-k-1}) \\
&= Cov(e_t, e_{t-k}) - Cov(e_t, e_{t-k-1}) - Cov(e_{t-1}, e_{t-k}) + Cov(e_{t-1}, e_{t-k-1}) \\
&= \begin{cases} 2\sigma_e^2 & \text{if } k = 0 \\ -2\sigma_e^2 & \text{if } k = 1 \\ -\sigma_e^2 & \text{if } k = -1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Now, we need the variance.

$$
\begin{aligned}
Var(Y_t) &= Var(Y_{t-k}) = Var(\mu + e_t - e_{t-k}) = Var(\mu + e_{t-k} - e_{t-k-1}) \text{ due to stationality} \\
&= Var(\mu) + Var(e_t) + Var(e_{t-k-1}) \\
&= 0 + \sigma_e^2 + \sigma_e^2 \\
&= 2\sigma_e^2
\end{aligned}
$$

Now, putting these together, we can find the autocorrelation function,

$$
\begin{aligned}
\rho_k &= \frac{Cov(e_t, e_{t-k})}{\sqrt{Var(e_t)Var(e_{t-k})}} \\
&= \begin{cases} 1 & \text{if } k = 0 \\ -1 & \text{if } k = 1 \\ -\frac{1}{2} & \text{if } k = -1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Finding the variance of $\overline{Y}$,

$$
\begin{aligned}
Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) &= Var\left(\frac{1}{n}[\mu + e_1 - e_0 + \mu + e_2 - e_1 + \ldots + \mu + e_{n-1} + e_n]\right) \\
&= Var\left(\frac{1}{n}[n\mu - e_0 + e_1 - e_1 + e_2 - e_2 + \ldots + e_{n-1} - e_{n-1} + e_n]\right) \\
&= \frac{1}{n^2}\left[Var(n\mu - e_0 + e_n)\right] \\
&= \frac{1}{n^2}\left[Var(n\mu) + Var(e_0) + Var(e_n)\right] \\
&= \frac{1}{n^2}\left[0 + \sigma_e^2 + \sigma_e^2\right] \\
&= \frac{1}{n^2}\left[2\sigma_e^2\right] \\
&= \frac{2\sigma_e^2}{n^2}
\end{aligned}
$$

(c) $Y_t = \mu + e_t + e_{t-1}$

Answer:

For the autocorrelation function, $\rho_k = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}}$, we first find the Covariance.

$$
\begin{aligned}
Cov(Y_t, Y_{t-k}) &= Cov(\mu + e_t + e_{t-1}, \mu + e_{t-k} + e_{t-k-1}) \\
&= Cov(e_t + e_{t-1}, e_{t-k} + e_{t-k-1}) \\
&= Cov(e_t, e_{t-k}) + Cov(e_t, e_{t-k-1}) + Cov(e_{t-1}, e_{t-k}) + Cov(e_{t-1}, e_{t-k-1}) \\
&= \begin{cases} 2\sigma_e^2 & \text{if } k = 0 \\ \sigma_e^2 & \text{if } k = \pm 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Now, we need the variance.

$$
\begin{aligned}
Var(Y_t) = Var(Y_{t-k}) = Var(\mu + e_t + e_{t-k}) &= Var(\mu + e_{t-k} + e_{t-k-1}) \text{ due to stationality} \\
&= Var(\mu) + Var(e_t) + Var(e_{t-k-1}) \\
&= 0 + \sigma_e^2 + \sigma_e^2 \\
&= 2\sigma_e^2
\end{aligned}
$$

Now, putting these together, we can find the autocorrelation function,

$$
\begin{aligned}
\rho_k &= \frac{Cov(e_t, e_{t-k})}{\sqrt{Var(e_t)Var(e_{t-k})}} \\
&= \begin{cases} 1 & \text{if } k = 0 \\ \frac{1}{2} & \text{if } k = -1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Finding the variance of $\overline{Y}$,

$$\begin{aligned}
Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) &= Var\left(\frac{1}{n}[n\mu + e_1 + e_0 + e_2 + e_1 + e_3 + e_2 + \ldots + e_{n-1} + e_n]\right) \\
&= Var\left(\frac{1}{n}[n\mu + e_0 + 2e_1 + 2e_2 + \ldots + 2e_{n-1} + e_n]\right) \\
&= Var\left(\frac{1}{n}\left[n\mu + e_0 + e_n + 2\sum_{t=1}^{n-1} e_t\right]\right) \\
&= \frac{1}{n^2} Var\left(\left[n\mu + e_0 + e_n + 2\sum_{t=1}^{n-1} e_t\right]\right) \\
&= \frac{1}{n^2}\left[Var(n\mu) + Var(e_0) + Var(e_n) + 4Var\left(\sum_{t=1}^{n-1} e_t\right)\right] \\
&= \frac{1}{n^2}[0 + \sigma_e^2 + \sigma_e^2 + 4(n-1)\sigma_e^2] \\
&= \frac{2\sigma_e^2 + 4\sigma_e^2 n - 4\sigma_e^2}{n^2} \\
&= \frac{\sigma_e^2[2 + 4n - 4]}{n^2} \\
&= \frac{[4n-2]}{n^2}\sigma_e^2
\end{aligned}$$

■

3. The data file `wages` contains monthly values of the average hourly wages (in dollars) for workers in the U.S. apparel and textile products industry for July 1981 through 1987.

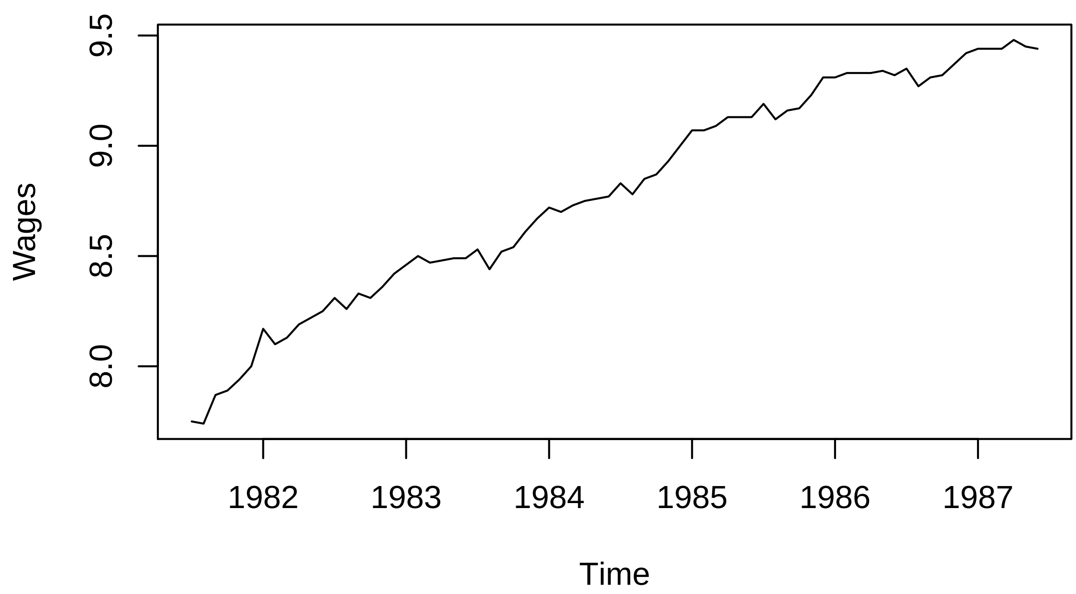
```
data(wages)
wages
```

```
##       Jan  Feb  Mar  Apr  May  Jun  Jul  Aug  Sep  Oct  Nov  Dec
## 1981                               7.75 7.74 7.87 7.89 7.94 8.00
## 1982 8.17 8.10 8.13 8.19 8.22 8.25 8.31 8.26 8.33 8.31 8.36 8.42
## 1983 8.46 8.50 8.47 8.48 8.49 8.49 8.53 8.44 8.52 8.54 8.61 8.67
## 1984 8.72 8.70 8.73 8.75 8.76 8.77 8.83 8.78 8.85 8.87 8.93 9.00
## 1985 9.07 9.07 9.09 9.13 9.13 9.13 9.19 9.12 9.16 9.17 9.23 9.31
## 1986 9.31 9.33 9.33 9.33 9.34 9.32 9.35 9.27 9.31 9.32 9.37 9.42
## 1987 9.44 9.44 9.44 9.48 9.45 9.44
```

(a) Display and interpret the time series plot for these data.

Answer:

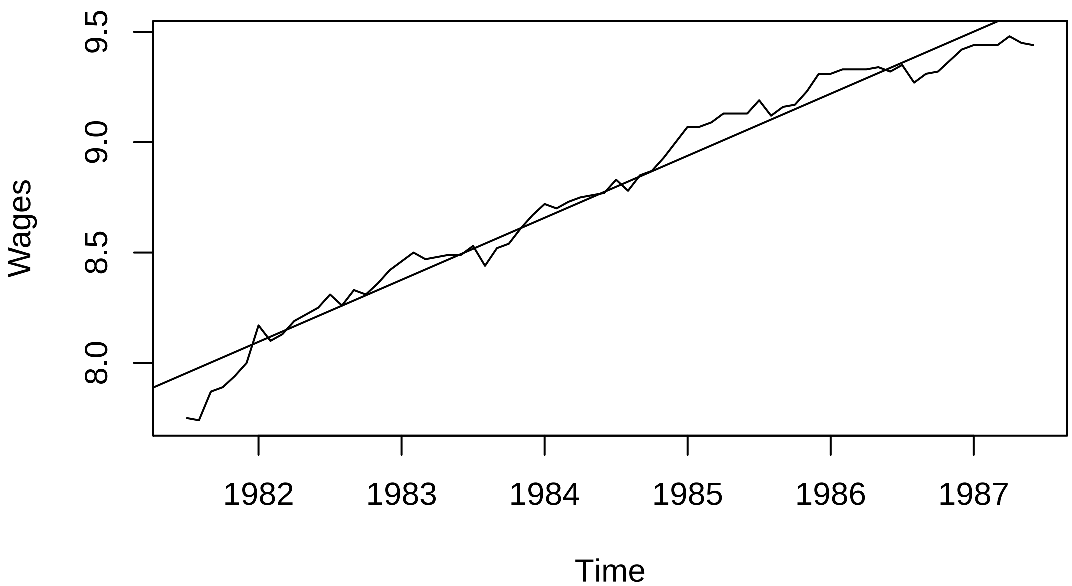
```
plot.ts(wages, frequency=12, start=c(1981, 7))
```



(b) Use least squares to fit a linear time trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for futher analysis.

Answer:

```
fit <- lm(wages ~ time(wages))
summ <- summary(fit)
plot.ts(wages, frequency=12, start=c(1981, 7))
abline(fit)
```




```
summm
```

```
##
## Call:
## lm(formula = wages ~ time(wages))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.23828 -0.04981  0.01942  0.05845  0.13136
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.490e+02  1.115e+01  -49.24   <2e-16 ***
## time(wages)  2.811e-01  5.618e-03   50.03   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08257 on 70 degrees of freedom
## Multiple R-squared:  0.9728, Adjusted R-squared:  0.9724
```

```
## F-statistic:  2503 on 1 and 70 DF,  p-value: < 2.2e-16
```

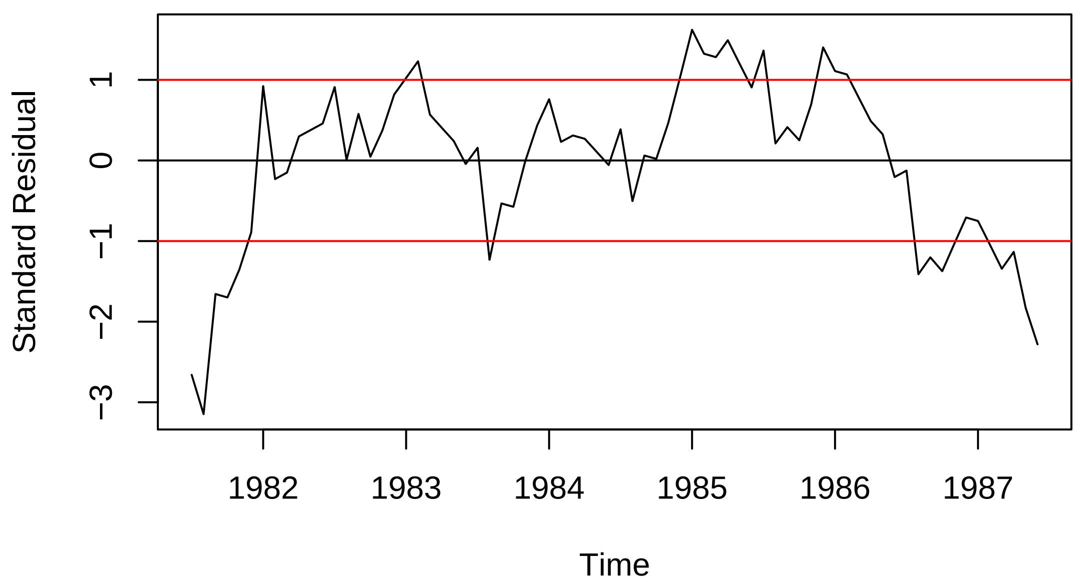
```r
std.resid <- rstudent(fit)
```

(c) Construct and interpret the time series plot of the standardized residuals from (b).

Answer:

It appears that the residuals are mostly zero-mean and fall within one standard deviation, except for the beginning and end of the plot.

```r
std.resid.ts <- ts(std.resid, start=c(1981, 7), frequency = 12)
plot.ts(std.resid.ts, ylab="Standard Residual")
abline(h=0)
abline(h=1, col="red")
abline(h=-1, col="red")
```



(d) Use least squares to fit a quadratic time trend to the wages time series. Interpret the regression output. Save the standardized residuals from the fit.

Answer:

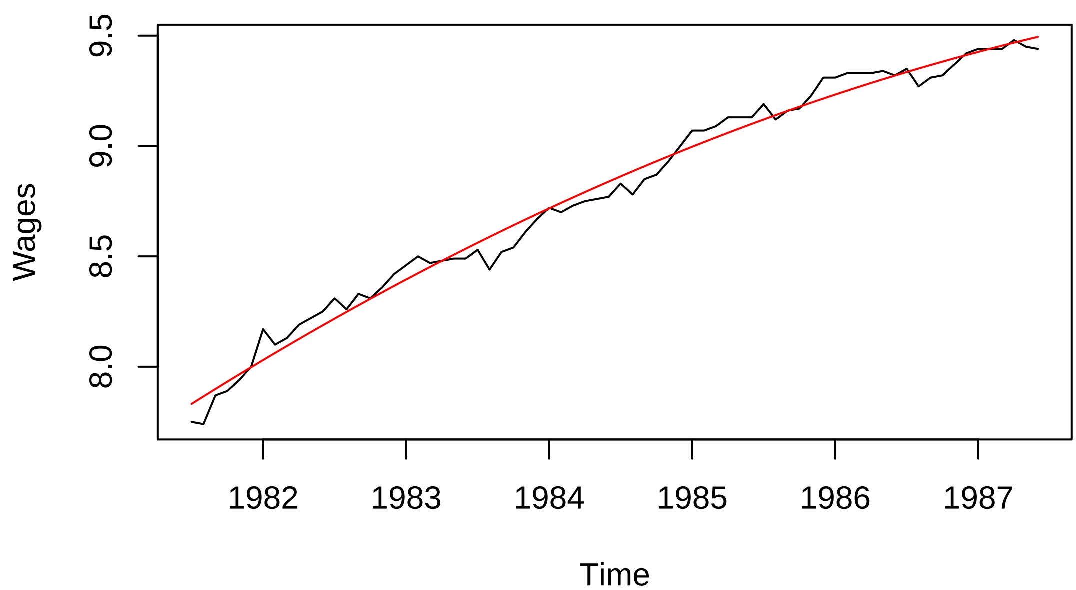
```r
fit.quad <- lm(wages ~ time(wages) + I(time(wages)^2))
summ.quad <- summary(fit.quad)
summ.quad
```

```
##
## Call:
```

```
## lm(formula = wages ~ time(wages) + I(time(wages)^2))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.148318 -0.041440  0.001563  0.050089  0.139839
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -8.495e+04  1.019e+04  -8.336 4.87e-12 ***
## time(wages)        8.534e+01  1.027e+01   8.309 5.44e-12 ***
## I(time(wages)^2) -2.143e-02  2.588e-03  -8.282 6.10e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05889 on 69 degrees of freedom
## Multiple R-squared:  0.9864, Adjusted R-squared:  0.986
## F-statistic:  2494 on 2 and 69 DF,  p-value: < 2.2e-16
```

```r
predicted.counts <- predict(fit.quad, list(Time=time(wages), Time2=time(wages)^2))
predicted.ts <- ts(predicted.counts, start=c(1981, 7), frequency=12)

plot.ts(wages)
lines(predicted.ts, col="red")
```

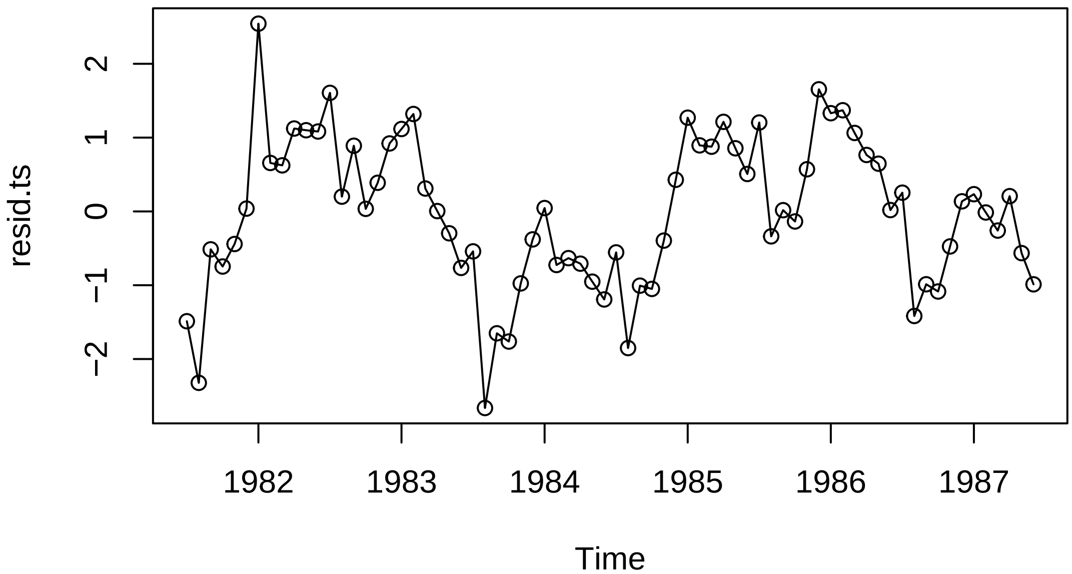


Comparing R^2 values from the linear model and the quadratic model, we can see that Model 1 has $R^2 = 0.972792$ and Model 2 has an adjusted $R^2 = 0.9859596$, which is slightly better.

(e) Construct and interpret the time series plot of the standardized residuals from part(d).

Answer:

```
resid <- rstudent(fit.quad)
resid.ts <- ts(resid, start=c(1981, 7), frequency=12)
plot(resid.ts, type="o")
```



The data show mean -0.0049486 and standard deviation 1.0208957, which would suggest that X_t is indeed white noise. However, we notice some patterns in the residuals, and a few outliers. We may need some more tests.

(f) Perform a runs test on the standardized residuals from part (d) and interpret the results.

Answer:

With the following null and alternative hypotheses, we perform a Runs Test:

$$H_0: \text{The data are from a standard normal}$$
$$H_a: \text{The data are not from a standard normal}$$

```
runs <- runs.test(factor(resid > median(resid)))
runs
```

```
##
##  Runs Test
```

```
##
## data:  factor(resid > median(resid))
## Standard Normal = -5.2224, p-value = 1.767e-07
## alternative hypothesis: two.sided
```

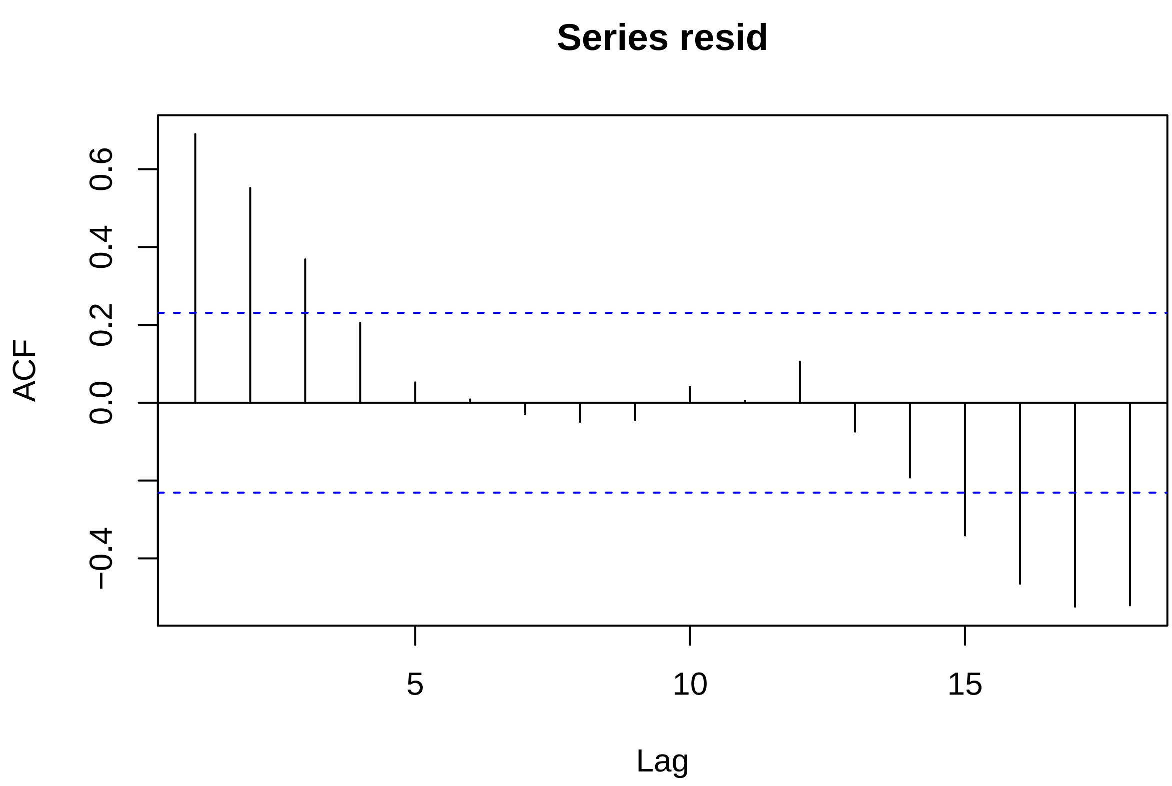
With a p-value of $1.7665967 \times 10^{-7} << 0.05$, we can say there are trends in the data.

(g) Calculate and interpret the sample autocorrelations for the standardized residuals from part (d).

Answer:

Here, we see that the autocorrelation at lag $k = 1, 2, 3$ and $k = 15, 16, 17, 18$ fall outside of the standard error (blue lines). The autocorrelation is also heteroskedastic, which is uncharacteristic of a white noise process.

```r
acf(resid)
```

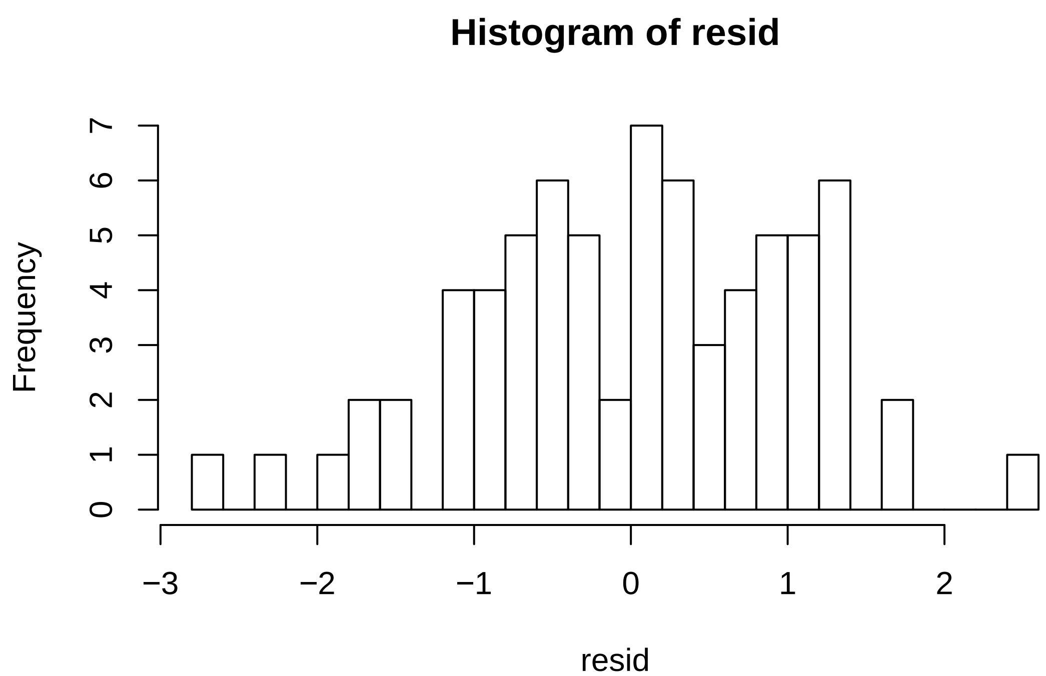


(h) Investigate the normality of the standardized residuals from part (d). Consider histograms and normal probability plots. Interpret the plots. Perform the Shapiro-Wilk test for Normality.

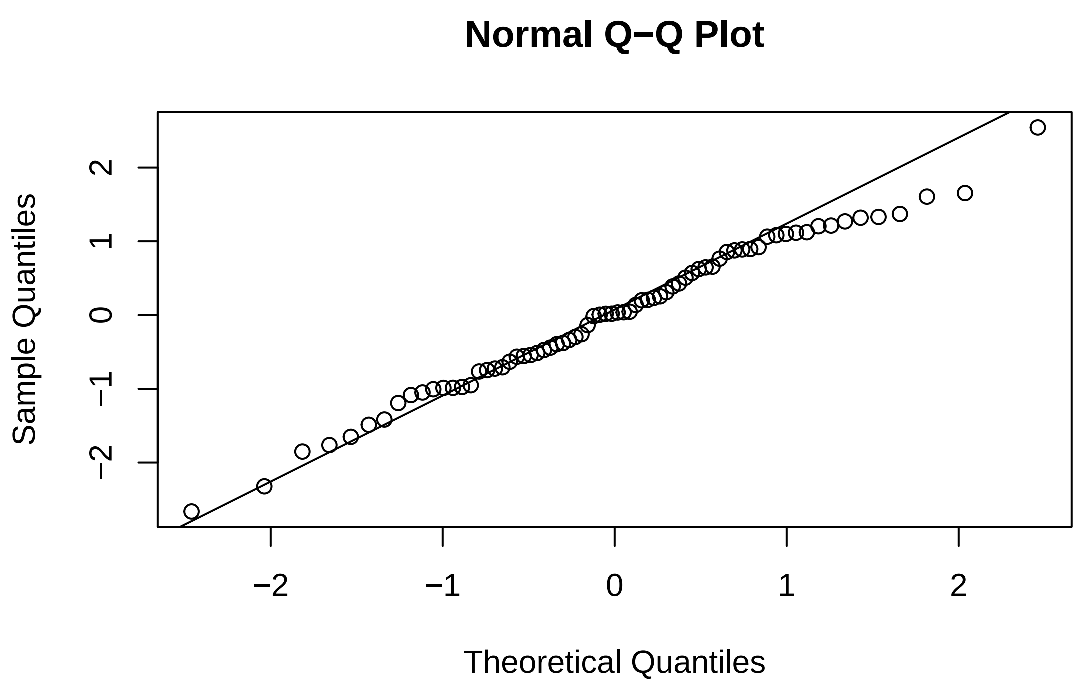
Answer:

Looking at the histogram, we can see how this would be problematic for describing a white noise process. We would expect to see a peak at 0, but instead we see a trough, and there are outliers on the left and right.

```r
hist(resid, breaks=30)
```




```r
qqn <- qqnorm(resid)
qqline(resid)
```



```r
r.sq <- cor(qqn$x, qqn$y)
```

We can further use a Q-Q plot to see how the data conform to a normal distribution. Here, we have $R^2 = 0.9940933$, but we can see some deviation in the upper quantiles. From here, we perform a Shapiro-Wilk test for Normality.

Again, using the following hypotheses:

$$
\begin{aligned}
H_0 &: \text{The data are from a standard normal} \\
H_a &: \text{The data are not from a standard normal}
\end{aligned}
$$

```r
shapiro.test(resid)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid
## W = 0.98868, p-value = 0.7693
```

Thus we do not have enough information to reject the null hypothesis, so we must conclude that the data are from a normal distribution.

■

4. The data file **retail** lists total U.K. retail sales (in billions of pounds) from January 1986 through March 2007. Note that year 2000 = 100 is the base year.

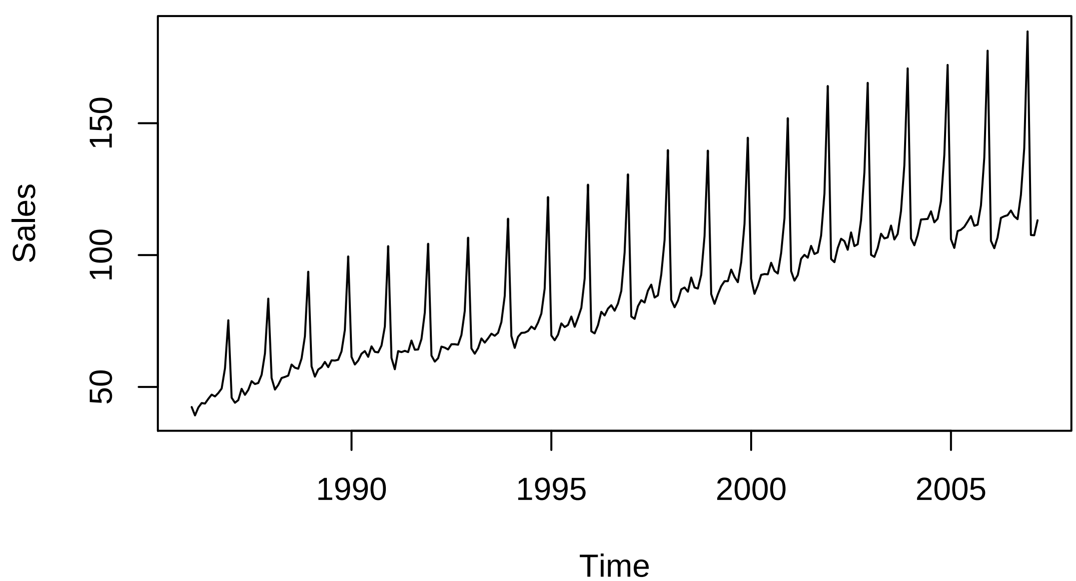
```
data(retail)
```

(a) Display and interpret the time series plot for these data. Do you see any seasonally trend from the data?

Answer:

Here we can see that sales are generally increasing over time, but there is also a serious seasonal effect.

```
plot(retail)
```



(b) Use least squares to fit a seasonal-means plus linear time trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

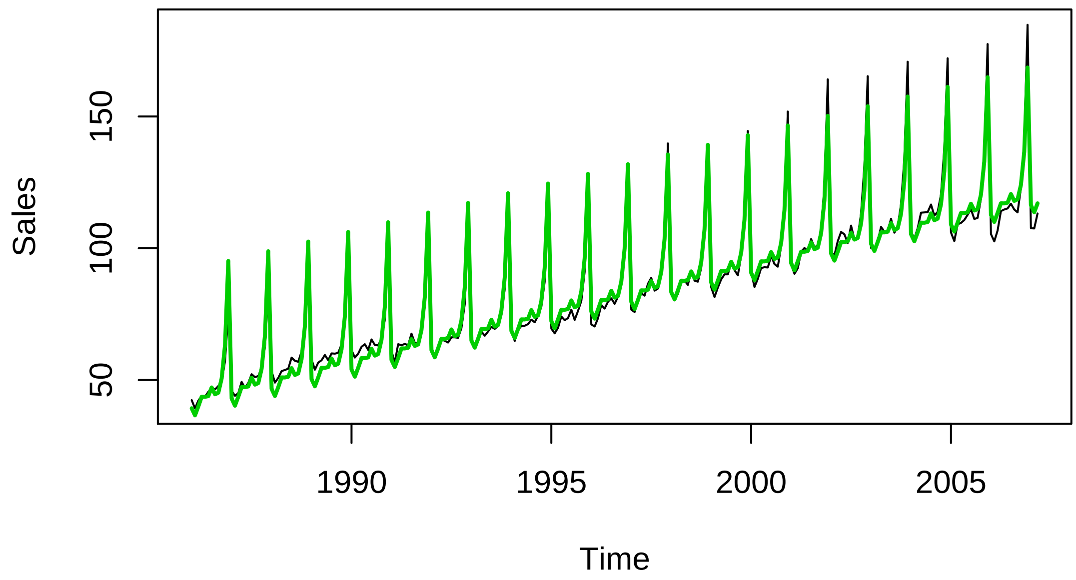
Answer:

```
fit1 = lm(retail ~ time(retail) + season(retail) - 1)
summ1 <- summary(fit1)
summ1
```

```
##
## Call:
```

```
## lm(formula = retail ~ time(retail) + season(retail) - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.8950  -2.4440  -0.3518   2.1971  16.2045
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## time(retail)             3.670e+00  4.369e-02   84.00   <2e-16 ***
## season(retail)January   -7.249e+03  8.724e+01  -83.10   <2e-16 ***
## season(retail)February  -7.252e+03  8.724e+01  -83.13   <2e-16 ***
## season(retail)March     -7.249e+03  8.725e+01  -83.09   <2e-16 ***
## season(retail)April     -7.246e+03  8.723e+01  -83.07   <2e-16 ***
## season(retail)May       -7.246e+03  8.723e+01  -83.07   <2e-16 ***
## season(retail)June      -7.246e+03  8.723e+01  -83.07   <2e-16 ***
## season(retail)July      -7.243e+03  8.724e+01  -83.03   <2e-16 ***
## season(retail)August    -7.246e+03  8.724e+01  -83.06   <2e-16 ***
## season(retail)September -7.246e+03  8.725e+01  -83.05   <2e-16 ***
## season(retail)October   -7.241e+03  8.725e+01  -82.99   <2e-16 ***
## season(retail)November  -7.229e+03  8.725e+01  -82.85   <2e-16 ***
## season(retail)December  -7.197e+03  8.726e+01  -82.48   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.278 on 242 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9978
## F-statistic:  8791 on 13 and 242 DF,  p-value: < 2.2e-16
```

```r
resid1 <- rstudent(fit1)
resid1.ts <- ts(resid1, start=c(1986, 1), freq=12)
plot(retail)
lines(as.vector(time(retail)), fitted.values(fit1), col = 3, lwd = 2)
```

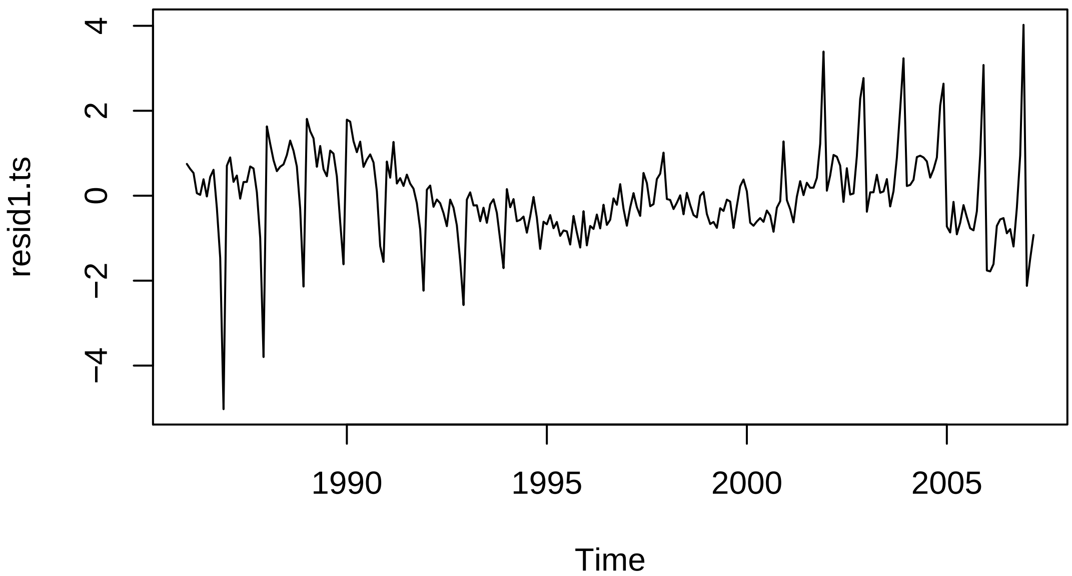


(c) Construct and interpret the time series plot of the standardized residuals from part (b). Are there still any seasonal trends in the residuals?

Answer:

Here, the residuals very much display a seasonal effect, and is very heteroskedastic.

```
plot(resid1.ts)
```



(d) Perform a runs test on the standardized residuals from part (b) and interpret the results.

Answer:

```
runs <- runs.test(factor(resid1 > median(resid1)))
runs
```

```
##
##  Runs Test
##
## data:  factor(resid1 > median(resid1))
## Standard Normal = -9.3491, p-value < 2.2e-16
## alternative hypothesis: two.sided
```

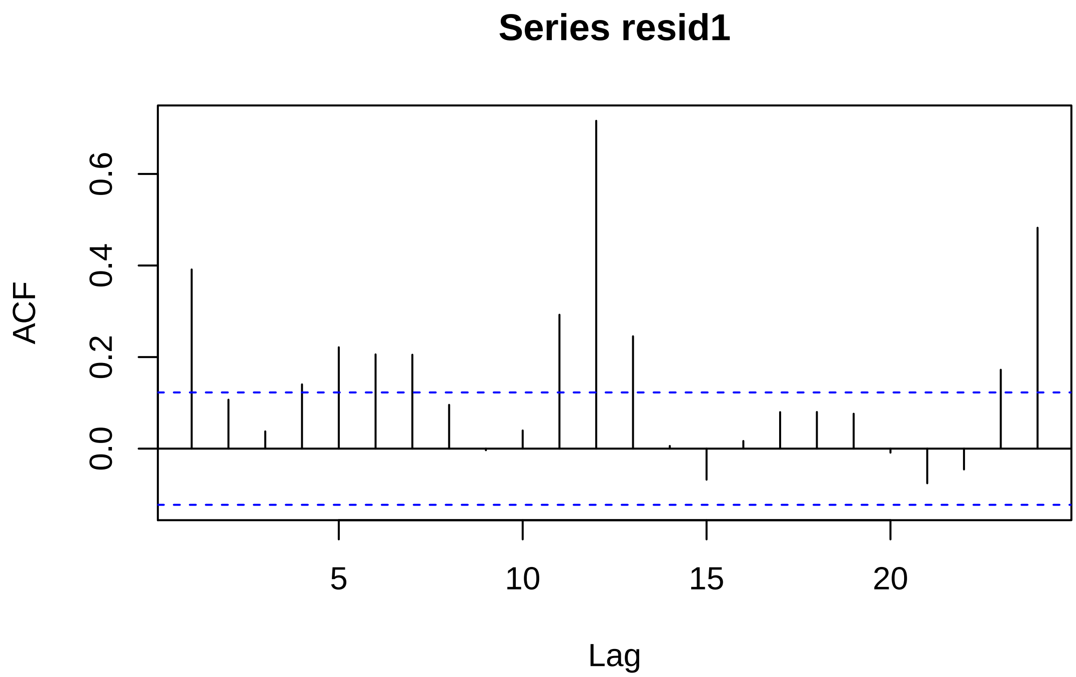
The runs test has a p-value of $8.8428247 \times 10^{-21} << 0.05$, so we can say there are definitely trends in the residuals.

(e) Calculate and interpret the sample autocorrelations for the standardized residuals from part (b).

Answer:

There are several values for lag where the autocorrelation falls outside of the standard error lines.

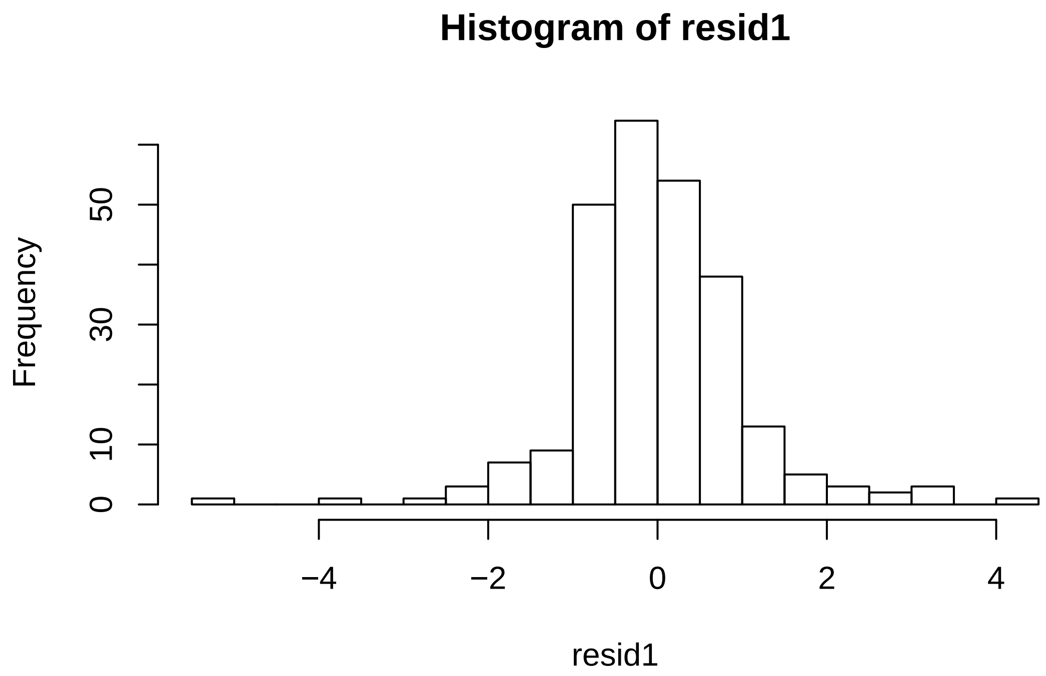
```
acf(resid1)
```



(f) Investigate the normality of the standardized residuals from part (b). Consider histograms and normal probability plots. Interpret the plots. Perform the Shapiro-Wilk test for Normality.

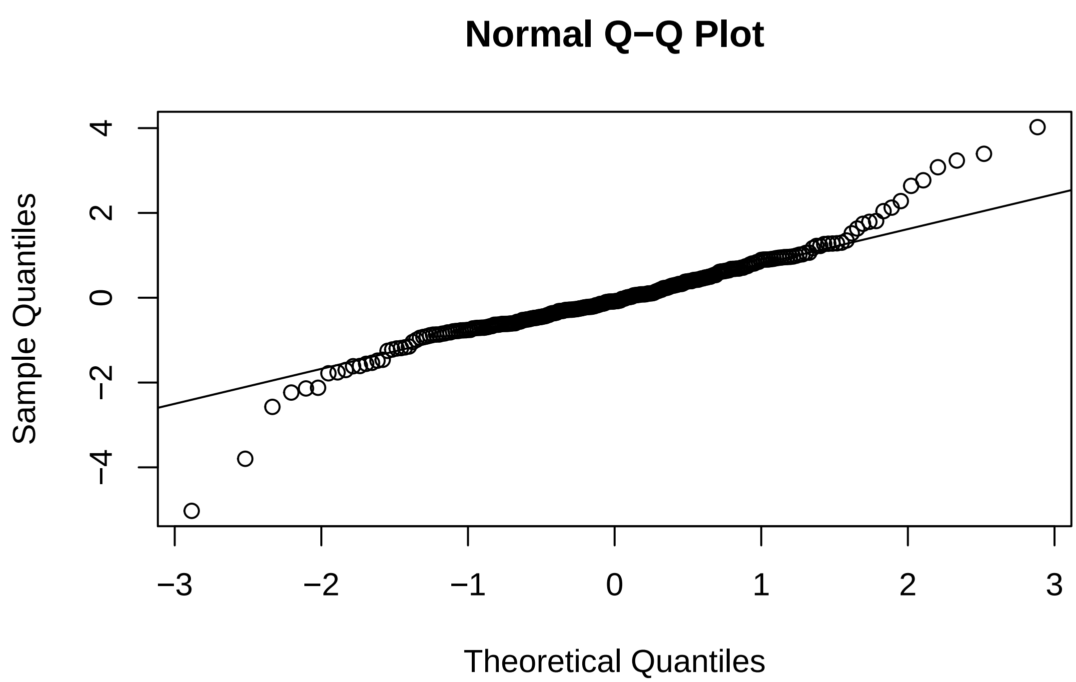
Answer: The histogram looks very normal, but there are outliers that may be throwing off the normality.

```
hist(resid1, breaks=30)
```



And here in the Q-Q plot, we can see this effect, throwing the data off from normal at the outliers.

```
qqn <- qqnorm(resid1)
qqline(resid1)
```



```
r.sq <- cor(qqn$x, qqn$y)
```

While the data fit a normal distribution with $R^2 = 0.9671387$, we may be hesitant to say for sure it is normal due to the outliers in the tails.

Finally, we perform a Shapiro-Wilk test for normality to confirm our suspicion:

$$
\begin{aligned}
H_0 &: \text{The data are from a standard normal} \\
H_a &: \text{The data are not from a standard normal}
\end{aligned}
$$

```
shap <- shapiro.test(resid1)
```

With a p-value of $8.5335864 \times 10^{-9} << 0.05$, so we must conclude that the residuals are not from a standard normal, and thus not from a white noise process.

■