Time Series Analysis - 478 - Exam 2

Alex Towell (atowell@siue.edu)

Contents

Problem 1.1

Suppose that simple exponential smoothing is being used to forecast the process $y_t = \mu + e_t$, where where $\{e_t\}$ are white noise with mean 0 and variance σ^2. At the start of period t^*, the mean of the process experiences a transient; that is, it shifts to a new level $\mu + \delta$, but reverts to its original level μ at the start of the next period $t^* + 1$. The mean remains at this level for subsequent time periods.

Part (a)

Find the expected value of the simple exponential smoother

$$\tilde{y}_T = (1-\theta)\sum_{t=0}^{\infty}\theta^t y_T.$$

We have a time series

$$y_t = \mu + e_t$$

except at y_{t^*} which is distributed

$$y_t^* = \mu + \delta + e_{t^*}$$

where the error terms are zero mean white noise with variance σ^2.

The expectation of the smoothed time series $\tilde{y}_T$ is given by

$$\begin{aligned}
\mathrm{E}(\tilde{y}_T) &= (1-\theta)\sum_{t=0}^{\infty}\theta^t\,\mathrm{E}(y_{T-t}) \\
&= (1-\theta)\left(\sum_{t=0}^{T-t^*+1}\theta^t\mu + \theta^{T-t^*}(\mu+\delta) + \sum_{t=T-t^*-1}^{\infty}\theta^t\mu\right) \\
&= (1-\theta)\left(\sum_{t=0}^{\infty}\theta^t\mu + \theta^{T-t^*}\delta\right) \\
&= \mu + (1-\theta)\theta^{T-t^*}\delta.
\end{aligned}$$

Part (b)

For $\theta = 0.5$, determine the number of periods that it will take following the impulse for the expected value of $\tilde{y}_T$ to return to within 0.1δ of the original level μ.

We wish to find $\tilde{y}_k$ such that it is expected to be within $\frac{1}{10}\delta$ of μ,

$$|\mathrm{E}(\tilde{y}_k) - \mu| \le \left|\frac{1}{10}\delta\right|.$$

Plugging in the definition of the expectation and simplifying,

$$\left|(1-\theta)\theta^{k-t^*}\delta\right| \le \left|\frac{1}{10}\delta\right|.$$

Since pulling all positive numbers (or symbols that stand for positive numbers) out of the absolute value function does not change the expression, we may rewrite the above as

$$(1-\theta)\theta^{k-t^*}|\delta| \le \frac{1}{10}|\delta|.$$

Dividing by $|\delta|$ on both sides,

$$(1-\theta)\theta^{k-t^*} \le \frac{1}{10},$$

which may be rewritten as

$$\theta^k \le \frac{\theta^{t^*}}{10(1-\theta)}.$$

Taking the logarithm of both sides

$$\begin{aligned} k \log \theta &\leq \log\left(\frac{\theta^{t^*}}{10(1-\theta)}\right) \\ &\leq t^* \log \theta - \log 10 - \log(1-\theta). \end{aligned}$$

Finally, we isolate k by dividing by $\log \theta$ on both sides. However, note that $\log \theta$ is negative, and so we must flip the inequality,

$$k \geq t^* - \frac{\log 10}{\log \theta} - \frac{\log(1-\theta)}{\log \theta}.$$

Letting $\theta = 0.5$,

$$k \geq t^* + \frac{\log 10}{\log 2} - \frac{\log 0.5}{\log 0.5}$$

which simplifies to

$$k \geq t^* + 2.32.$$

We wish to take the *smallest* k that is an integer that satisfies the equation. Thus, $k = t^* + 3$. Or, in other words, 3 periods after t^*, $\tilde{y}_T$ has an expectation that is within the specified distance of μ.

Problem 1.2

Let $\{Y_t\}$ be an AR(1) process with $|\varphi| < 1$. That is $Y_t = \varphi Y_{t-1} + e_t$, where $\{e_t\}$ are white noise with mean 0 and variance σ^2. Also note e_t's are independent of $Y_{t-1}, Y_{t-2}, ...$

Part (a)

Find the autocorrelation function for $W_t = Y_t - Y_{t-1}$ in terms of φ and σ^2.

Observe that

$$W_t = Y_t - Y_{t-1} = \varphi Y_{t-1} + e_t - Y_{t-1}$$

and thus

$$W_t = (\varphi - 1)Y_{t-1} + e_t.$$

The autocovariance function for W_t, denoted by $\gamma_{\{W_t\}}$, is defined as

$$\gamma_{\{W_t\}}(k) = \operatorname{Cov}(W_t, W_{t-k}).$$

Assuming $k \neq 0$ (we solve directly for variance in that case) and replacing W_t and W_{t-k} with their respective definitions yields

$$\begin{aligned} \gamma_{\{W_t\}}(k) &= \operatorname{Cov}((\varphi-1)Y_{t-1} + e_t, (\varphi-1)Y_{t-k-1} + e_{t-k}) \\ &= \operatorname{Cov}((\varphi-1)Y_{t-1}, (\varphi-1)Y_{t-k-1}) \\ &= (\varphi-1)^2 \operatorname{Cov}(Y_{t-1}, Y_{t-k-1}). \end{aligned}$$

Observe that $\operatorname{Cov}(Y_{t-1}, Y_{t-k-1}) = \gamma_{\{Y_t\}}(k)$. Since $\{Y_t\}$ is AR(1),

$$\gamma_{\{Y_t\}}(k) = \sigma^2 \frac{\varphi^k}{1-\varphi^2}.$$

Thus,

$$\gamma_{\{W_t\}}(k) = \gamma_{\{Y_t\}}(k) = \sigma^2 \frac{\varphi^k}{1-\varphi^2}.$$

The variance of $\{W_t\}$ is given by

$$\begin{aligned}
\operatorname{Cov}(W_t, W_t) &= \operatorname{Cov}((\varphi - 1)Y_{t-1} + e_t, (\varphi - 1)Y_{t-1} + e_t) \\
&= (\varphi - 1)^2 \operatorname{Cov}(Y_{t-1}, Y_{t-1}) + \operatorname{Cov}(e_t, e_t) \\
&= (\varphi - 1)^2 \frac{\sigma^2}{1 - \varphi^2} + \sigma^2 \\
&= \sigma^2 \left(1 + \frac{(\varphi - 1)^2}{1 - \varphi^2}\right).
\end{aligned}$$

Thus, the autocorrelation function is given by

$$\rho_k = \frac{\gamma_{\{W_t\}}(k)}{\gamma_{\{W_t\}}(0)} = \frac{\sigma^2 \frac{\varphi^k}{1-\varphi^2}}{\sigma^2 \left(1 + \frac{(\varphi-1)^2}{1-\varphi^2}\right)},$$

which simplifies to

$$\rho_k = \frac{\frac{\varphi^k}{1-\varphi^2}}{1 + \frac{(\varphi-1)^2}{1-\varphi^2}} = \frac{\varphi^k}{2(1 - \varphi)}.$$

Part (b)

In part (a), we found that

$$\operatorname{Var}(W_t) = \sigma^2 \left(1 + \frac{(\varphi - 1)^2}{1 - \varphi^2}\right).$$

Problem 1.3

Suppose $Y_t = X_t + e_t$, where $\{e_t\}$ are normal white noise with mean 0 and variance σ_e^2. The $\{X_t\}$ process is a stationary AR(1) defined by $X_t = \varphi X_{t-1} + Z_t$, where $\{Z_t\}$ is a zero mean normal white noise process with variance σ_Z^2. As usual, in the AR(1) process, assume that Z_t is independent of $X_{t-1}, X_{t-2}, \ldots$. Assume additionally that $\operatorname{E}(e_t Z_s) = 0$ for all t and s.

Part (a)

Show that $\{Y_t\}$ is stationary and find its autocovariance function, γ_k.

To be stationary, $\{Y_t\}$ must have a constant mean and a a autocovariance that is strictly a function of the lag.

The mean is given by

$$\operatorname{E}(Y_t) = \operatorname{E}(X_t) + \operatorname{E}(e_t).$$

Since X_t is AR(1) with mean $\delta/(1 - \varphi) = 0$, we see that $\operatorname{E}(Y_t) = 0$, i.e., is a constant zero.

The variance is given by

$$\operatorname{Var}(Y_t) = \operatorname{Var}(X_t) + \sigma^2.$$

Since X_t is AR(1), its variance is $\sigma_Z^2/(1 - \varphi^2)$, thus

$$\operatorname{Var}(Y_t) = \sigma_Z^2/(1 - \varphi^2) + \sigma^2.$$

The autocovariance of $\{Y_t\}$ is given by

$$\gamma_k = \operatorname{Cov}(Y_t, Y_{t-k}) = \operatorname{E}(Y_t Y_{t-k}) - \operatorname{E}(Y_t)\operatorname{E}(Y_{t-k}).$$

Since $\{Y_t\}$ has a constant expectation of zero, this simplies to

$$\gamma_k = \mathrm{Cov}(Y_t, Y_{t-k}) = \mathrm{E}(Y_t Y_{t-k}).$$

Observe that $Y_t = \varphi X_{t-1} + Z_t + e_t$ and

$$Y_t Y_{t-k} = (\varphi X_{t-1} + Z_t + e_t) Y_{t-k} = \varphi Y_{t-k} X_{t-1} + Y_{t-k} Z_t + Y_{t-k} e_t.$$

The expectation of $Y_t Y_{t-k}$ is given by

$$\begin{aligned}
\mathrm{E}(Y_t Y_{t-k}) &= \varphi\,\mathrm{E}(Y_{t-k} X_{t-1}) + \mathrm{E}(Y_{t-k} Z_t) + \mathrm{E}(Y_{t-k} e_t) \\
&= \varphi\,\mathrm{E}(Y_{t-k} X_{t-1}) + \mathrm{E}(Y_{t-k})\,\mathrm{E}(Z_t) + \mathrm{E}(Y_{t-k})\,\mathrm{E}(e_t) \\
&= \varphi\,\mathrm{E}(Y_{t-k} X_{t-1}) \\
&= \varphi\,\mathrm{E}((X_{t-k} + e_t) X_{t-1}) \\
&= \varphi\,\mathrm{E}(X_{t-1} X_{t-k} + e_t X_{t-1}) \\
&= \varphi\,(\mathrm{E}(X_{t-1} X_{t-k}) + \mathrm{E}(e_t X_{t-1})) \\
&= \varphi\,\mathrm{E}(X_{t-1} X_{t-k}).
\end{aligned}$$

Since $\{X_t\}$ is AR(1), observe that the autocovariance function for $\{X_t\}$ is $\gamma_{\{X_t\}}(k) = \varphi\,\mathrm{E}(X_{t-1} X_{t-k})$, which has a closed-form solution

$$\gamma_{\{X_t\}}(k) = \begin{cases} \frac{\sigma_Z^2}{1-\varphi^2} & k = 0 \\ \varphi \gamma_{\{X_t\}}(k-1) & k > 0. \end{cases} \tag{1}$$

Thus, the autocovariance function for $\{Y_t\}$ is given by

$$\gamma_k = \begin{cases} \frac{\sigma_Z^2}{1-\varphi^2} + \sigma_e^2 & k = 0 \\ \gamma_{\{X_t\}}(k) & k > 0. \end{cases} \tag{2}$$

Since its autocovariance function is strictly a function of lag and its mean is a constant zero, $\{Y_t\}$ is stationary. Note that it is not just weakly stationary, but strongly stationary given the normally distributed random errors.

Part (b)

Show that the process $\{U_t\}$, where $U_t = Y_t - \varphi Y_{t-1} = (1 - \varphi B) Y_t$, has nonzero correlation only at lag 1 (excluding lag 0, of course!).

The autocovariance is given by

$$\begin{aligned}
\gamma_{\{U_t\}}(k) &= \mathrm{Cov}(U_t, U_{t-k}) \\
&= \mathrm{Cov}(Y_t - \varphi Y_{t-1}, Y_{t-k} - \varphi Y_{t-k-1}).
\end{aligned}$$

Observe that $Y_t - \varphi Y_{t-1} = X_t + e_t - \varphi(X_{t-1} + e_{t-1})$. Since $Z_t = X_t - \varphi X_{t-1}$, we see that

$$Y_t - \varphi Y_{t-1} = e_t + Z_t - \varphi e_{t-1}$$

and

$$Y_{t-k} - \varphi Y_{t-k-1} = e_{t-k} + Z_{t-k} - \varphi e_{t-k-1}.$$

Thus,

$$\gamma_{\{X_t\}}(k) = \mathrm{Cov}(e_t + Z_t - \varphi e_{t-1}, e_{t-k} + Z_{t-k} - \varphi e_{t-k-1}).$$

If $k > 1$, then $\gamma_{\{X_t\}}(k) = \operatorname{Cov}(e_t + Z_t - \varphi e_{t-1}, e_{t-k} + Z_{t-k} - \varphi e_{t-k-1}) = 0$ since they have no terms in common. If $k = 1$, then

$$\begin{aligned}
\gamma_{\{X_t\}}(1) &= \operatorname{Cov}(e_t + Z_t - \varphi e_{t-1}, e_{t-1} + Z_{t-1} - \varphi e_{t-2}) \\
&= \operatorname{Cov}(-\varphi e_{t-1}, e_{t-1}) \\
&= -\varphi \operatorname{Var}(e_{t-1}) \\
&= -\varphi \sigma_e^2,
\end{aligned}$$

which is the only lag that is non-zero.

Problem 1.4

Suppose that $\{e_t\}$ is a zero mean white noise process with variance σ^2. Consider:

(i) $y_t = 0.80y_{t-1} - 0.15y_{t-2} + e_t - 0.30e_{t-1}$

(ii) $y_t = y_{t-1} - 0.50y_{t-2} + e_t - 1.2e_{t-1}$.

Part (a)

Identify each model as an ARMA(p, q) process; that is, specify p and q.

1. We rewrite equation (i),
$$y_t = 0.80\,\mathrm{B}\,y_t - 0.15\,\mathrm{B}^2\,y_t + e_t - 0.30\,\mathrm{B}\,e_t.$$
Now, we rewrite it into the form
$$\begin{aligned}
(1 - 0.8B + 0.15B^2)y_t &= (1 - 0.3B)e_t \\
-20(1 - 0.5B)(1 - 0.3B)y_t &= (1 - 0.3B)e_t \\
-20(1 - 0.5B)y_t &= e_t.
\end{aligned}$$
We see that $y_t = 0.5y_{t-1} - \frac{e_t}{20}$. Two things should be pointed out. First, assuming e_t is symmetric with zero mean, $-\frac{e_t}{20}$ is distributed the same as $\frac{e_t}{20}$. Next, the variance of $\frac{e_t}{20}$ is $\frac{1}{400}\sigma^2$.
We let $W_t = \frac{1}{20}e_t$, and thus
$$y_t = 0.5y_{t-1} + W_t,$$
where $\{W_t\}$ is a zero mean white noise process with variance $\frac{1}{400}\sigma^2$ and $\{y_t\}$ is AR(1).
2. We rewrite equation (ii),
$$y_t = By_t - 0.5B^2y_t + e_t - 1.2Be_t.$$
Now, we rewrite it into the form
$$\begin{aligned}
(1 - B + 0.5B^2)y_t &= (1 - 1.2B)e_t \\
0.5(B - 1 + i)(B - 1 - i)y_t &= (1 - 1.2B)e_t.
\end{aligned}$$
We see that this is an ARMA(2,1) process.

Part (b)

Determine whether each model is stationary and/or invertible.

Time series (i) is AR(1) and is thus invertible. We also know that it is stationary since $|\varphi| = |0.5| < 1$.

Time series (ii) is ARMA(2,1). Let $\varphi(x) = (x - 1 + i)(x - 1 - i)$ which has roots $1 + i$ and $1 - i$, which both modulus $\sqrt{2}$. This is larger than 1, so it is invertible. Let $\theta(x) = 1 - 1.2x$ which has root $0.8\bar{3}$. Since $|0.8\bar{3}| < 1$, it is not stationary.

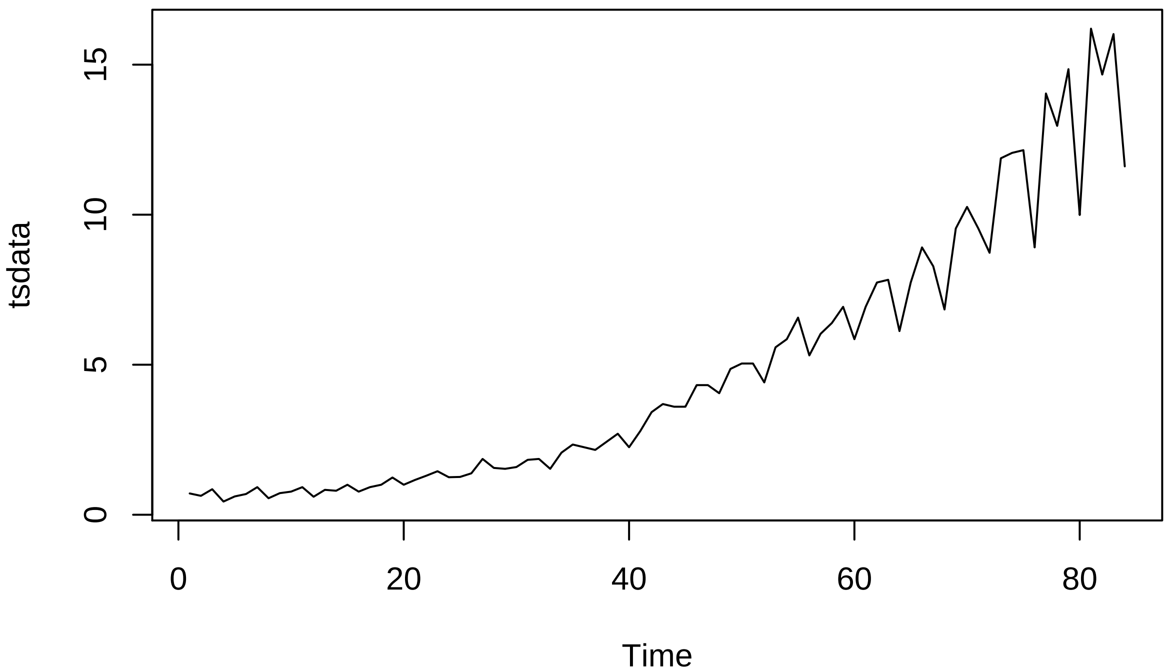
Problem 2.1

The Johnson and Johnson dataset contains quarterly earnings per share for the U.S. company Johnson & Johnson. There are 84 quarters (21 years) measured from the first quarter of 1960 to the last quarter of 1980. To load the dataset, run the following: install.packages("astsa"); library(astsa). The dataset is under the name jj. Do a log transformation of the original time series before answering the following.

Preliminary analysis

We would like to take a look at a simple plot of the data, prior to any transformations.

```
library(astsa)
tsdata <- ts(data=jj)
plot(tsdata)
```



We see that the variance increases over time. The log-transformation will fix this problem, as computed in the following code:

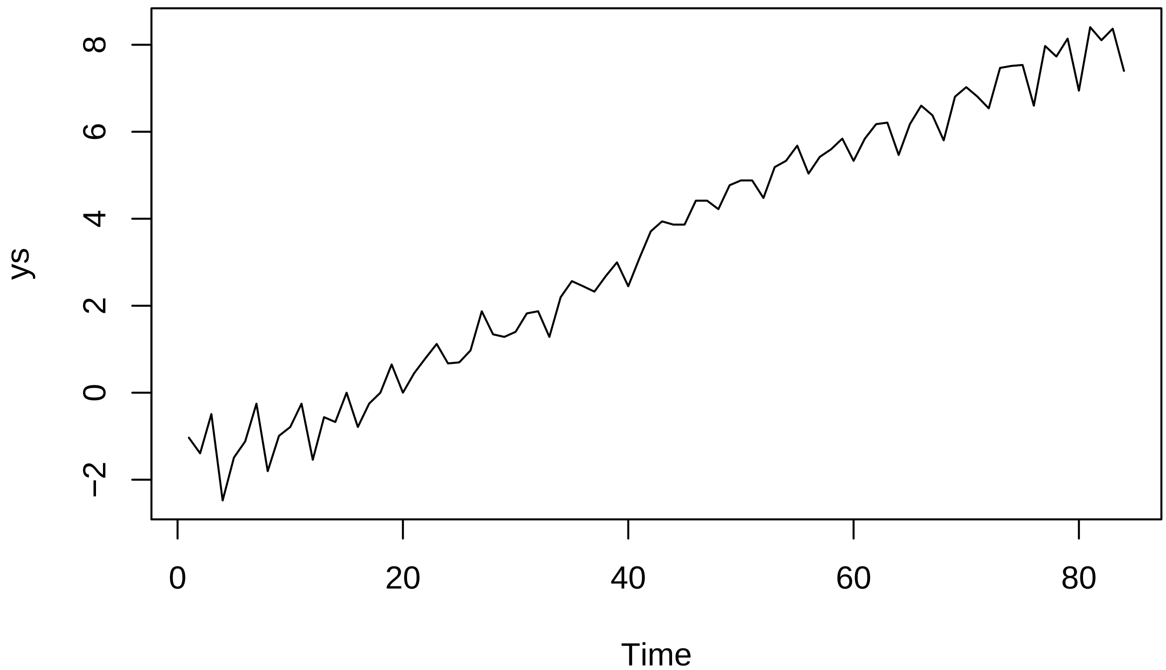
```
n <- length(tsdata)
A <- exp((1/n)*sum(log(tsdata)))
ys <- A*log(tsdata)
```

Part (a)

Construct a time series plot for the logged data. Comment on overall trend and seasonality variation.

We generate the plot with the following R code:

```
plot(ys)
```



The data has both seasonality and a (positive) trend.

Part (b)

Fit the a regression model on the logged data

$$y_t = \beta_0 + \beta_1 t + \alpha_1 Q_2(t) + \alpha_2 Q_3(t) + \alpha_3 Q_4(t) + e_t,$$

where $Q_i(t) = 1$ if time t corresponds to quarter $i = 1, 2, 3$ and zero otherwise. Assume e_t is a normal white noise sequence. Report model coefficients estimates. Superimpose the fitted values on the time plot in part (a). Note: you will need to first create a variable for time and quarter. To do that, you may use: t=1:84; qt=as.factor(rep(1:4,21)).

We perform the model fitting using the following R code:

```
t <- 1:n
qt <- as.factor(rep(1:4,(n/4)))
q1 <- qt==1
q2 <- qt==2
q3 <- qt==3
m <- cbind(t,q1,q2,q3,ys)

# fit regression model to data
fit <- lm(ys~t+q1+q2+q3, data=m)
```

The model coefficients are given by:

```
summary(fit)
```

```
##
## Call:
## lm(formula = ys ~ t + q1 + q2 + q3, data = m)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

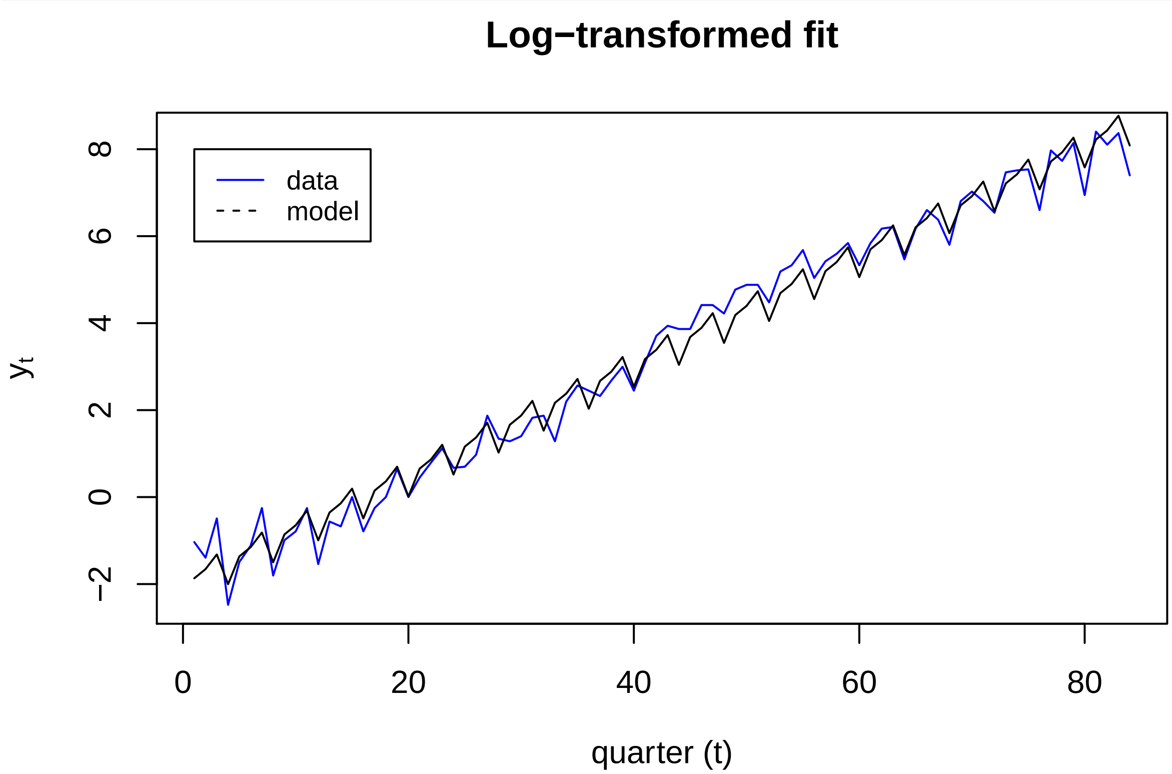
```
## -0.8847 -0.2735 -0.0356  0.2553  0.8342
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.508319   0.111529 -22.490  < 2e-16 ***
## t            0.126112   0.001704  73.999  < 2e-16 ***
## q1           0.514570   0.116866   4.403 3.31e-05 ***
## q2           0.599431   0.116803   5.132 2.01e-06 ***
## q3           0.810985   0.116766   6.945 9.50e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3783 on 79 degrees of freedom
## Multiple R-squared:  0.9859, Adjusted R-squared:  0.9852
## F-statistic:  1379 on 4 and 79 DF,  p-value: < 2.2e-16
```

In other words, the estimate is given by

$$\hat{y}_t = -2.508 + 0.126t + 0.514Q_1(t) + 0.599Q_2(t) + 0.811Q_3(t).$$

The plot of the data with $\hat{y}_t$ superimosed onto it is given by:

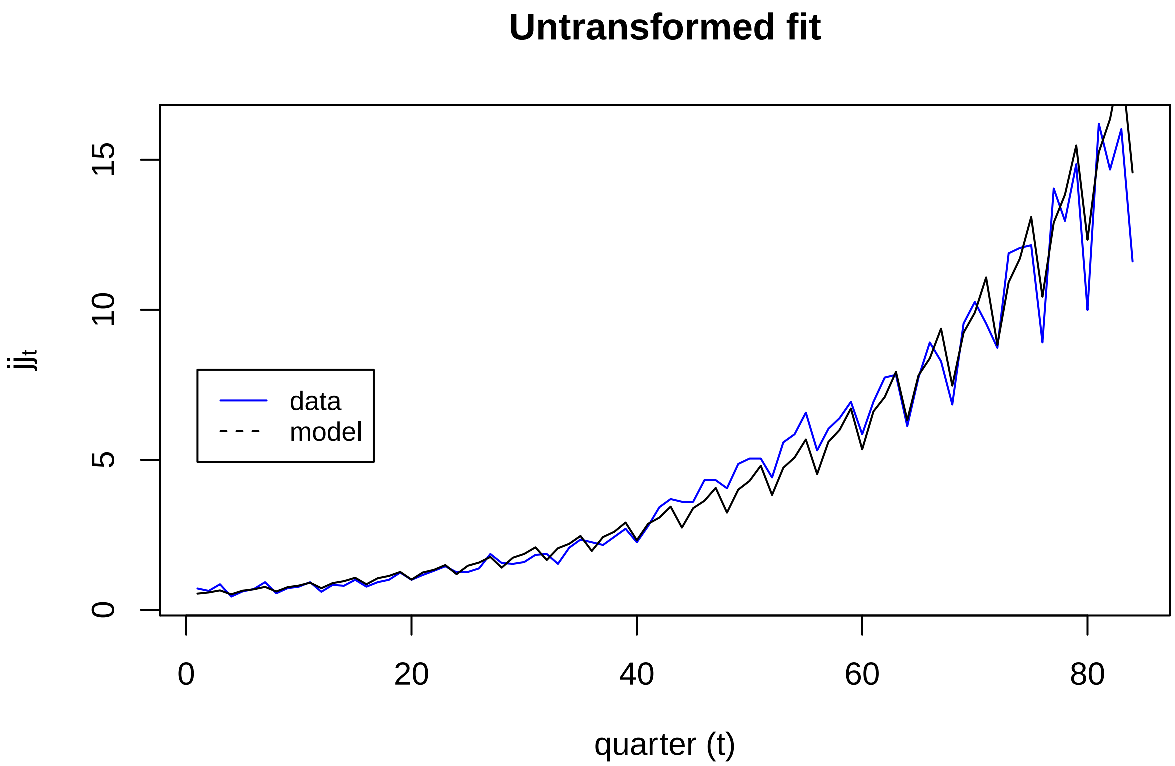
```
library(latex2exp)
plot(ys,col="blue", pch=19,xlab=TeX("quarter ($t$)"),ylab=TeX("$y_t$"),, main="Log-transformed fit")
lines(fitted.values(fit),type="l")
legend(1,8,legend=c("data","model"),col=c("blue","black"),lty=1:2,cex=0.8)
```



We use the following R code to show the model fit with the log-transformation reverse.

```
library(latex2exp)
plot(tsdata,col="blue", pch=19,xlab=TeX("quarter ($t$)"),ylab=TeX("$jj_t$"), main="Untransformed fit")
```

```r
lines(exp(fit$fitted.values/A),type="l")
legend(1,8,legend=c("data","model"),col=c("blue","black"),lty=1:2,cex=0.8)
```



This looks pretty good.

Part (c)

Calculate the MSE. Make a time plot, a ACF plot and a histogram for the residuals. Does the residuals look like a normal white noise process?

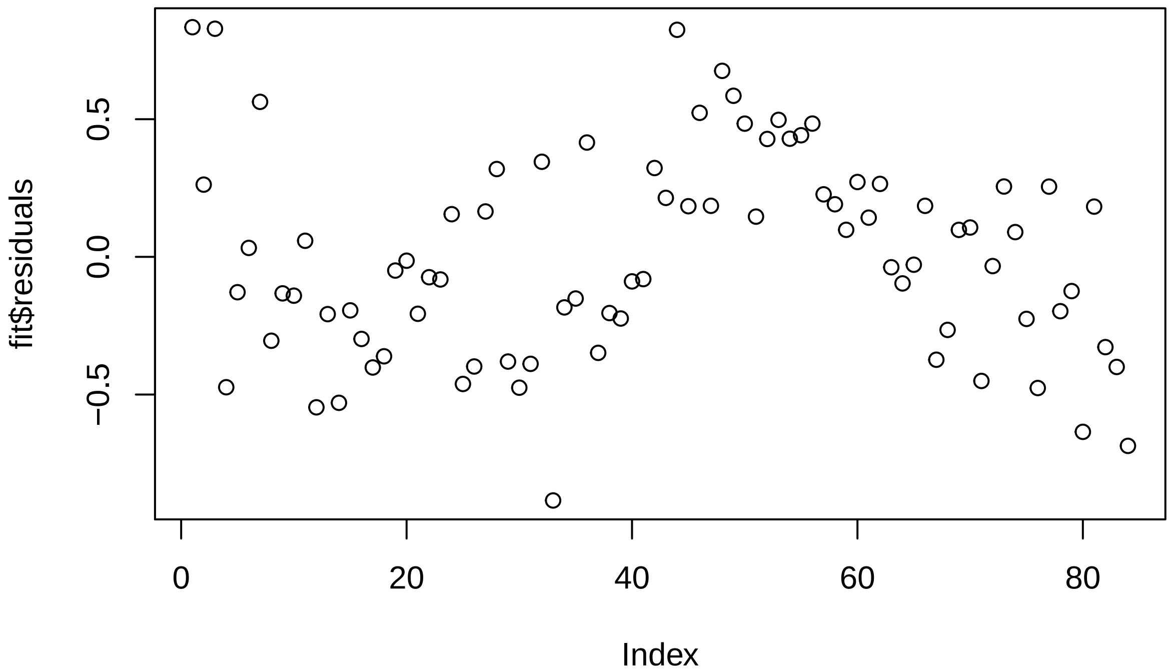
We compute the MSE in two different ways.

```r
df <- length(fit$residuals)-5
sse <- sum(fit$residuals^2)
mse <- sse/df
mse_alt <-summary(fit)$sigma^2  # agrees with mse calculation above
```

We see that the MSE is 0.14313.

The plot of the residuals is given by:

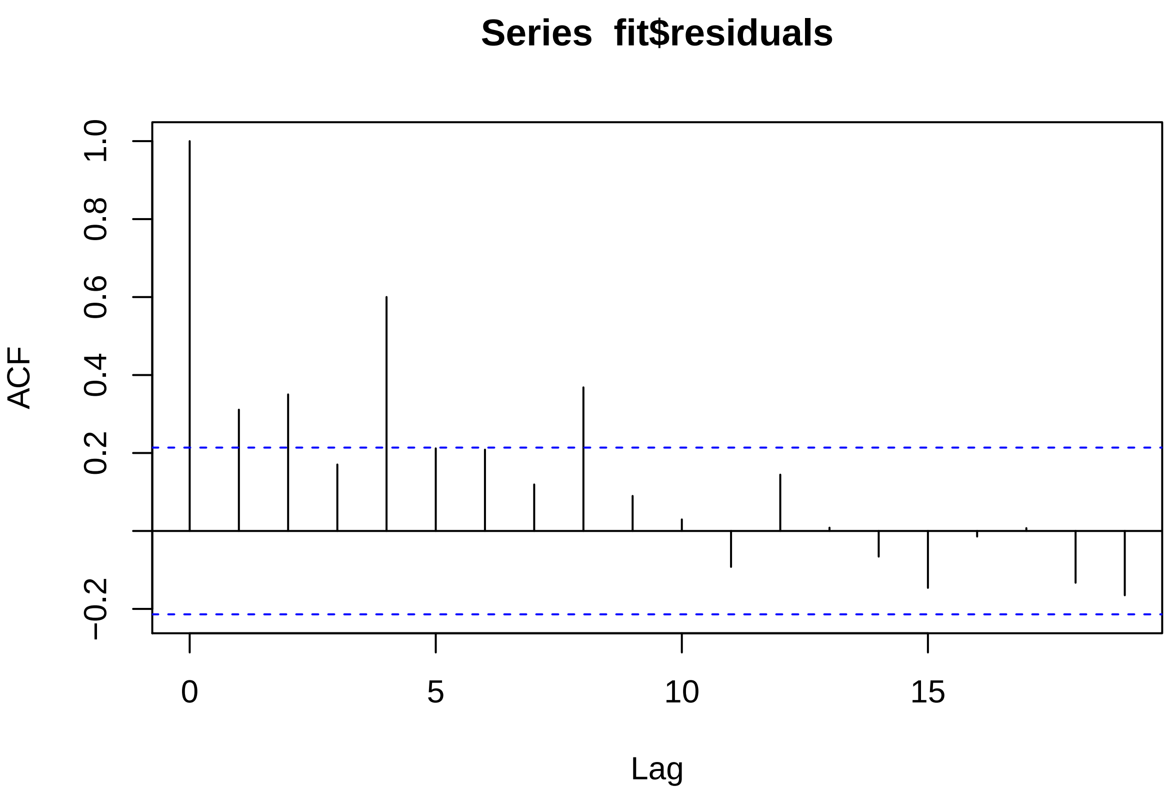
```r
plot(fit$residuals)
```



The residual plot is not especially demonstrative of white noise. In particular, time units 40 to 65 seem to have a non-zero positive expectation. They should hover above and below more or less equally, but the expectation seems to be roughly 0.5 there.

The ACF of the residuals is given by:

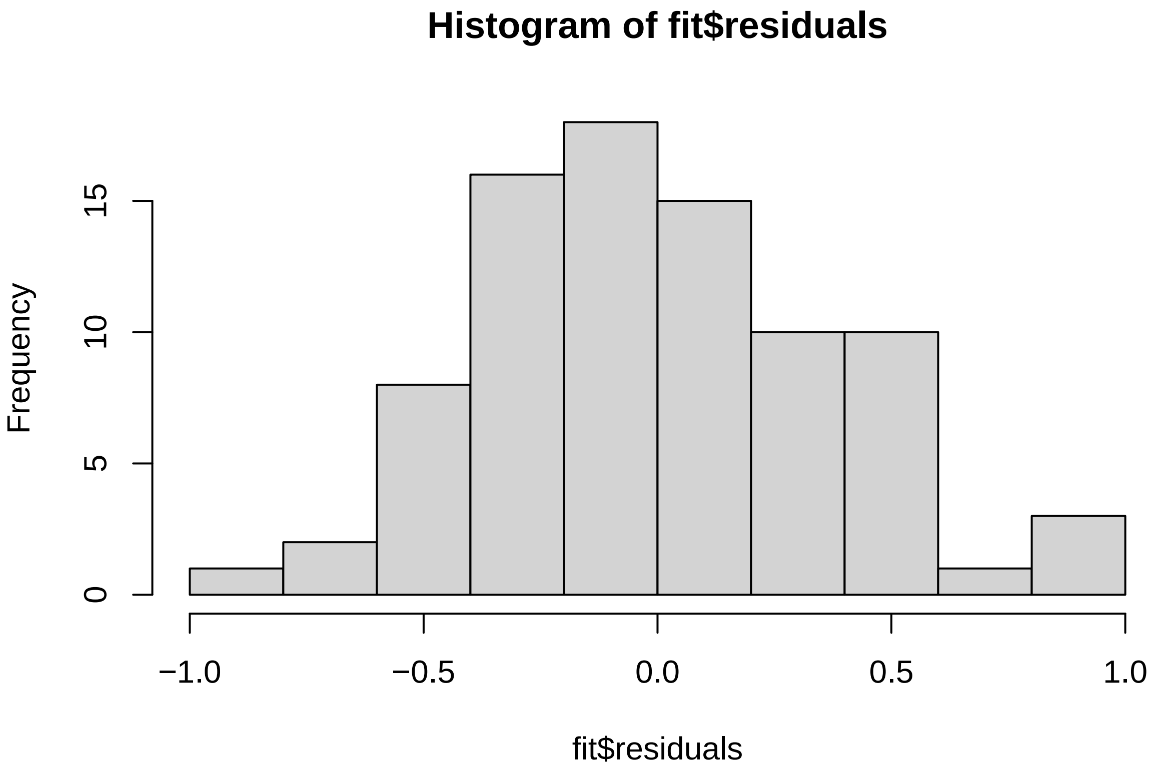
```
acf(fit$residuals)
```



According to the ACF, there seems to be some correlation. Particularly, the periods seem to be correlated, i.e., lags 4 and 8 are positively correlated, but it has positive correlations for other lags also.

The histogram of the residuals is given by:

```
hist(fit$residuals)
```



By itself, this seems alright. It has a mean of around 0 and it is symmetric with a basic bell-shaped curve.

Part (d)

Make predictions for the first quarter in 1981. Construct the 95% prediction interval.

```
part_d_pred <- predict(fit,
                       newdata=data.frame(t=85,q1=1,q2=0,q3=0),
                       interval="prediction")
part_d_pred
```

```
##       fit      lwr      upr
## 1 8.72576 7.940683 9.510837
```

Note that this is the prediction for the log-transformed data. To compute the actual prediction, we must take the inverse of the transformation.

Let

$$A = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log y_i\right)$$

then we transform the data with

$$y_i = A \log \mathrm{jj}_i,$$

where jj is the original data set (prior to the log transformation). If we wish to undo the transformation, we do the inverse,

$$\hat{jj}_i = \exp\left(\frac{y_i}{A}\right).$$

We apply this transformation to the predictions in the following R code:

```
exp(part_d_pred/A)
```

```
##        fit      lwr      upr
## 1 18.02369 13.89481 23.37948
```

Part (e)

Fit a additive model using the Holt-Winters method. Let the function choose the optimal smoothing parameters automatically. Report the smoothing parameters and coefficients. Superimpose the fitted values on the time plot in part (a).

First, we put the data into the format expected by hw by converting the frequency to 4 (4 quarters per year), and then use the Holt-Winters seasonal additive function to generate the model. Finally, we report the model.

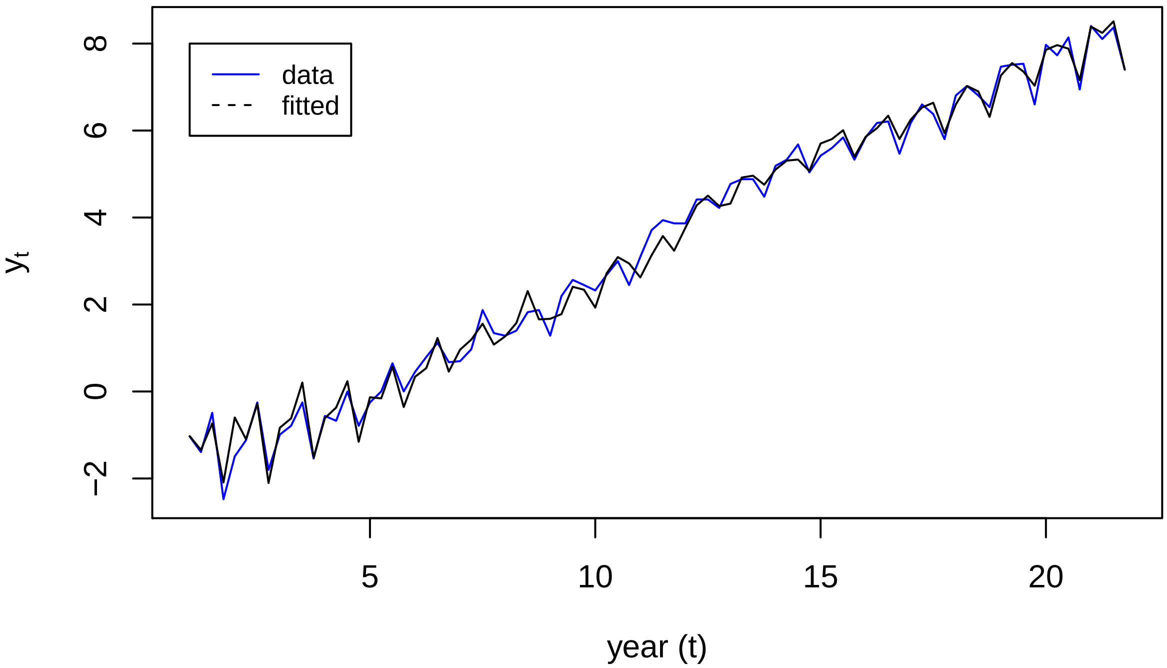
```
library(forecast)
seasonal <- ts(data=ys, frequency=4)
holt_model <- hw(seasonal, level=c(95),h=1, seasonal="additive", initial="optimal")
summary(holt_model)
```

```
##
## Forecast method: Holt-Winters' additive method
##
## Model Information:
## Holt-Winters' additive method
##
## Call:
##  hw(y = seasonal, h = 1, seasonal = "additive", level = c(95),
##
##  Call:
##      initial = "optimal")
##
##   Smoothing parameters:
##     alpha = 0.1731
##     beta  = 1e-04
##     gamma = 0.6741
##
##   Initial states:
##     l = -1.5965
##     b = 0.117
##     s = -0.9962 0.5214 0.0216 0.4532
##
##   sigma:  0.2777
##
##      AIC     AICc      BIC
## 166.5136 168.9461 188.3910
##
## Error measures:
##                         ME      RMSE       MAE MPE MAPE      MASE      ACF1
## Training set 0.0001060719 0.2641076 0.2055873 NaN  Inf 0.4251315 0.1068085
##
## Forecasts:
##       Point Forecast    Lo 95    Hi 95
## 22 Q1       8.821088 8.276884 9.365292
```

When we view this as a regression, we get, for instance $\hat{\beta}_{0,T} = -1.5965$ and $\hat{\beta}_{1,T} = 0.117$. The seasonal coefficients can be found in the s initial state.

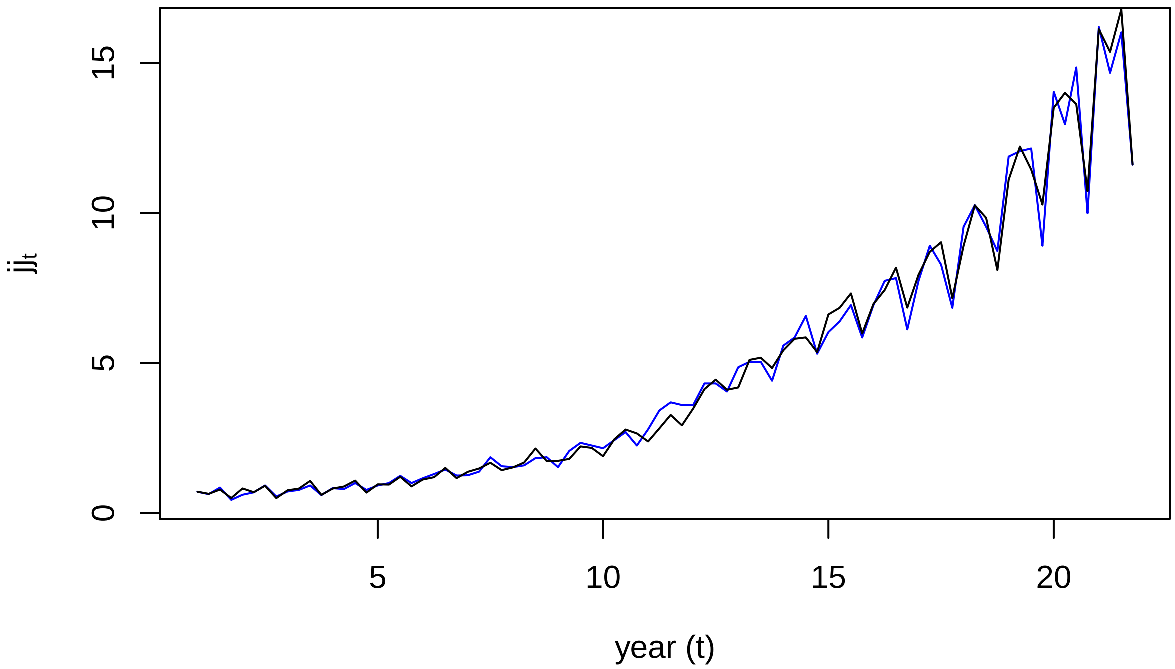
We plot the seasonal time series (now time is with respect to years, four quarters per year) with the fitted model superimposed onto it with the following R code:

```
library(latex2exp)
plot(seasonal,col="blue",pch=19,xlab=TeX("year ($t$)"),ylab=TeX("$y_t$"))
lines(holt_model$fitted)
legend(1,8,legend=c("data","fitted"),col=c("blue","black"),lty=1:2,cex=0.8)
```



Now we show a plot of the untransformed fit.

```
library(latex2exp)
plot(exp(seasonal/A),col="blue",pch=19,xlab=TeX("year ($t$)"),ylab=TeX("$jj_t$"))
lines(exp(holt_model$fitted/A))
```



Part (f)

> Calculate the MSE for the Holt-Winters additive model and compare it with part (c). Make a time plot, a ACF plot and a histogram for the residuals. Does the residuals look like a normal white noise process?

The MSE for the Holt-Winters model is given by

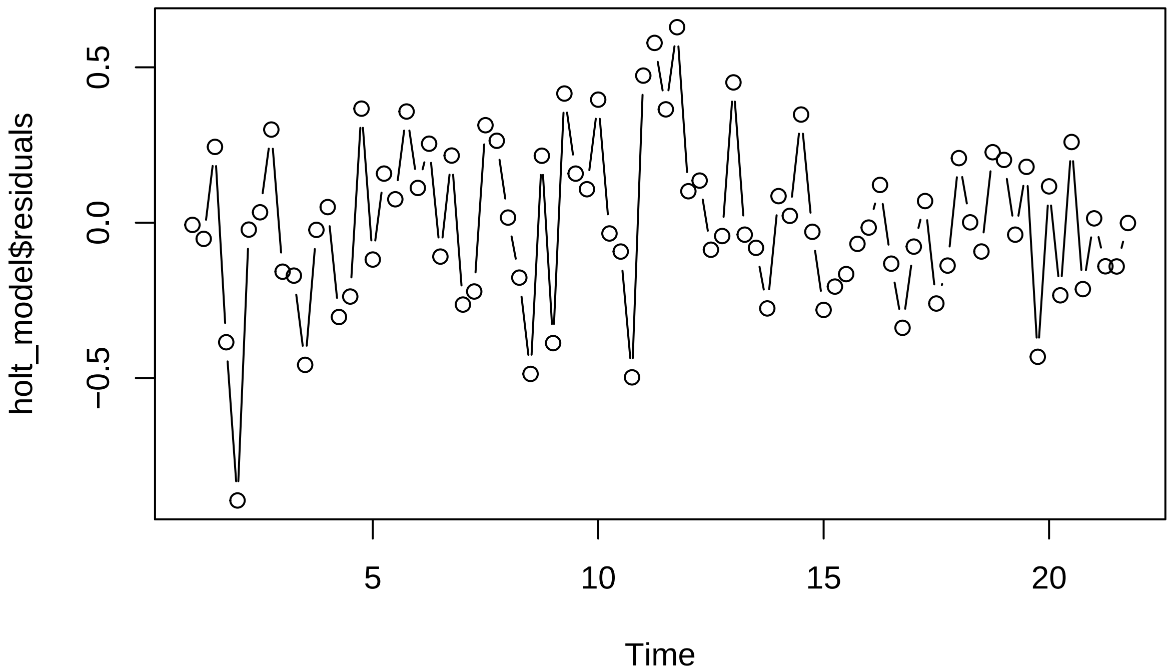
```
round(holt_model$model$mse,digits=3)
```

```
## [1] 0.07
```

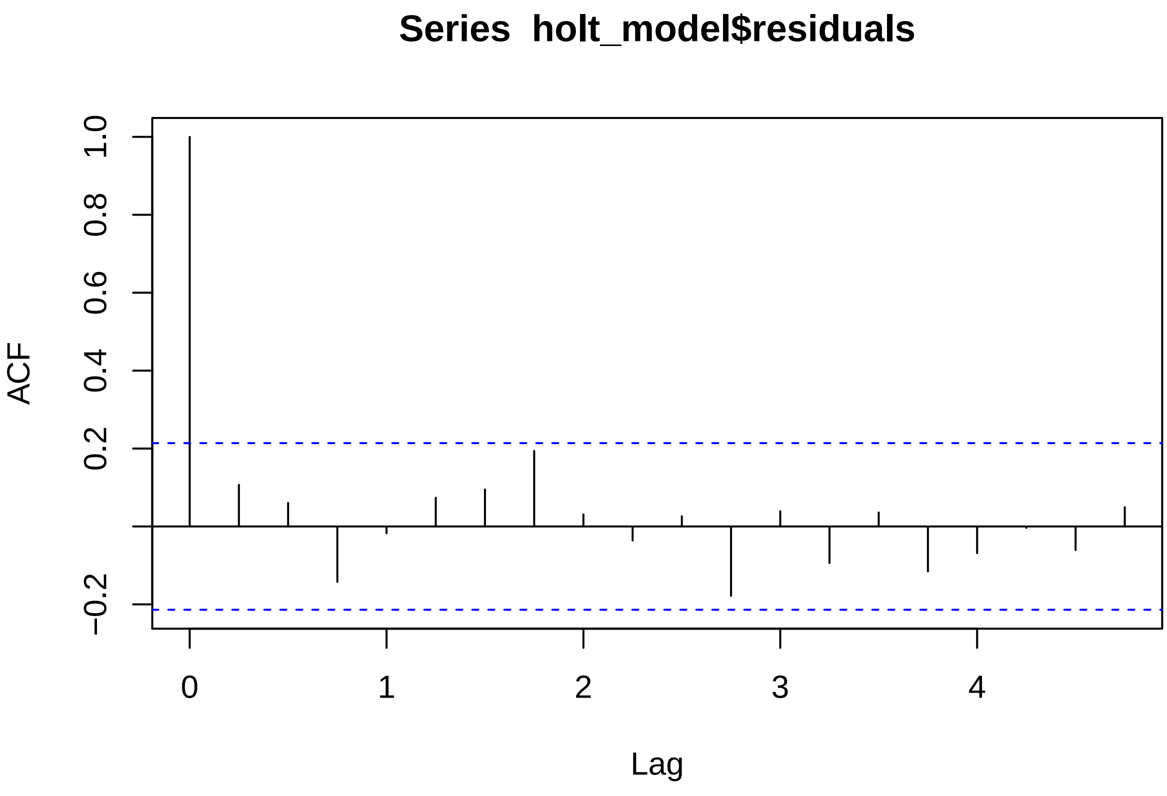
In part (d), we had an MSE of 0.143, which is larger than the MSE for the Holt fit's MSE of 0.07. This suggests a better fit was obtained by the Holt method.

We plot the residuals using the R code:

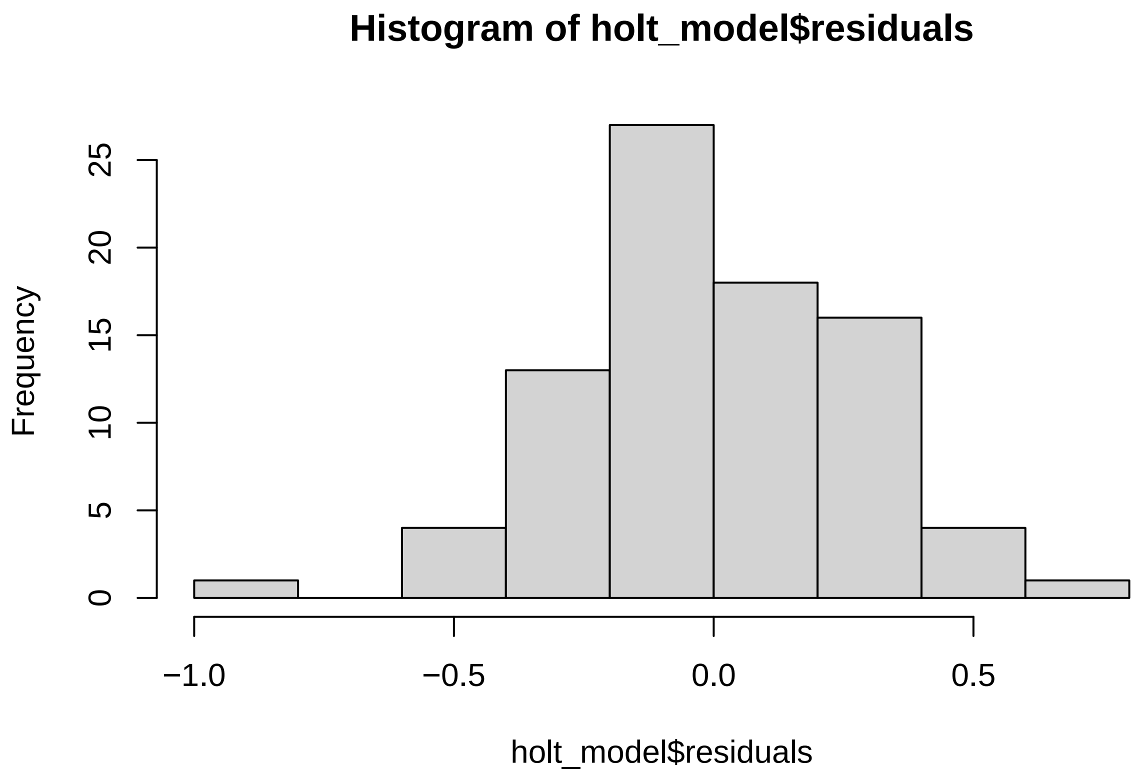
```
plot(holt_model$residuals,type="b")
```




```
acf(holt_model$residuals)
```




```
hist(holt_model$residuals)
```



These residuals are compatible with white noise. The ACF is compatible with lag times being uncorrelated, the plot seems to jump up and down around the mean with constant variance, and the histogram is more or less symmetric with a bell shape and indicates a mean of 0.

Part (g)

Make forecast using the Holt-Winters model for the first quarter in 1981. Also report the 95% prediction interval. Compare the result with part (d).

From the summary output in party (a), the forecast is 8.821 with a 95% prediction interval

$$(8.277, 9.365).$$

By comparison, with the linear regression in part (d), we obtained the prediction interval given by:

```
part_d_pred
```

```
##       fit      lwr      upr
## 1 8.72576 7.940683 9.510837
```

The interval in part (d) is a bit larger (the prediction has more uncertainty), but they are in reasonably close agreement.

The simpler smoothing model seems to model the data relatively well with less uncertainty. Needing to estimate the extra parameters in the regression model may be what led to its larger variance.

We now show the prediction intervals with the log-transformation undone:

```
exp(part_d_pred/A)
```

```
##        fit      lwr      upr
## 1 18.02369 13.89481 23.37948
```

```
exp(c(holt_model$mean,holt_model$lower,holt_model$upper)/A)
```

```
## [1] 18.60217 15.53244 22.27858
```

Problem 2.2

Suppose that $\{e_t\}$ is a zero mean white noise process with variance σ^2. Let B denote the backshift operator. Consider the processes:

(i) $(1+0.4\,\mathrm{B})Y_t = e_t$

(ii) $(1-0.9\,\mathrm{B})(1-\mathrm{B})Y_t = (1-0.5\,\mathrm{B})(1+0.4\,\mathrm{B})e_t$

(iii) $(1-0.4\,\mathrm{B}-0.45\,\mathrm{B}^2)Y_t = (1+\mathrm{B}+0.25\,\mathrm{B}^2)e_t$

Part (a)

Identify each model as an ARMA(p,q) process; that is, specify p and q.

1. The process
$$(1+0.4\,\mathrm{B})Y_t = e_t$$
has no redundancies in its representation. We can determine the ARIMA model from this representation, but for clarity we rewrite the process into its "canonical" form,
$$Y_t = -0.4Y_{t-1} + e_t.$$
We see that it is a zero mean ARMA$(p=1,q=0) \equiv$ AR(1) process.

2. The process
$$(1-0.9\,\mathrm{B})(1-\mathrm{B})Y_t = (1-0.5\,\mathrm{B})(1+0.4\,\mathrm{B})e_t.$$
has no redundancies in its representation. We can determine the ARIMA model from this representation, but for clarity we seek its "canonical" form by expanding the factored form into
$$(1-1.9\,\mathrm{B}+0.9\,\mathrm{B}^2)Y_t = (1-0.1\,\mathrm{B}-0.2\,\mathrm{B}^2)e_t.$$
and then rewriting as
$$Y_t = 1.9Y_{t-1} - 0.9Y_{t-2} + e_t - 0.1e_{t-1} - 0.2e_{t-2},$$
which models a zero mean ARIMA$(p=2,q=2)$ process.

3. The process
$$(1-0.4\,\mathrm{B}-0.45\,\mathrm{B}^2)Y_t = (1+\mathrm{B}+0.25\,\mathrm{B}^2)e_t.$$
can be factored as
$$(1-0.9\,\mathrm{B})(1+0.5\,\mathrm{B})Y_t = (1+0.5\,\mathrm{B})(1+0.5\,\mathrm{B})e_t,$$
which shows that the representation has redundancy. We may remove this redundancy by dividing both sides by the common factor $1+0.5\,\mathrm{B}$, resulting in the equivalent expression
$$(1-0.9\,\mathrm{B})Y_t = (1+0.5\,\mathrm{B})e_t.$$
We can determine the ARIMA model from this representation, but for clarity we rewrite the process into its "canonical" form,
$$Y_t = 0.9Y_{t-1} + e_t + 0.5e_{t-1},$$
which models a zero mean ARIMA$(p=1,q=1)$ process.

Part (b)

Give the autocorrelation function ρ_k for those processes which are stationary. If you want, you can use the ARMAacf function in R to see the first dozen or so correlations.

Process (i)

Process $Y_t = -0.4Y_{t-1} + e_t$ is stationary since $|-0.4| < 1$. Its autocorrelation function is

$$\rho_k = (-0.4)^k.$$

We output the first dozen lags using the following R code:

```
p1 <- ARMAacf(ar=c(-.4),lag.max=12,pacf=F)
p1
```

```
##             0             1             2             3             4
##  1.000000e+00 -4.000000e-01  1.600000e-01 -6.400000e-02  2.560000e-02
##             5             6             7             8             9
## -1.024000e-02  4.096000e-03 -1.638400e-03  6.553600e-04 -2.621440e-04
##            10            11            12
##  1.048576e-04 -4.194304e-05  1.677722e-05
```

Process (ii)

Process $Y_t = 1.9Y_{t-1} - 0.9Y_{t-2} + e_t - 0.1e_{t-1} - 0.2e_{t-2}$ has the more convenient representation

$$(1 - 0.9\,\mathrm{B})(1 - \mathrm{B})Y_t = (1 - 0.5\,\mathrm{B})(1 + 0.4\,\mathrm{B})e_t.$$

for determining stationarity condition. We let $\varphi(x) = (1 - 0.9x)(1 - x)$ and see that the roots of φ are $\{1, 10/9\}$. Since one of the roots is not greater than 1, the process is non-stationary.

Process (iii)

Process $Y_t = 0.9Y_{t-1} + e_t + 0.5e_{t-1}$ has the more convenient representation

$$(1 - 0.9\,\mathrm{B})Y_t = (1 + 0.5\,\mathrm{B})e_t.$$

for determining stationary condition. We let $\varphi(x) = (1 - 0.9x)$ and see that the root of φ is 10/9, which is greater than 1 and therefore the process is stationary.

The autocorrelation function is given by

$$\rho_k = 0.6984(0.9)^k$$

for $k > 0$.

Proof. For an ARIMA$(1, 1)$ model, its form is given by

$$Y_t = \varphi Y_{t-1} + e_t - \theta e_{t-1}.$$

We match this to $Y_t = 0.9Y_{t-1} + e_t - (-0.5)e_{t-1}$ and see that $\varphi = 0.9$ and $\theta = -0.5$.

Note that

$$\begin{aligned}\gamma_0 &= \sigma^2\left(1 + \frac{(\varphi+\theta)^2}{1-\varphi^2}\right)\\ &= \sigma^2\left(1 + \frac{(0.9-0.5)^2}{1-0.9^2}\right)\\ &= 1.8421\sigma^2\end{aligned}$$

and

$$\begin{aligned}\gamma_1 &= \sigma^2 \left(\varphi + \theta + \varphi \frac{(\varphi+\theta)^2}{1-\varphi^2} \right) \\ &= 1.1579\sigma^2\end{aligned}$$

Then, $\rho_1 = \gamma_0/\gamma_1 = 0.6286$ and

$$\rho_k = \varphi^{k-1}\rho_1 = 0.6984(0.9)^k$$

for $k > 0$. □

We output the first dozen lags using the following R code:

```
ARMAacf(ar=c(.9),ma=c(-.5),lag.max=12,pacf=F)
```

```
##         0         1         2         3         4         5         6         7
## 1.0000000 0.6285714 0.5657143 0.5091429 0.4582286 0.4124057 0.3711651 0.3340486
##         8         9        10        11        12
## 0.3006438 0.2705794 0.2435215 0.2191693 0.1972524
```

Part (c)

Simulate a data set from each process identified above show the data set in a time series plot and also show the sample ACF for it. In each case, does the time plot agree with the stationarity? Does the sample ACF agree with what we know to be true from the theory? You may pick your favorite sample size (anything larger than 100) and the white noise variance.

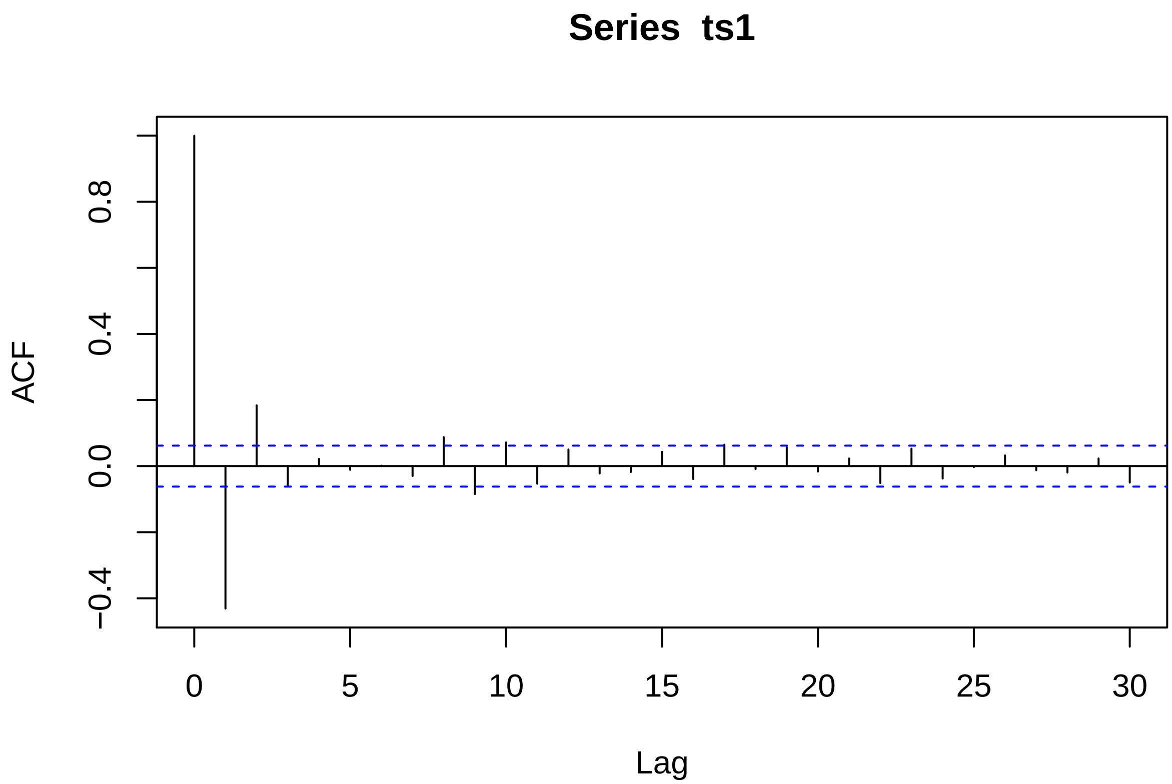
Process (i)

The process defined as

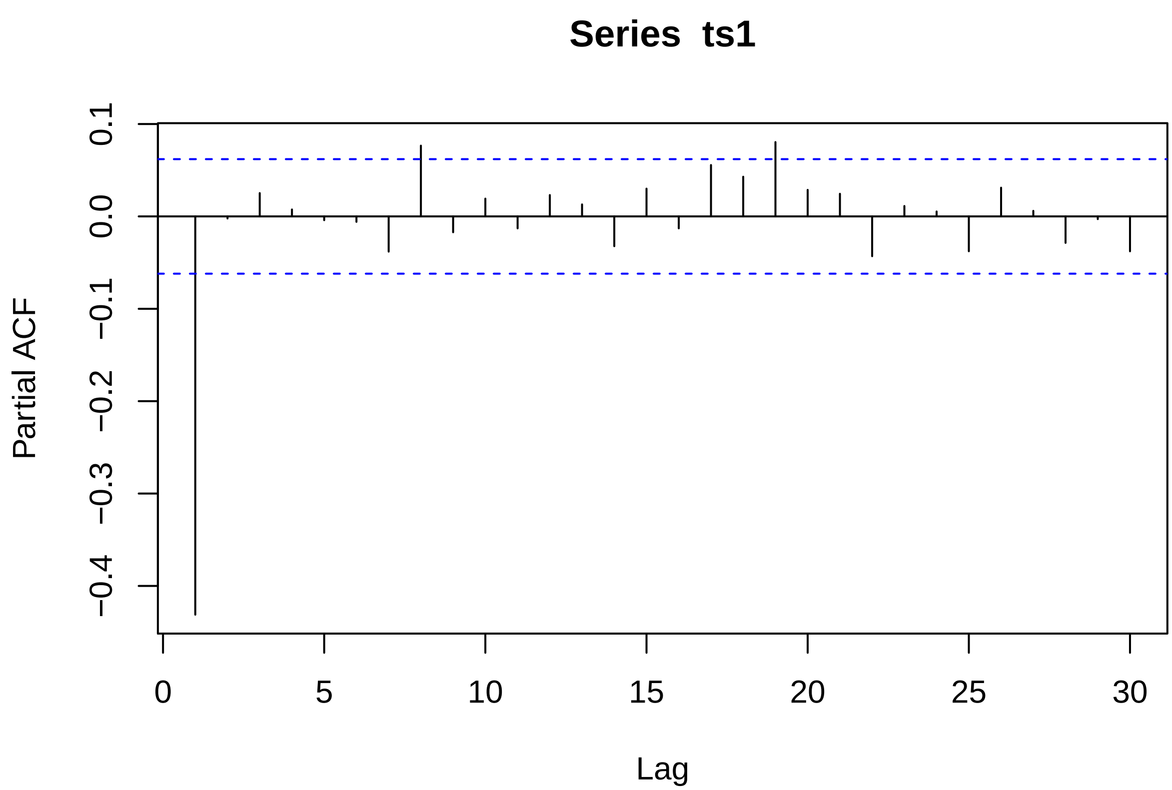
$$Y_t = -0.4Y_{t-1} + e_t.$$

is an AR(1) model. We simulate drawing a sample and plotting the sample ACF with the following R code:

```
sigma <- 1
ts1 <- arima.sim(n = 1000,
          list(ar = c(-0.4)),
          sd = sigma)
acf(ts1)
```




```
pacf(ts1)
```



We see that the sample ACF of the AR(1) process has the following characteristics:

1. Oscillates as expected since it has a negative coefficient -0.4.

2. Seems to exponentially decay, as expected of an AR model.

More importantly, the sample PACF is compatible with a non-zero correlation only at lag 1, as expected.

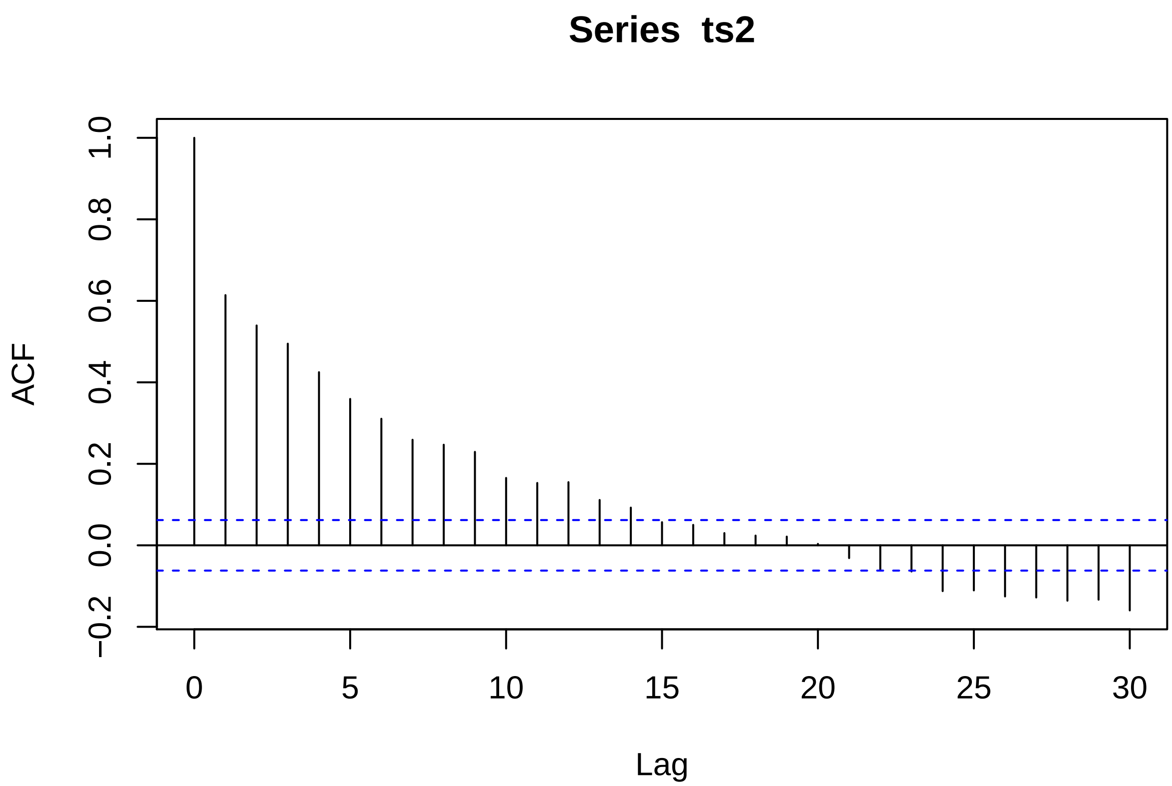
Process (iii)

The process defined as

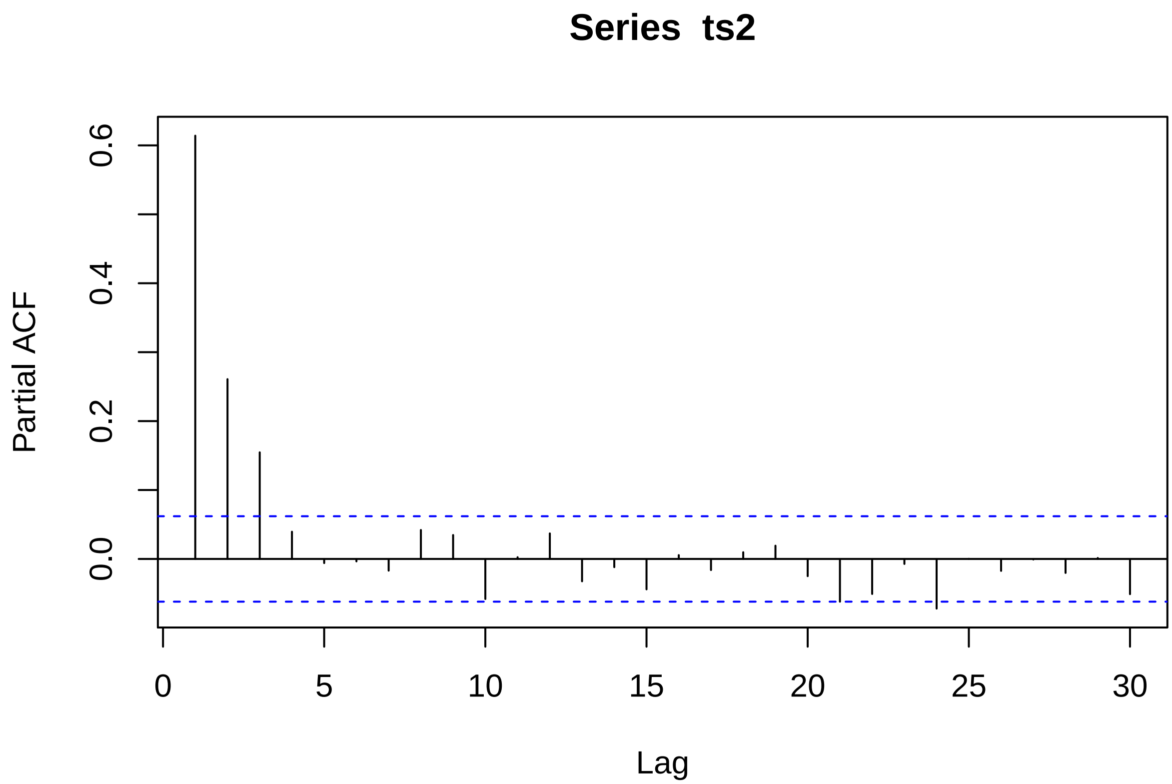
$$Y_t = 0.9Y_{t-1} + e_t - (-.5)e_{t-1}.$$

is an ARMA$(1, 1)$ model. We simulate drawing a sample and plotting the sample ACF with the following R code:

```
ts2 <- arima.sim(n = 1000,
                 list(ar = c(0.9),
                      ma = c(-0.5)),
                 sd = sigma)
acf(ts2)
```




```
pacf(ts2)
```



We see that the sample ACF of ARMA$(1, 1)$ process has the following characteristics:

1. ACF has no cut-off. The observed correlation is compatible with being non-zero for many lags, as expected for a time series with an autoregressive component. This looks like an AR process.

2. Both the ACF and PACF exponentially decays (with a damped sinusoidal component), as expected. It is a mixture of AR and MA.