Question 1

3.5 Let $\{Y_t\}$ be the ARMA plus noise time series defined by

$$Y_t = X_t + W_t,$$

where $\{W_t\} \sim \text{WN}(0, \sigma_w^2)$, $\{X_t\}$ is the ARMA(p, q) process satisfying

$$\phi(B)X_t = \theta(B)Z_t, \quad \{Z_t\} \sim \text{WN}(0, \sigma_z^2),$$

and $E(W_s Z_t) = 0$ for all s and t.

(a) Show that $\{Y_t\}$ is stationary and find its autocovariance function in terms of σ_W^2 and the ACVF of $\{X_t\}$.

AVCF of $\{X_t\}$: $X_t - \phi_1 X_{t-1} \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$

$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \cdots + \theta_q Z_{t-q}$.

$E[Y_t] = E[X_t] + E[W_t] = \mu_x$ is independent of t

$\text{Cov}(Y_t, Y_{t+h}) = E[Y_t \cdot Y_{t+h}] = E[(X_t + W_t)(X_{t+h} + W_{t+h})]$

$= E[X_t X_{t+h} + X_t W_{t+h} + W_t X_{t+h} + W_t W_{t+h}]$

$= E[X_t \cdot X_{t+h}] + E[X_t \cdot W_{t+h}] + E[W_t X_{t+h}] + E[W_t W_{t+h}]$

Since $E[W_s Z_t] = 0 \Rightarrow E[W_s X_t] = 0$

then if $h = 0$: $\text{Cov}(Y_t, Y_{t+h}) = E[X_t^2] + E[W_t^2] = \gamma_x(0) + \sigma_w^2$

if $h > 0$: $\text{Cov}(Y_t, Y_{t+h}) = E[X_t \cdot X_{t+h}] = \gamma_x(h)$

(b) Show that the process $U_t := \phi(B)Y_t$ is r-correlated, where $r = \max(p, q)$ and hence, by Proposition 2.1.1, is an MA(r) process. Conclude that $\{Y_t\}$ is an ARMA(p, r) process.

$U_t = \phi(B) Y_t = \phi(B)[X_t + W_t] = \theta(B) Z_t + \phi(B) W_t$

$= (Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}) + (W_t - \phi_1 W_{t-1} \cdots - \phi_p W_{t-p})$

If $p > q$ then $\theta_{q+1}, \ldots, \theta_r = 0$

If $q > p$ then $\phi_{p+1}, \ldots, \phi_r = 0$

Then we can rewrite $Cov(U_t, U_{t+h})$ as follows:

$$Cov(U_t, U_{t+h}) = E[(Z_t + \cdots + \theta_q Z_{t-q} + W_t - \cdots - \phi_p W_{t-p})(Z_{t+h} + \cdots + \theta_q Z_{t+h-q} + W_{t+h} - \cdots - \phi_p W_{t+h-p})]$$

$$= E[(Z_t + \cdots + \theta_r Z_{t-r} + W_t - \cdots - \phi_r W_{t-r})(Z_{t+h} + \cdots + \theta_r Z_{t+h-r} + W_{t+h} - \cdots - \phi_r W_{t+h-r})]$$

Then for $h > r$ $\quad E[U_t] = 0 \quad Cov(U_t, U_{t+h}) = E[U_t \cdot U_{t+h}]$

Since $t+h-q > t$ and $t+h-p > t$ and $E[Z_t \cdot Z_s] = 0$ for $t \neq s$ and $E[W_t \cdot W_s] = 0$ for $t \neq s$

$\Rightarrow \quad Cov(U_t, U_{t+h}) = 0$

However, if $h \leq r$:

$Cov(U_t, U_{t+h})$

$$= E[\theta_{r-h}\theta_r Z_{t+h-r}^2 + \cdots + \theta_0 \cdot \theta_h Z_{t-r+r}^2] + E[\phi_{r-h}\phi_r W_{t+h-r}^2 + \cdots + \phi_h \cdot \phi_0 W_{t-r+r}^2]$$

$$= \sum_{i=0}^{r-h} \theta_i \theta_{h+i} \gamma_Z(0) + \sum_{i=0}^{r-h} \phi_i \phi_{h+i} \gamma_W(0)$$

$$= \sum_{i=0}^{r-h} \theta_i \theta_{h+i} \sigma_Z^2 + \sum_{i=0}^{r-h} \phi_i \phi_{h+i} \sigma_W^2$$ is uncorrelated with t.

$\Rightarrow$ U_t is r-correlated, hence U_t is an MA(R) process, so

$U_t = Z_t + \theta_1 Z_{t-1} + \cdots \theta_r Z_{t-r}$ for some $\{Z_t\}$ and $\underset{\sim}{\theta}$

$\Rightarrow \quad \phi(B) Y_t = \Phi(B) Z_t$

$\Rightarrow \quad Y_t$ is an ARMA(p, r) process

Question 2

3.7 Suppose that $\{X_t\}$ is the noninvertible MA(1) process

$$X_t = Z_t + \theta Z_{t-1}, \quad \{Z_t\} \sim \mathrm{WN}\left(0, \sigma^2\right),$$

where $|\theta| > 1$. Define a new process $\{W_t\}$ as

$$W_t = \sum_{j=0}^{\infty} (-\theta)^{-j} X_{t-j}$$

and show that $\{W_t\} \sim \mathrm{WN}\left(0, \sigma_W^2\right)$. Express σ_W^2 in terms of θ and σ^2 and show that $\{X_t\}$ has the *invertible* representation (in terms of $\{W_t\}$)

$$X_t = W_t + \frac{1}{\theta} W_{t-1}.$$

$$E[X_t] = E[Z_t + \theta Z_{t-1}] = 0$$

$$E[X_t \cdot X_{t+h}] = E[(Z_t + \theta Z_{t-1})(Z_{t+h} + \theta Z_{t+h-1})]$$

$$= E[Z_t \cdot Z_{t+h} + \theta Z_{t-1} Z_{t+h} + \theta Z_t Z_{t+h-1} + \theta^2 Z_{t-1} Z_{t+h-1}]$$

then $\gamma_X(h) = \begin{cases} \sigma^2 + \theta^2 \sigma^2 & \text{for } h = 0 \\ \theta \sigma^2 & \text{for } h = 1 \\ \theta \sigma^2 & \text{for } h = -1 \end{cases}$

$$\Rightarrow E[W_t] = \sum_{j=0}^{\infty} (-\theta)^{-j} E[X_{t-j}] = 0$$

$$\mathrm{Cov}(W_t, W_{t+h}) = E[W_t \cdot W_{t+h}]$$

$$= E\left[\left(\sum_{j=0}^{\infty} (-\theta)^{-j} X_{t-j}\right)\left(\sum_{k=0}^{\infty} (-\theta)^{-j} X_{t+h-k}\right)\right]$$

$$= E\left[\sum_{k=0}^{\infty} \sum_{j=0}^{\infty} (-\theta)^{-j} (-\theta)^{-k} X_{t-j} X_{t+h-k}\right]$$

$$= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \gamma_X(h-k+j)$$

$$= \sum_{k=h+1}^{\infty} (-\theta)^{-2k+h+1} \sigma^2 \theta + \sum_{\substack{k=h-1, \\ k>0}}^{\infty} (-\theta)^{-2k+h-1} \sigma^2 \theta + \sum_{k=h}^{\infty} (-\theta)^{-2j+k} \sigma^2 (1+\theta^2)$$

$$= \sigma^2 (1+\theta^2)(-\theta)^{-h} \sum_{k=h}^{\infty} (-\theta)^{-2k+h} + (-\theta)^{-h+1} \sigma^2 \theta \sum_{k=h-1}^{\infty} (-\theta)^{-2k+2(h-1)} + (-\theta)^{-h-1} \theta \sigma^2 \sum_{k=h+1}^{\infty} (-\theta)^{-2k+2h+2}$$

$$= \sigma^2(-\theta)^{-h}\frac{\theta^2}{\theta^2-1} + \sigma^2\theta^2(-\theta)^{h}\frac{\theta^2}{\theta^2-1} + \sigma^2\theta(-\theta)^{-h}(-\theta)^{-1}\frac{\theta^2}{\theta^2-1} + \sigma^2\theta(-\theta)^{-h}\cdot -\theta\frac{\theta^2}{\theta^2-1}$$

$$= \sigma^2(-\theta)^{h}\frac{\theta^2}{\theta^2-1}(\cancel{1}+\cancel{\theta^2}-\cancel{\theta^2}-\cancel{1}) + \sigma^2\theta^2\,\mathbb{1}_{\{h=0\}}$$

$$= \sigma^2\theta^2\,\mathbb{1}_{\{h=0\}}$$

$$\Rightarrow \gamma_W(h) = \begin{cases} \sigma^2\theta^2 & \text{if } h=0 \\ 0 & \text{if } h\neq 0 \end{cases} \qquad \Rightarrow W \sim WN(0, \sigma^2\theta^2)$$

W_t is a White Noise Process.

Now let $\pi_j = (-\theta)^{-j}$ then W_t can be writen as $W_t = \sum_{j=0}^{\infty}\pi_j X_{t-j}$. Also $\sum_{j=0}^{\infty}\theta^{-j} < \infty$. Since invertible and we want $\phi(B)X_t = \theta(B)Z_t$

$$\pi(B) = \sum_{j=0}^{\infty}\frac{\phi(B)}{\theta(B)} \Rightarrow \sum_{j=0}^{\infty}\pi_j z^j = \frac{\phi(z)}{\theta(z)}$$

$$\Rightarrow \phi(z) = 1 \text{ and } \theta(z) = 1+\frac{z}{\theta} \Rightarrow X_t = W_t + \theta^{-1}W_{t-1}$$

Question 3

Notes for 3.9 below:

(i) The accidental deaths data is available in the *itsmr* package.

(ii) You can save the sample autocorrelations from *ggAcf*. For example, using the *lynx* data, we have that:

(iii) Using *diff* twice (appropriately) with different lags will allow you to obtain the series of interest.

3.9 (a) Calculate the autocovariance function $\gamma(\cdot)$ of the stationary time series

$$Y_t = \mu + Z_t + \theta_1 Z_{t-1} + \theta_{12}Z_{t-12}, \quad \{Z_t\} \sim \text{WN}(0, \sigma^2).$$

$$E[Y_t] = E[\mu + Z_t + \theta_1 Z_{t-1} + \theta_{12}Z_{t-12}] = \mu$$

$$\text{Cov}(Y_t, Y_{t+h}) = E[(Y_t-\mu)(Y_{t+h}-\mu)]$$

$$= E[Y_tY_{t+h} - \mu Y_t - \mu Y_{t+h} + \mu^2]$$

$$= E[Y_tY_{t+h}] - \mu^2 - \cancel{\mu^2} + \cancel{\mu^2}$$

Then:

$$E[Y_t \cdot Y_{t+h}] = E[(\mu + Z_t + \theta_1 Z_{t-1} + \theta_{12} Z_{t-12})(\mu + Z_{t+h} + \theta_1 Z_{t+h-1} + \theta_{12} Z_{t+h-12})]$$

$$= \mu^2 + E[Z_t(\mu + Z_{t+h} + \theta_1 Z_{t+h-1} + \theta_{12} Z_{t+h-12}) + E[\theta_1 Z_{t-1}(\mu + Z_{t+h} + \theta_1 Z_{t+h-1} + \theta_{12} Z_{t+h-12})$$

$$+ E[\theta_{12} Z_{t-12}(\mu + Z_{t+h} + \theta_1 Z_{t+h-1} + \theta_{12} Z_{t+h-12})]$$

then if $h = 0$:

$$E[Y_t \cdot Y_{t+h}] = \mu^2 + \sigma^2 + \theta_1^2 \sigma^2 + \theta_{12}^2 \sigma^2$$

if $h = -1$:

$$E[Y_t \cdot Y_{t+h}] = \mu^2 + \theta_1 \sigma^2$$

if $h = -12$:

$$E[Y_t \cdot Y_{t+h}] = \mu^2 + \theta_{12} \sigma^2$$

if $h = 1$: $E[Y_t \cdot Y_{t+h}] = \mu^2 + \theta_1 \sigma^2$

if $h = 12$: $E[Y_t \cdot Y_{t+h}] = \mu^2 + \theta_{12} \sigma^2$

if $h = 11$: $E[Y_t \cdot Y_{t+h}] = \mu^2 + \theta_1 \theta_{12} \sigma^2$

if $h = -11$: $E[Y_t \cdot Y_{t+h}] = \mu^2 + \theta_1 \theta_{12} \sigma^2$

$$\Rightarrow \gamma_Y(h) = \begin{cases} \sigma^2 + \theta_1^2 \sigma^2 + \theta_{12}^2 \sigma^2 & \text{if } h = 0 \\ \theta_1 \sigma^2 & \text{if } h = \pm 1 \\ \theta_1 \theta_{12} \sigma^2 & \text{if } h = \pm 11 \\ \theta_{12} \sigma^2 & \text{if } h = \pm 12 \\ 0 & h \neq 0, \pm 1, \pm 11, \pm 12 \end{cases}$$

(b) Use the program ITSM to compute the sample mean and sample autocovariances $\hat{\gamma}(h)$, $0 \le h \le 20$, of $\{\nabla\nabla_{12} X_t\}$, where $\{X_t, t = 1, \ldots, 72\}$ is the accidental deaths series DEATHS.TSM of Example 1.1.3.

Q3_a3

Question 3.9(b)

```r
library(itsmr)
library(fpp2)
```

```
## Loading required package: ggplot2

## Loading required package: forecast

##
## Attaching package: 'forecast'

## The following object is masked from 'package:itsmr':
##
##     forecast

## Loading required package: fma

##
## Attaching package: 'fma'

## The following objects are masked from 'package:itsmr':
##
##     airpass, strikes

## Loading required package: expsmooth
```

```r
library(expsmooth)
library(forecast)
library(tibbletime)
```

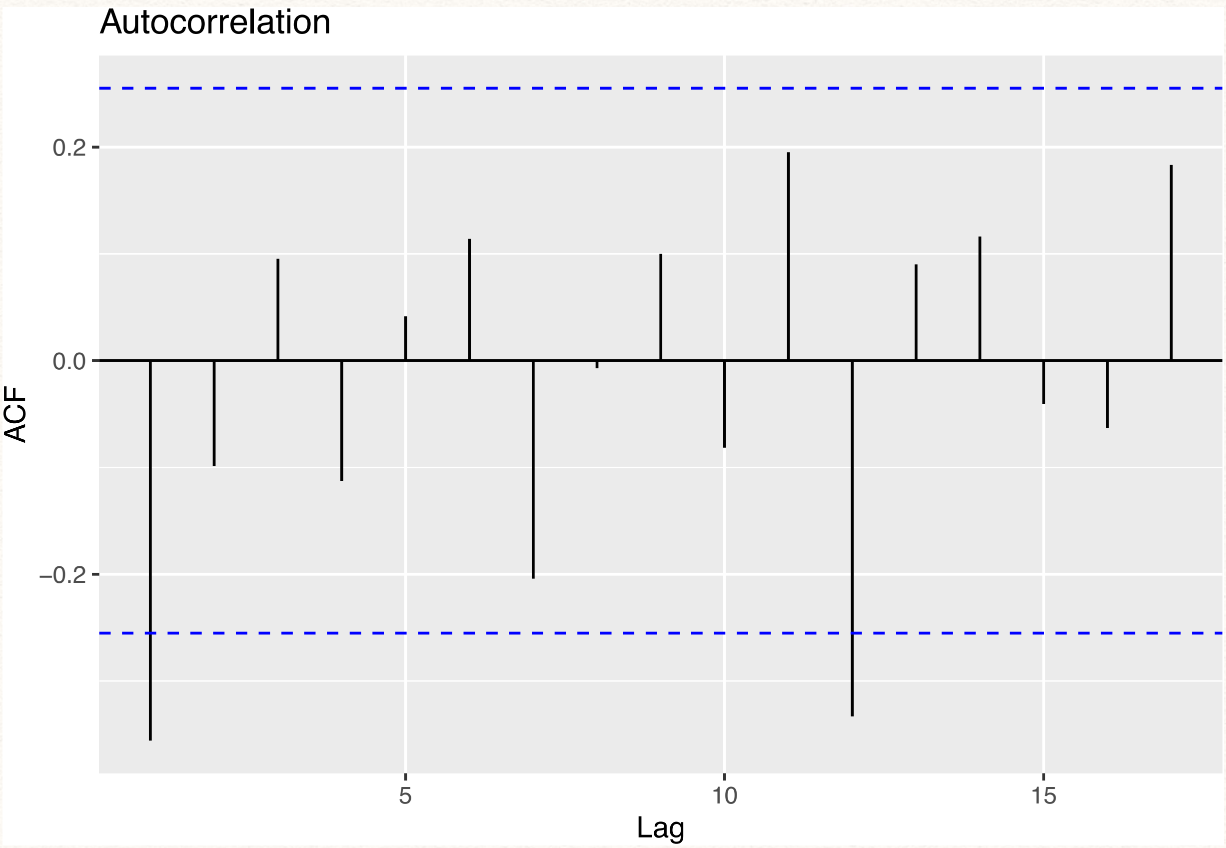
```
##
## Attaching package: 'tibbletime'

## The following object is masked from 'package:stats':
##
##     filter
```

```r
library(tidyverse)
```

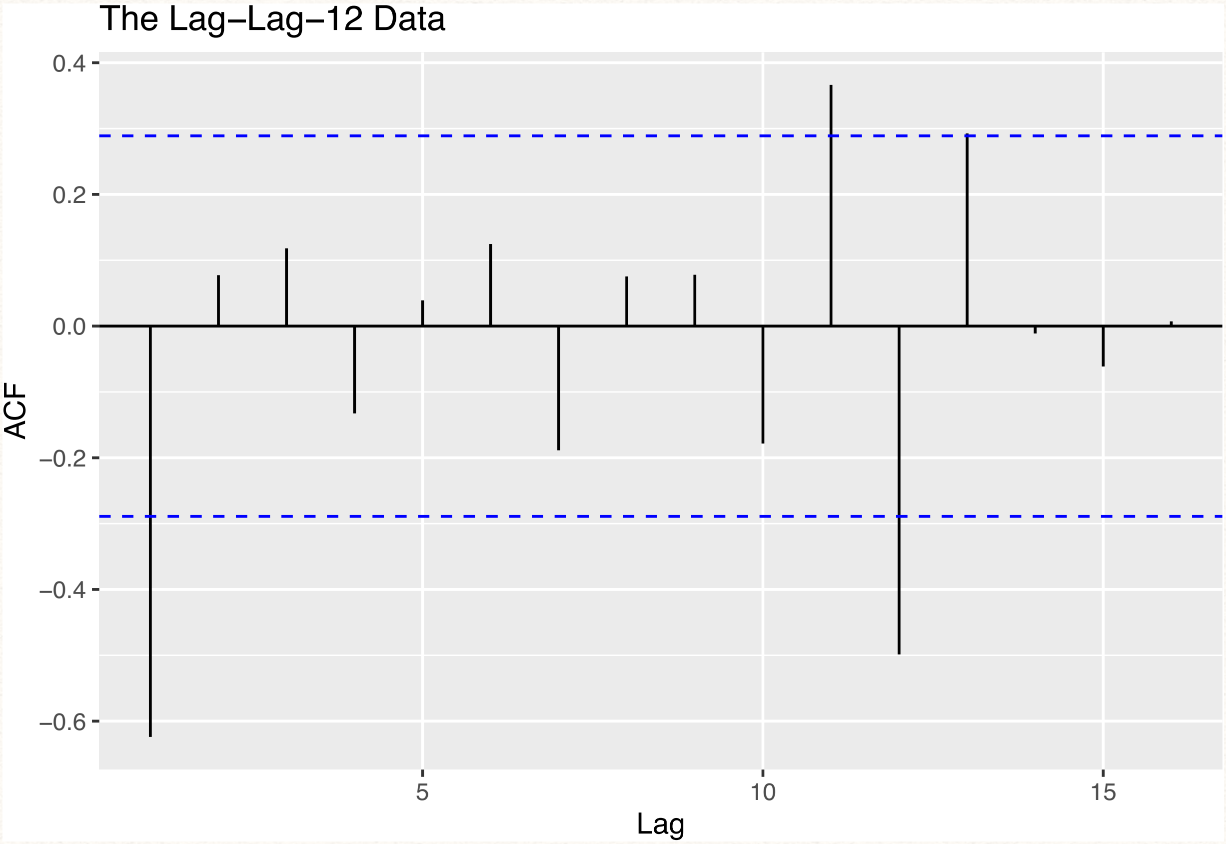
```
## -- Attaching packages ------------------------------------------------------------------------

## v tibble  2.1.3     v dplyr   0.8.4
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
## v purrr   0.3.3

## -- Conflicts ---------------------------------------------------------------------------------
## x dplyr::filter() masks tibbletime::filter(), stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(tsbox)
library(knitr)

#death<-ts(deaths, frequency=12, start=c(1973,1))
death<-diff(diff(deaths,12),1)
ACF<-ggAcf(death)
```

```r
#autoplot(death)+ggplot("Death Data")
ggAcf(death) + ggtitle("Autocorrelation")
```




```r
laglag<-diff(diff(death,lag=12))
ggAcf(laglag)+ggtitle("The Lag-Lag-12 Data")
```




```
#Now get the sample mean of data
mean(laglag)
```

```
## [1] 11.67391
```

```
#Now get the sample covariance
#acf<-data.frame(Lag=0:20,AutoCov=acf(laglag,lag.max=20,plot=FALSE, type="covariance")[["acf"]][, , 1])
#kable(acf)
data.frame("Lag"=ACF$data$lag,"autocovariance"=ACF$data$Freq)
```

```
##    Lag autocovariance
## 1    1   -0.355843707
## 2    2   -0.098720893
## 3    3    0.095530249
## 4    4   -0.112515462
## 5    5    0.041529223
## 6    6    0.114108535
## 7    7   -0.204130053
## 8    8   -0.007123311
## 9    9    0.100066891
## 10  10   -0.081448224
## 11  11    0.195205609
## 12  12   -0.333182813
## 13  13    0.090181278
## 14  14    0.116314298
## 15  15   -0.040607730
## 16  16   -0.063250388
## 17  17    0.183280640
```

(c) By equating $\hat{\gamma}(1)$, $\hat{\gamma}(11)$, and $\hat{\gamma}(12)$ from part (b) to $\gamma(1)$, $\gamma(11)$, and $\gamma(12)$, respectively, from part (a), find a model of the form defined in (a) to represent $\{\nabla\nabla_{12}X_t\}$.

Let $\hat{\gamma}(1) = -0.356 = \gamma(1) = \theta_1 \sigma^2$

$\hat{\gamma}(11) = 0.195 = \gamma(11) = \theta_1 \theta_{12} \sigma^2$

$\hat{\gamma}(12) = 0.333 = \gamma(12) = \theta_{12} \sigma^2$

$\Rightarrow$

$\theta_1 \theta_{12} \sigma^4 = -0.1185 \quad \Rightarrow \quad \hat{\sigma}^2 = -0.608$

$$\Rightarrow \begin{cases} \hat{\theta}_1 = 0.586 \\ \hat{\theta}_{12} = -0.548 \end{cases}$$

$\Rightarrow \nabla_{12} X_{12} = -0.586 (Z_{t-1} - Z_{t-13}) - 0.548 (Z_{t-12} - Z_{t-24})$

$\Rightarrow \nabla\nabla_{12} = -0.586 (Z_{t-1} - Z_{t-2} - Z_{t-13} + Z_{t-14}) - 0.548 (Z_{t-12} - Z_{t-13} - Z_{t-24} - Z_{t-25})$

Question 4

5.3 Consider the AR(2) process $\{X_t\}$ satisfying

$$X_t - \phi X_{t-1} - \phi^2 X_{t-2} = Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2).$$

a. For what values of ϕ is this a causal process?

Since $\Phi(B) = 1 - \phi B - \phi^2 B^2 = -\phi^2 B^2 - \phi B + 1 = 0$

$$B = \frac{-b \pm \sqrt{b^2 - 4ac}}{2b}$$

$$= \frac{\phi \pm \sqrt{\phi^2 4 4\phi^2}}{-2\phi} = \frac{\phi \pm \phi\sqrt{5}}{-2\phi} = \frac{1 \pm \sqrt{5}}{-2\phi}$$

for $\left|\frac{1+\sqrt{5}}{-2\phi}\right| > 1$ $\qquad$ $\left|\frac{1-\sqrt{5}}{-2\phi}\right| > 1$

$$1+\sqrt{5} > 2|\phi| \qquad \left|\frac{\sqrt{5}-1}{2\phi}\right| > 1$$

$$|\phi| < \frac{1+\sqrt{5}}{2} \qquad \sqrt{5}-1 > |2\phi|$$

$$|\phi| < \frac{\sqrt{5}-1}{2}$$

$$\Rightarrow \quad |\phi| < 0.618$$

b. The following sample moments were computed after observing $X_1, \ldots, X_{200}$:

$$\hat{\gamma}(0) = 6.06, \quad \hat{\rho}(1) = 0.687.$$

Find estimates of ϕ and σ^2 by solving the Yule–Walker equations. (If you find more than one solution, choose the one that is causal.)

Here we have AR (2) :

$$\because \gamma_X(h) = E[X_t \cdot X_{t+h}] = \sum_{r=1}^{2} \phi_r \gamma(|r-h|)$$

Also
$$\gamma_X(0) = \sum_{r=1}^{2} \phi_r \gamma(r) + \sigma^2 = \phi_1 \gamma(1) + \phi_2 \gamma(2) + \sigma^2$$

$$\gamma_X(1) = \sum_{r=1}^{2} \phi_r \gamma(|r-1|) = \phi_1 \gamma(0) + \phi_2 \gamma(1)$$

$$\gamma_X(2) = \sum_{r=1}^{2} \phi_r \gamma(|r-2|) = \phi_1 \gamma(1) + \phi_2 \gamma(0)$$

$$\Rightarrow \begin{cases} 6.06 = +\phi \cdot 4.163 + \phi^2 \gamma(2) + \sigma^2 \\ 4.163 = +\phi\, 6.06 + \phi^2 \cdot 4.163 \\ \gamma(2) = +\phi \cdot 4.163 + \phi^2 * 6.06 \end{cases} \qquad \Rightarrow \phi = \frac{1}{2 * 0.687} \pm \sqrt{\frac{1}{4 * 0.472^2} + 1}$$

$$\hat{\phi} = \begin{cases} 0.509 \\ -1.965 \end{cases} \quad \& \ X_t \text{ is a causal process.} \qquad \therefore \ \hat{\phi} = 0.509$$

then

$$\hat{\sigma}^2 = 6.06 - 0.509 * 4.163 - (0.509)^2 * (0.509 * 4.163 + 0.509^2 * 6.06)$$

$$= 2.985$$

Question 5

5.4 Two hundred observations of a time series, $X_1, \ldots, X_{200}$, gave the following sample statistics:

sample mean: $\bar{x}_{200} = 3.82$;

sample variance: $\hat{\gamma}(0) = 1.15$;

sample ACF: $\hat{\rho}(1) = 0.427$;

$\hat{\rho}(2) = 0.475$;

$\hat{\rho}(3) = 0.169$.

a. Based on these sample statistics, is it reasonable to suppose that $\{X_t - \mu\}$ is white noise?

Assume $\{X_t - \mu\} \sim WN(0, \sigma^2)$ then $\{X_t\}$ should have $E[X_t] = \mu$, $Var(X_t) = \sigma^2$ and $Cov(X_t, X_{t+h}) = 0$ for $h \neq 0$. However, here $\hat{p}(1) \neq \hat{p}(2) \neq \hat{p}(3) \neq 0$. Let's see the 95% confidence interval of them: that is $\hat{p}(h) \pm \frac{1.96}{\sqrt{n}}$:

$$\hat{p}(1) \pm \frac{1.96}{\sqrt{200}} = (0.427 - 0.139,\ 0.427 + 0.139)$$

$$\hat{p}(2) \pm \frac{1.96}{\sqrt{200}} = (0.475 - 0.139,\ 0.475 + 0.139)$$

$$\hat{p}(3) \pm \frac{1.96}{\sqrt{200}} = (0.169 - 0.139,\ 0.169 + 0.139)$$

It is obvious that the 95% confidence interval does not contain $\hat{p}(h) = 0$. We should reject the null hypothesis that $\{X_t - \mu\}$ is white Noise process.

b. Assuming that $\{X_t - \mu\}$ can be modeled as the AR(2) process

$$(X_t - \mu) - \phi_1(X_{t-1} - \mu) - \phi_2(X_{t-2} - \mu) = Z_t,$$

where $\{Z_t\} \sim \text{IID}(0, \sigma^2)$, find estimates of μ, ϕ_1, ϕ_2, and σ^2.

Since $\overline{X_{200}} = 3.82$, we can estimate $\mu = 3.82$ then since

$$\hat{\gamma}_m = \hat{\Gamma}_m \underline{\phi}$$

$$\hat{\Gamma}_m = \begin{pmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) \end{pmatrix} = \begin{pmatrix} \hat{\gamma}(0) & \hat{p}(1)\hat{\gamma}(0) \\ \hat{\gamma}(0) & \hat{\gamma}(0)\hat{p}(1) \end{pmatrix} = \begin{pmatrix} 1.15 & 1.15 * 0.427 \\ 1.15 * 0.427 & 1.15 \end{pmatrix}$$

$$\hat{\gamma}_m = (\hat{\gamma}(1), \hat{\gamma}(2)) = (1.15 * 0.427 \quad , \quad 1.15 * 0.475)$$

$\Rightarrow$

$$\hat{\phi} = \hat{\Gamma}_m^{-1} \hat{\gamma}_m$$

$$= \frac{1}{1.15^2 - 1.15^2 * 0.427^2} \begin{pmatrix} 1.15 & -1.15 * 0.427 \\ -1.15 * 0.427 & 1.15 \end{pmatrix} \begin{pmatrix} 1.15 * 0.427 \\ 1.15 * 0.475 \end{pmatrix}$$

$$= \frac{1}{1.15^2 - 1.15^2 * 0.427^2} \begin{pmatrix} 0.296 \\ 0.387 \end{pmatrix}$$

$$= \frac{1}{1.08} \begin{pmatrix} 0.296 \\ 0.387 \end{pmatrix} \Rightarrow \begin{cases} \hat{\phi}_1 = 0.274 \\ \hat{\phi}_2 = 0.358 \end{cases}$$

$$\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}_1 \hat{\gamma}(1) - \hat{\phi}_2 \hat{\gamma}(2)$$

$$= 1.15 - 0.274 * 0.427 * 1.15 - 0.358 * 0.475 * 1.15$$

$$= 0.8199$$

c. Would you conclude that $\mu = 0$?

Find the 95% confidence interval of X_{200} :

Calculate $\hat{v} = \sum_{h=-\infty}^{\infty} \hat{\gamma}(h) = \hat{\gamma}(-3) + \hat{\gamma}(-2) + \hat{\gamma}(-1) + \hat{\gamma}(0) + \hat{\gamma}(1) + \hat{\gamma}(2) + \hat{\gamma}(3)$

$$= 1.15 + (0.427 * 2 + 0.475 * 2 + 0.169 * 2) * 1.15$$

$$= 3.613$$

$$\Rightarrow \bar{x}_n \pm 1.96 * \sqrt{\frac{v}{n}}$$

$$= 3.82 \pm 1.96 * \sqrt{\frac{3.613}{200}} = 3.82 \pm 0.263$$

$\Rightarrow$ $\mu = 0$ is not in 95% CI

$\Rightarrow$ $\mu \neq 0$

d. Construct 95 % confidence intervals for ϕ_1 and ϕ_2.

Since $\hat{\phi}_p \sim N(\phi_p, \frac{1}{n}\sigma^2 \Gamma_p^{-1})$

$\Rightarrow \hat{\phi}_2 \sim N(\phi_2, \frac{\sigma^2}{n}\Gamma_2^{-1})$

Since $\hat{\sigma}^2 = 0.8199$ and $\hat{\Gamma}_2^{-1} = \frac{1}{1.15^2 - 1.15^2 * 0.427^2}\begin{pmatrix} 1.15 & -1.15 * 0.427 \\ -1.15 * 0.427 & 1.15 \end{pmatrix}$

$$\Rightarrow \frac{\sigma^2}{n}\Gamma_2^{-1} = \frac{1}{1.08} * \frac{0.8199}{200}\begin{pmatrix} 1.15 & -1.15 * 0.427 \\ -1.15 * 0.427 & 1.15 \end{pmatrix}$$

$$= \begin{pmatrix} 0.005 & -0.0021 \\ -0.0021 & 0.005 \end{pmatrix}$$

$\Rightarrow \hat{\phi}_1 \pm 1.96\sqrt{0.005} = 0.274 \pm 0.139 = [0.135, 0.413]$

$\hat{\phi}_2 \pm 1.96\sqrt{0.005} = 0.358 \pm 0.139 = [0.219, 0.454]$

e. Assuming that the data were generated from an AR(2) model, derive estimates of the PACF for all lags $h \geq 1$.

Since $\alpha(0) = 1$ and $\alpha(h) = \phi_{hh} = \hat{\phi}_h$

Then we can conclude that $\alpha(1) = \hat{\phi}_1 = 0.427$

$\alpha(2) = \hat{\phi}_2 = 0.358$

$\alpha(h) = \hat{\phi}_h = 0$ for $h \geq 3$

Question.6

5.11 Given two observations x_1 and x_2 from the causal AR(1) process satisfying

$$X_t = \phi X_{t-1} + Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2),$$

and assuming that $|x_1| \neq |x_2|$, find the maximum likelihood estimates of ϕ and σ^2.

Since $L(\phi, \theta, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^n r_0 \cdots r_{n-1}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}\frac{(x_j - \hat{x}_j)^2}{r_{j-1}}\right)$

$$= \frac{1}{(2\pi\sigma^2)^2 r_0 r_1} \exp\left(-\frac{1}{2\sigma^2}\left(\frac{(x_1-\hat{x}_1)^2}{r_0} + \frac{(x_2-\hat{x}_2)^2}{r_1}\right)\right)$$

and $\ell(\phi, \theta, \sigma^2) = -\log 2\pi\sigma^2 (\log r_0 + \log r_1) - \frac{1}{2\sigma^2}\left(\frac{(x_1-\hat{x}_1)^2}{r_0} + \frac{(x_2-\hat{x}_2)^2}{r_1}\right)$

since $E[(x_n - \hat{x}_n)^2] = r_n \sigma^2 = V_n$

$$\Rightarrow \begin{cases} E[(x_1-\hat{x}_1)^2] = r_0\sigma^2 = V_0 \\ E[(x_2-\hat{x}_2)^2] = r_1\sigma^2 = V_1 \end{cases}$$

$x_1 - \hat{x}_1 = x_1 \sim N(0, V_0)$ and $x_2 - \hat{x}_2 = x_2 - \phi x_1 \sim N(0, V_1)$

$\Rightarrow \quad \sigma^2 r_0 = E[x_1^2] = \gamma(0) \quad \therefore \quad r_0 = \frac{\gamma(0)}{\sigma^2}$

$\sigma^2 r_1 = E[x_2^2 + \phi^2 x_1 - 2\phi x_1 x_2] = \gamma(0) + \phi^2\gamma(0) - 2\phi\gamma(1)$

$r_1 = (\gamma(0) + \phi^2\gamma(0) - 2\phi\gamma(1))/\sigma^2$

$\Rightarrow$

$$\ell(\phi, \theta, \sigma^2) = -\log(2\pi\sigma^2)\log\left(\frac{\gamma(0)}{\sigma^2}\ \frac{\gamma(0)+\phi^2\gamma(0)-2\phi\gamma(1)}{\sigma^2}\right) - \frac{1}{2\sigma^2}\left(\frac{x_1^2}{r_0} + \frac{(x_2-\phi x_1)^2}{r_1}\right)$$

$$= -\log(2\pi\sigma^2)\log\left(\frac{\gamma(0)}{\sigma^4}[\gamma(0)+\phi^2\gamma(0)-2\phi\gamma(1)]\right) - \frac{1}{2\sigma^2}\left(\frac{x_1^2}{\gamma(0)/\sigma^2} + \frac{(x_2-\phi x_1)^2}{(\gamma(0)+\phi^2\gamma(0)-2\phi\gamma(1))/\sigma^2}\right)$$

$$= -\log(2\pi\sigma^2) + \frac{1}{2}\log(1-\phi^2) - \frac{1}{2\sigma^2}(x_1^2(1-\phi^2) + (x_2-\phi x_1)^2)$$

then

$$\begin{cases} \frac{\partial \ell(\phi,\theta,\sigma^2)}{\partial\sigma^2} = -\frac{1}{2\pi\sigma^2} * 4\pi\sigma - \frac{1}{2}*(-2)*\frac{1}{\sigma^3}(x_1^2(1-\phi^2) + (x_2-\phi x_1)^2) \\ \qquad = \frac{1}{\sigma^2} + \frac{1}{2\sigma^4}(x_1^2(1-\phi^2) + (x_2-\phi x_1)^2) = 0 \\ \frac{\partial \ell(\phi,\theta,\sigma^2)}{\partial\sigma^2} = \frac{-\phi}{1-\phi^2} + \frac{x_1x_2}{\sigma^2} = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \hat{\phi} = \frac{2x_1x_2}{x_1^2+x_2^2} \\ \hat{\sigma}^2 = \frac{\frac{1}{2}(x_1^2-x_2^2)^2}{(x_2^2+x_1^2)} \end{cases}$$

Question 7

5.12 Derive a cubic equation for the maximum likelihood estimate of the coefficient ϕ of a causal AR(1) process based on the observations $X_1, \ldots, X_n$.

For AR(1) : $X_t - \phi X_{t-1} = Z_t$ & X_t is Gaussian Time series.

with mean = 0 and $Z_t \sim N(0, \sigma^2)$.

Then the likelihood would be

$$L(\phi, \sigma^2) = \prod_{j=1}^{n} f_{X_j - \hat{X}_j}$$

$$= f(X_1) f(X_2 | X_1) \cdots f(X_n | X_{n-1} \cdots X_{n-1})$$

$$= f(X_1) \prod_{j=2}^{n} f_{X_i | X_{i-1}} (X_i | X_{i-1})$$

Now let's calculate $f(X_j | X_{j-1})$ then

$$E[X_j | X_{j-1}] = E[\phi X_{t-1} + Z_t | X_{j-1}] = \phi E[X_{t-1}] = \phi X_{t-1}$$

$$Var[X_j | X_{j-1}] = E[(X_j - \phi X_{j-1})^2 | X_{j-1}] = E[X_j^2 + \phi^2 X_{j-1}^2 - 2X_j X_{j-1} | X_{j-1}]$$

$$= \phi^2 X_{j-1}^2 - 2X_{j-1}^2 * \phi + E[(\phi X_{j-1} + Z_t)^2 | X_{j-1}]$$

$$= \sigma^2$$

Since $X_1 - \hat{X}_1 = X_1 \sim N(0, V_0)$ by innovation, then

$$E[X_1 - \hat{X}_1] = 0 \qquad V_0 = E[(X_1 - \hat{X}_1)(X_1 - \hat{X}_1)] = E[X_1^2] = \gamma(0)$$

Then $\gamma_X(0) = \frac{\sigma^2}{1-\phi^2} \Rightarrow X_1 \sim N(0, \frac{\sigma^2}{1-\phi^2})$

For $N(\mu, \sigma^2) \Rightarrow f = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{\frac{-(x-m)^2}{-2\sigma^2}\}$.

$\Rightarrow$ The likelihood $L(\phi, \sigma^2)$

$$= \frac{1}{\sqrt{2\pi\sigma^2/1-\phi^2}} \exp\left\{\frac{-X_1^2}{-2\frac{\sigma^2}{1-\phi^2}}\right\} \prod_{j=2}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{(X_j - \phi X_{j-1})^2}{-2\sigma^2}\right\}$$

$$= \frac{\sqrt{1-\phi^2}}{(\sqrt{2\pi\sigma^2})^n} \exp\left\{-\frac{X_1^2(1-\phi^2)}{2\sigma^2} + \sum_{j=2}^{n} \frac{(X_j - \phi X_{j-1})^2}{-2\sigma^2}\right\}$$

$$\Rightarrow \ell(\phi, \sigma^2) = \log L(\phi, \sigma^2)$$
$$= -\frac{1}{2\sigma^2}\left\{ -x_1^2(1-\phi^2) + \sum_{j=2}^{n}(x_j - \phi x_{j-1})^2 \right\} + \log\sqrt{1-\phi^2} - n\log\sqrt{2\pi\sigma^2}$$

By taking $\frac{\partial \ell(\phi, \sigma^2)}{\partial \phi}$:

$$\frac{\partial}{\partial\phi}\left\{ \frac{x_1^2(1-\phi^2)}{2\sigma^2} - \frac{1}{2\sigma^2}\sum_{j=2}^{n}(x_j - \phi x_{j-1})^2 \right\} + \log\sqrt{1-\phi^2} - \frac{n}{2}\log 2\pi\sigma^2 = 0$$

$$\phi x_1^2 + \sum_{i=2}^{n} x_i x_{i-1} - \phi\sum_{j=2}^{n} x_{j-1}^2 = \frac{\sigma^2\phi}{1-\phi^2}$$

$$\Rightarrow \sum_{i=2}^{n} x_i x_{i-1} = \frac{1}{1-\phi^2}\left(\sigma^2\phi + \phi(1-\phi^2)\sum_{i=3}^{n} x_{i-1}^2\right)$$

$$(1-\phi^2)\sum_{i=2}^{n} x_i x_{i-1} = \sigma^2\phi + \phi\sum_{i=3}^{n} x_{i-1}^2 - \phi^3\sum_{i=3}^{n} x_{i-1}^2$$

This gives us a cubic equation.

finalquestion

```r
library(itsmr)
library(fpp2)
```

```
## Loading required package: ggplot2

## Loading required package: forecast

##
## Attaching package: 'forecast'

## The following object is masked from 'package:itsmr':
##
##     forecast

## Loading required package: fma

##
## Attaching package: 'fma'

## The following objects are masked from 'package:itsmr':
##
##     airpass, strikes

## Loading required package: expsmooth
```

```r
library(expsmooth)
library(forecast)
library(tibbletime)
```

```
##
## Attaching package: 'tibbletime'

## The following object is masked from 'package:stats':
##
##     filter
```

```r
library(tidyverse)
```

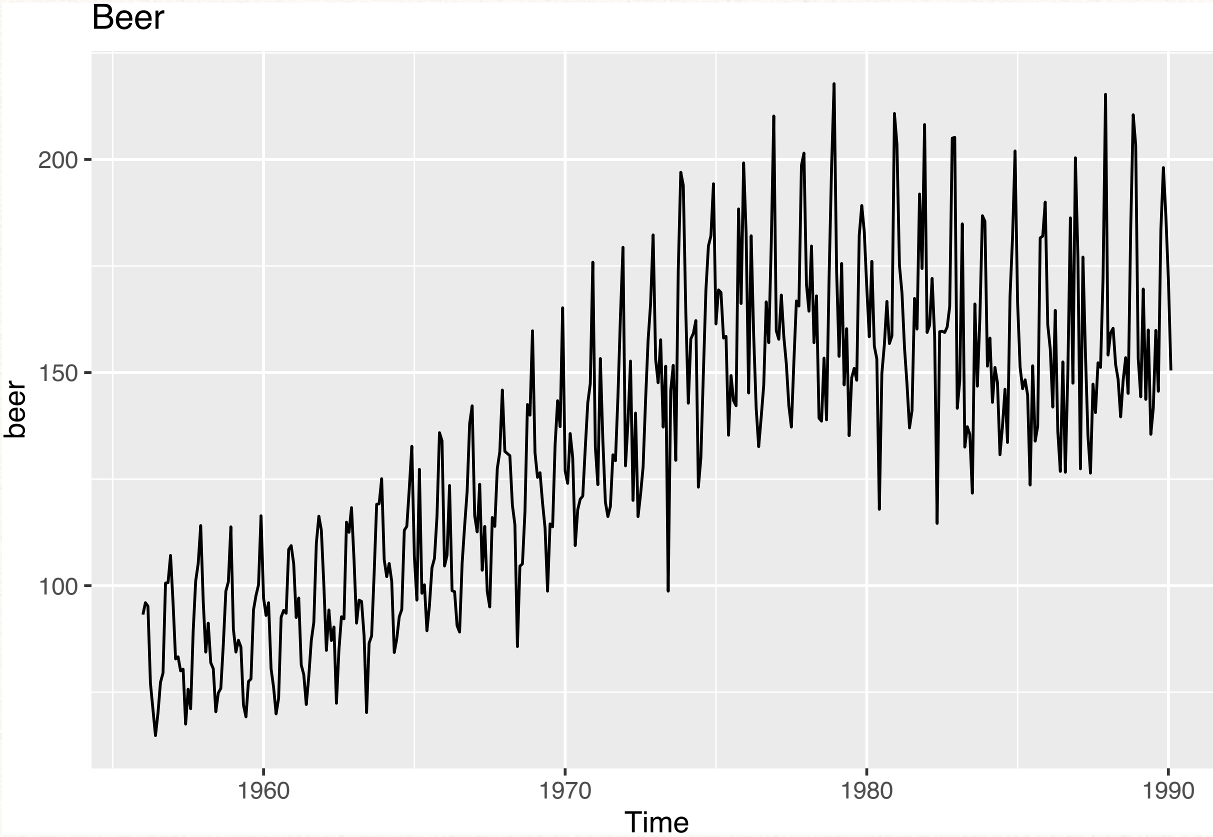
```
## -- Attaching packages ------------------------------------------------------------------------------------

## v tibble  2.1.3     v dplyr   0.8.4
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
## v purrr   0.3.3

## -- Conflicts ---------------------------------------------------------------------------------------------
## x dplyr::filter() masks tibbletime::filter(), stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(tsbox)
library(knitr)
```

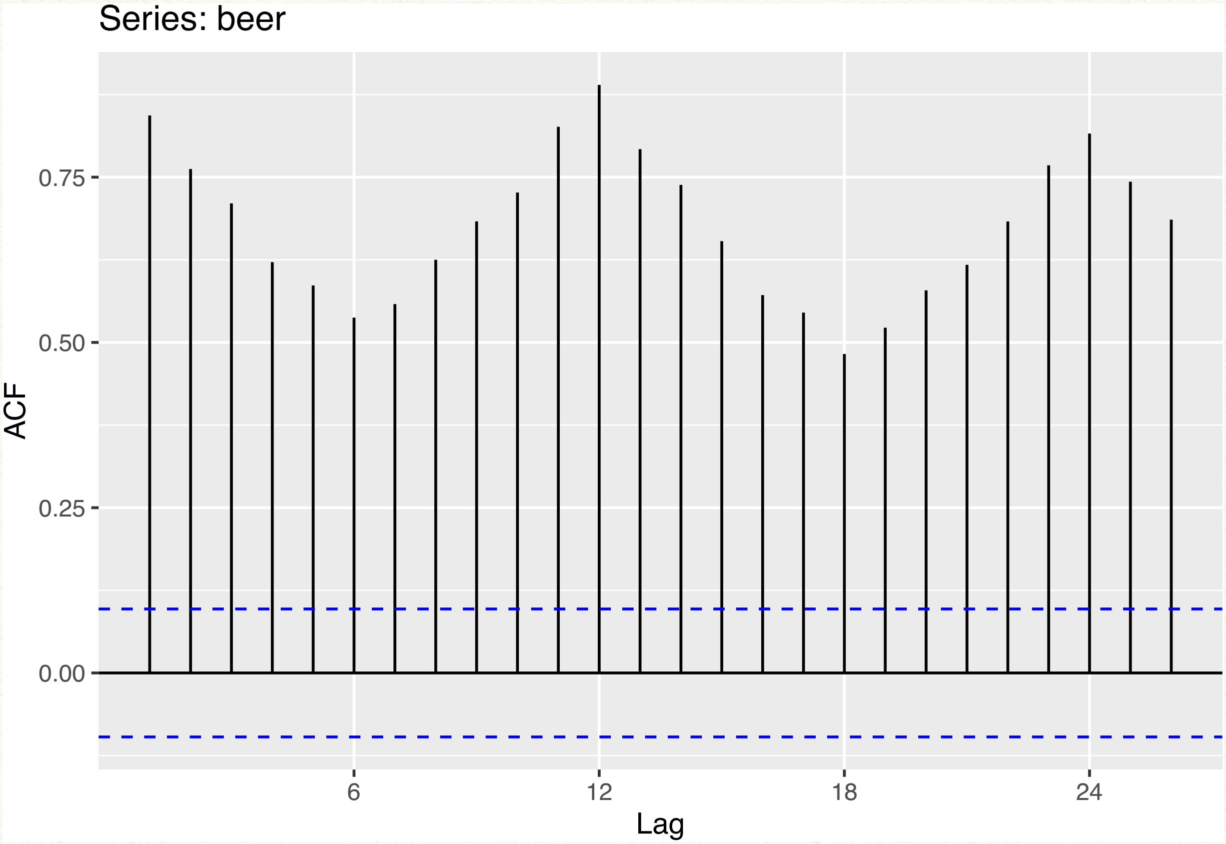
##(a)

```r
beer_dat=dget("beer.Rput")
# remove the last 12 values
beer<-head(beer_dat,-12)
```

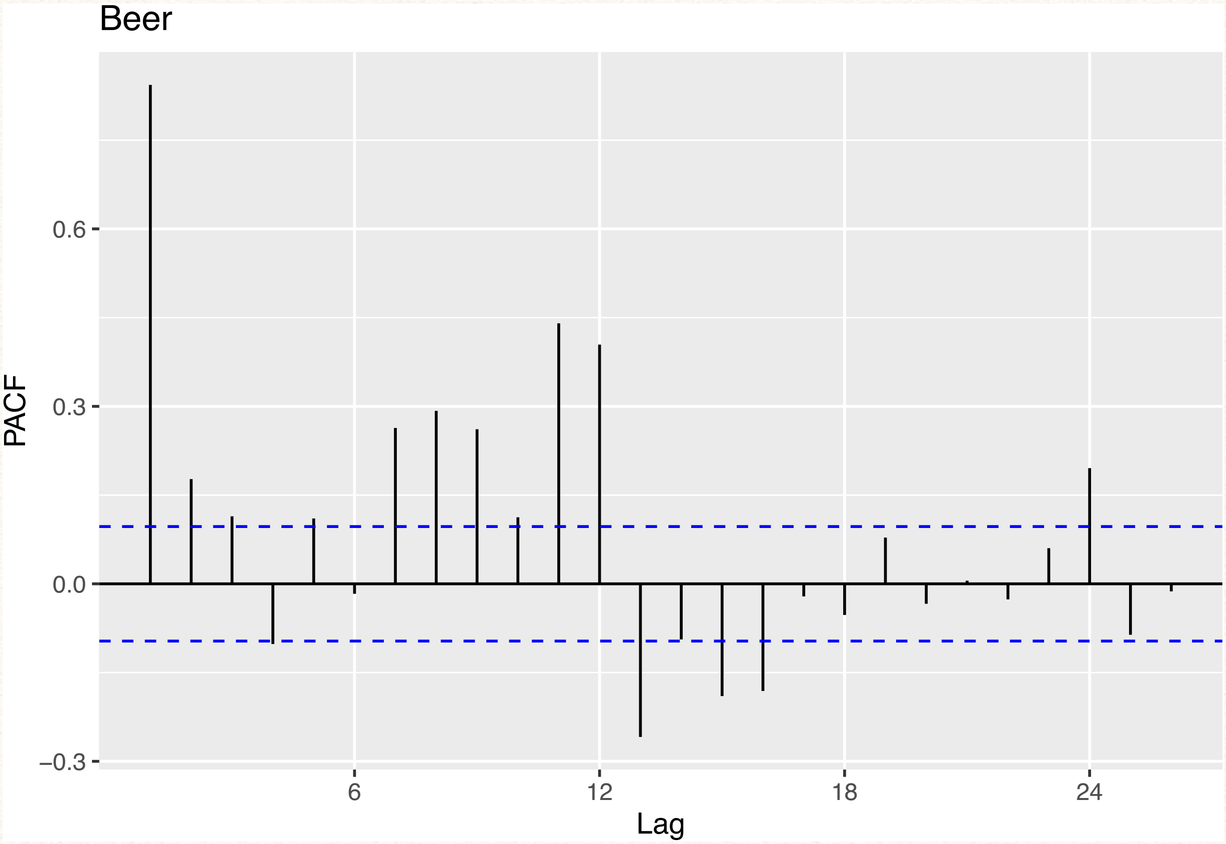
```r
# plot the data
autoplot(beer)+ggtitle("Beer")
```




```r
ggAcf(beer)
```



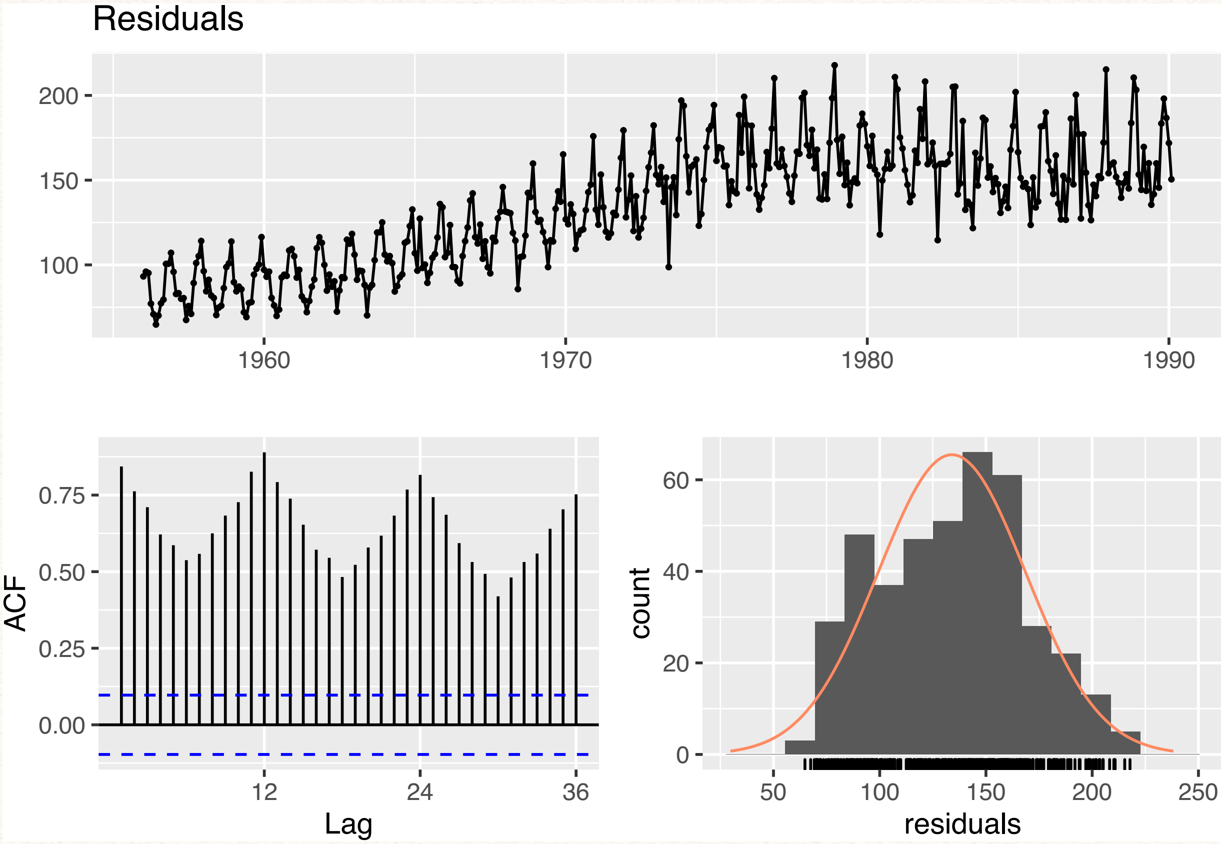

```r
ggPacf(beer)+ggtitle("Beer")
```



It shows that the data is not stationay but contain the seasonal component and noise.Also the above ACF plot shows the seasonal componet and the data are correlated. It is obvious that this dataset is not a white noise process.

```
# check the residual by deseasonalizing
#sales_ma5<-ma(beer,order=5)
#tslm<-tslm(beer~trend+I(trend^2)+season)
#stl<-stl(beer,s.window=12)
#checkresiduals(stl$time.series[,'remainder'])
checkresiduals(beer)
```

```
## Warning in modeldf.default(object): Could not find appropriate degrees of
## freedom for this model.
```



It shows that the data is not stationay but contain the seasonal component and noise.Also the above ACF plot and the residual plot shows the seasonal componet and the data are correlated. It is obvious that this dataset is not a white noise process.

```
# find the arima model
beer_ari<-auto.arima(beer,stepwise=FALSE, seasonal=FALSE, ic='aic',trace=TRUE)
```

```
##
##  Fitting models using approximations to speed things up...
##
##  ARIMA(0,1,0)                                : 3579.226
##  ARIMA(0,1,0)            with drift          : 3581.204
##  ARIMA(0,1,1)                                : 3547.203
##  ARIMA(0,1,1)            with drift          : 3549.134
##  ARIMA(0,1,2)                                : 3493.171
##  ARIMA(0,1,2)            with drift          : 3489.323
##  ARIMA(0,1,3)                                : 3490.168
##  ARIMA(0,1,3)            with drift          : 3486.407
##  ARIMA(0,1,4)                                : 3468.429
##  ARIMA(0,1,4)            with drift          : 3464.997
##  ARIMA(0,1,5)                                : 3469.946
##  ARIMA(0,1,5)            with drift          : 3466.611
##  ARIMA(1,1,0)                                : 3556.574
##  ARIMA(1,1,0)            with drift          : 3558.536
##  ARIMA(1,1,1)                                : 3483.419
##  ARIMA(1,1,1)            with drift          : 3480.55
##  ARIMA(1,1,2)                                : 3485.297
##  ARIMA(1,1,2)            with drift          : 3482.359
```

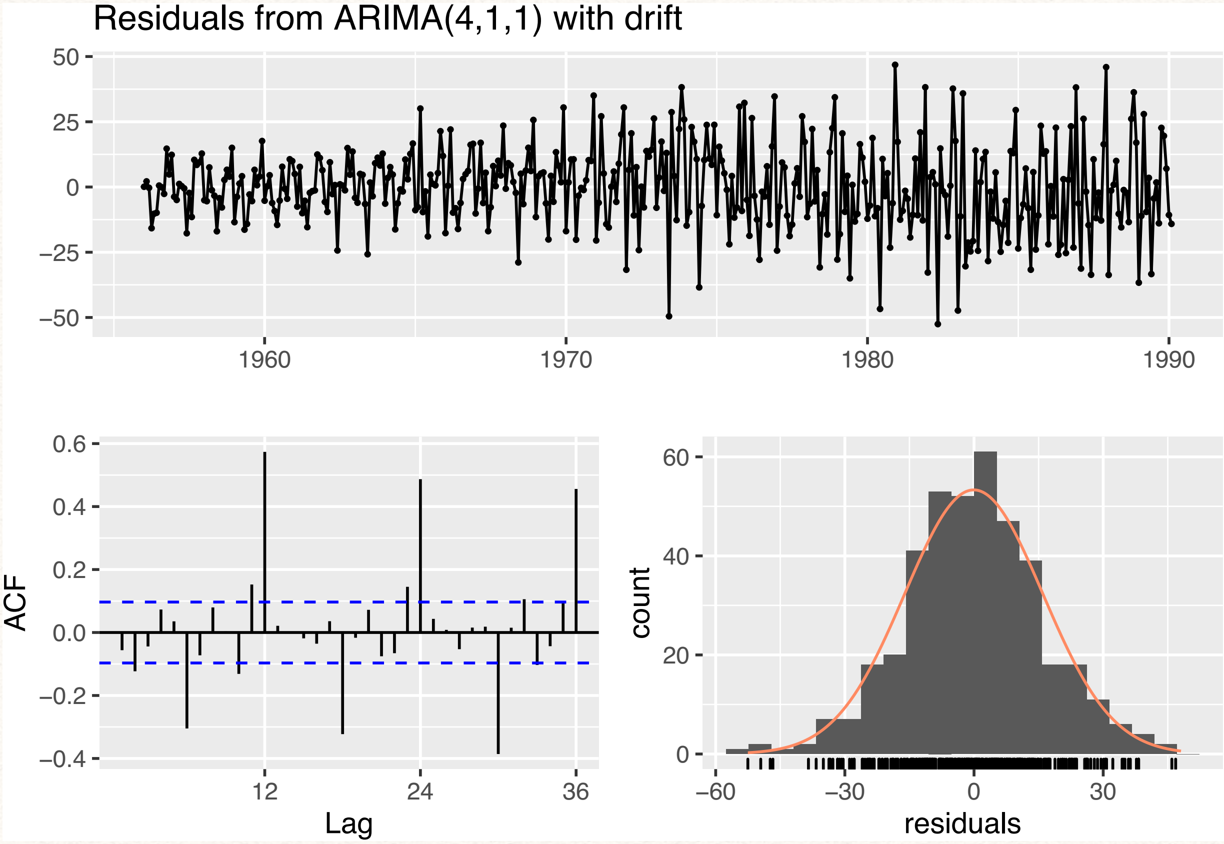
```
##  ARIMA(1,1,3)                                : 3486.94
##  ARIMA(1,1,3)            with drift          : 3483.973
##  ARIMA(1,1,4)                                : 3471.249
##  ARIMA(1,1,4)            with drift          : 3468.278
##  ARIMA(2,1,0)                                : 3547.96
##  ARIMA(2,1,0)            with drift          : 3549.898
##  ARIMA(2,1,1)                                : 3540.196
##  ARIMA(2,1,1)            with drift          : 3542.157
##  ARIMA(2,1,2)                                : 3481.216
##  ARIMA(2,1,2)            with drift          : 3477.923
##  ARIMA(2,1,3)                                : 3482.351
##  ARIMA(2,1,3)            with drift          : 3479.135
##  ARIMA(3,1,0)                                : 3549.09
##  ARIMA(3,1,0)            with drift          : 3551.011
##  ARIMA(3,1,1)                                : 3543.732
##  ARIMA(3,1,1)            with drift          : 3545.695
##  ARIMA(3,1,2)                                : 3481.842
##  ARIMA(3,1,2)            with drift          : 3476.954
##  ARIMA(4,1,0)                                : 3541.237
##  ARIMA(4,1,0)            with drift          : 3543.099
##  ARIMA(4,1,1)                                : 3449.546
##  ARIMA(4,1,1)            with drift          : 3443.193
##  ARIMA(5,1,0)                                : 3543.721
##  ARIMA(5,1,0)            with drift          : 3545.553
##
##  Now re-fitting the best model(s) without approximations...
##
##
##
##
##  Best model: ARIMA(4,1,1)            with drift
```

Now analyze the model:

```r
summary(beer_ari)
```

```
## Series: beer
## ARIMA(4,1,1) with drift
##
## Coefficients:
##          ar1     ar2     ar3      ar4      ma1   drift
##       0.4462  0.0028  0.0763  -0.3170  -0.9395  0.1977
## s.e.  0.0477  0.0518  0.0517   0.0475   0.0127  0.0632
##
## sigma^2 estimated as 261.2:  log likelihood=-1716.54
## AIC=3447.08   AICc=3447.36   BIC=3475.18
##
## Training set error measures:
##                      ME     RMSE      MAE       MPE     MAPE     MASE
## Training set -0.0588431 16.02236 12.37635 -1.226062 9.171119 1.306263
##                     ACF1
## Training set -0.05614064
```

```r
checkresiduals(beer_ari)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(4,1,1) with drift
## Q* = 377.63, df = 18, p-value < 2.2e-16
##
## Model df: 6.   Total lags used: 24
```

```r
#test(checkresiduals(beer_ari))
```

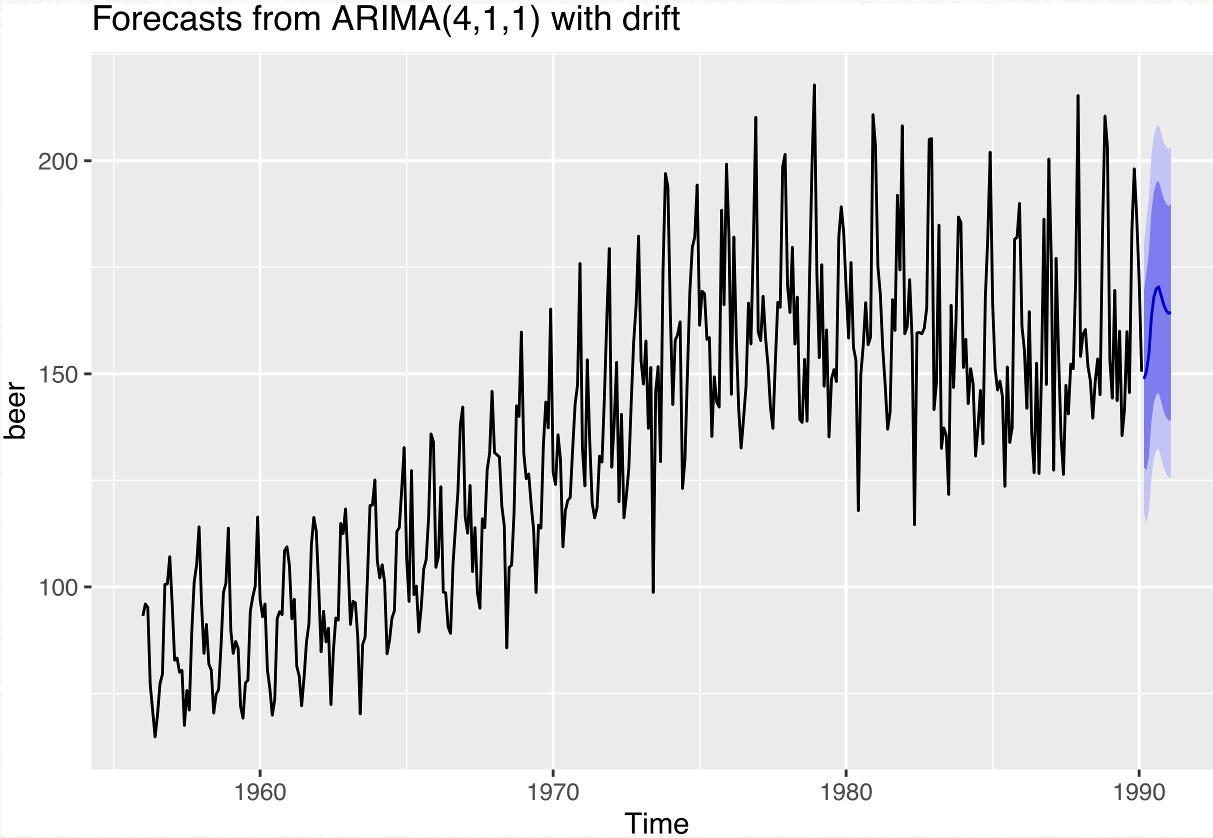
Construct the 95% confidence interval. Showing the bounds for component θ and ϕ

```r
confint(beer_ari)
```

```
##              2.5 %     97.5 %
## ar1     0.35276931  0.5396158
## ar2    -0.09861421  0.1043077
## ar3    -0.02492603  0.1775627
## ar4    -0.41008956 -0.2239423
## ma1    -0.96429447 -0.9146189
## drift   0.07376610  0.3216664
```

Now forecast on the 12 removed data

```r
beer_forecast<-forecast(beer_ari,h=12)
autoplot(beer_forecast)
```



Here's the result of prediction presented in table:

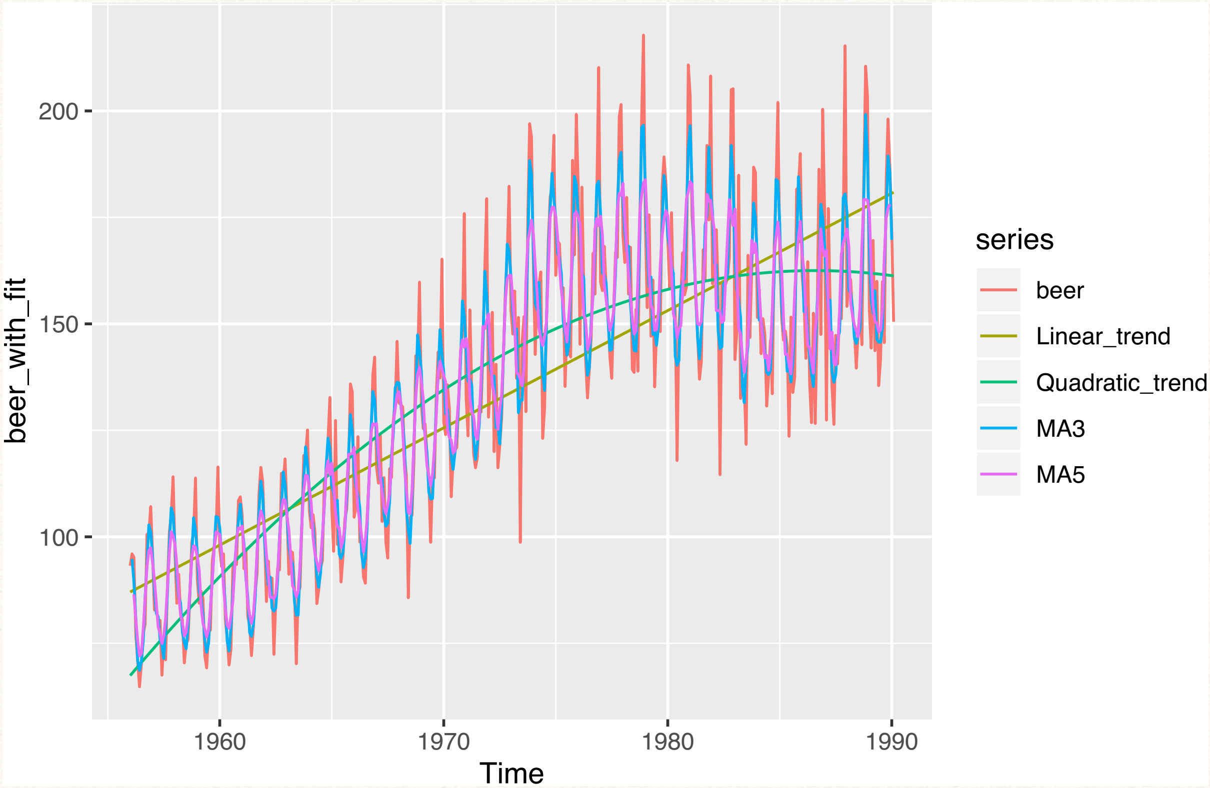
```
bt<-log(tail(beer_dat,12))
error<-bt-beer_forecast$mean
data<-data.frame(error, beer_forecast$lower, beer_forecast$upper,beer_forecast$mean)
colnames(data)<-c('error', '2.5%','97.5%','predict')
data
```

```
##         error     2.5%    97.5%  predict       NA       NA
## 1  -143.7117 128.0945 117.1307 169.5165 180.4803 148.8055
## 2  -145.5951 127.4111 115.1201 173.8479 186.1389 150.6295
## 3  -149.6245 130.6737 117.9794 178.6337 191.3279 154.6537
## 4  -158.3575 138.6542 125.6256 187.8773 200.9059 163.2658
## 5  -162.9536 143.2583 130.1858 192.6473 205.7197 167.9528
## 6  -164.9548 145.1331 131.9937 194.7749 207.9143 169.9540
## 7  -165.5043 145.5410 132.3823 195.2557 208.4145 170.3984
## 8  -163.1195 143.4968 130.3211 193.2759 206.4516 168.3864
## 9  -160.9738 141.3979 128.2084 191.2287 204.4182 166.3133
## 10 -159.6554 139.9032 126.6503 189.9740 203.2269 164.9386
## 11 -159.0816 139.0031 125.6744 189.3598 202.6884 164.1814
## 12 -159.4786 139.1259 125.7065 189.8257 203.2451 164.4758
```

##(c) First deseasonalizing

```
beer_linear<- tslm(beer~trend)
beer_quadratic<- tslm(beer~trend+I(trend^2))
beer_ma3<-ma(beer,order=3)
beer_ma5<-ma(beer,order=5)
#plot the fitted grapth
```

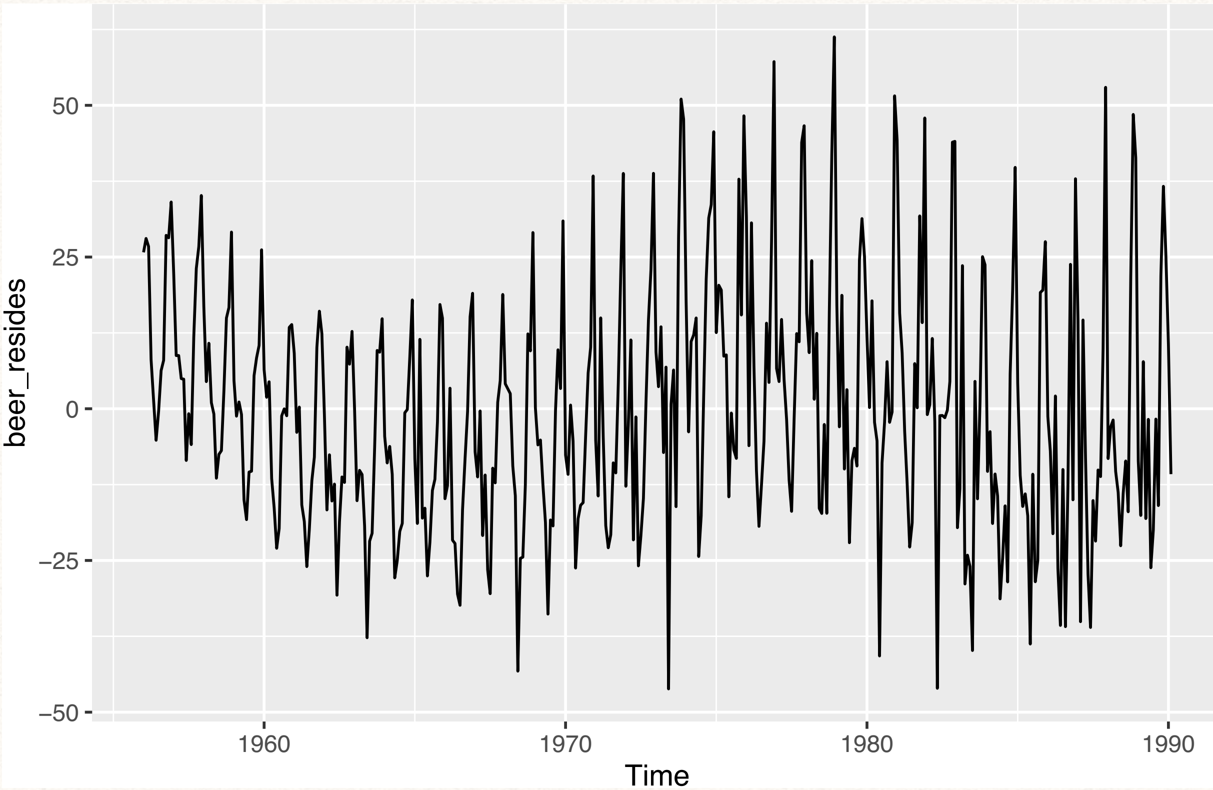
```r
beer_with_fit<-cbind(beer,Linear_trend=fitted(beer_linear),Quadratic_trend=fitted(beer_quadratic),MA3=be
autoplot(beer_with_fit)
```



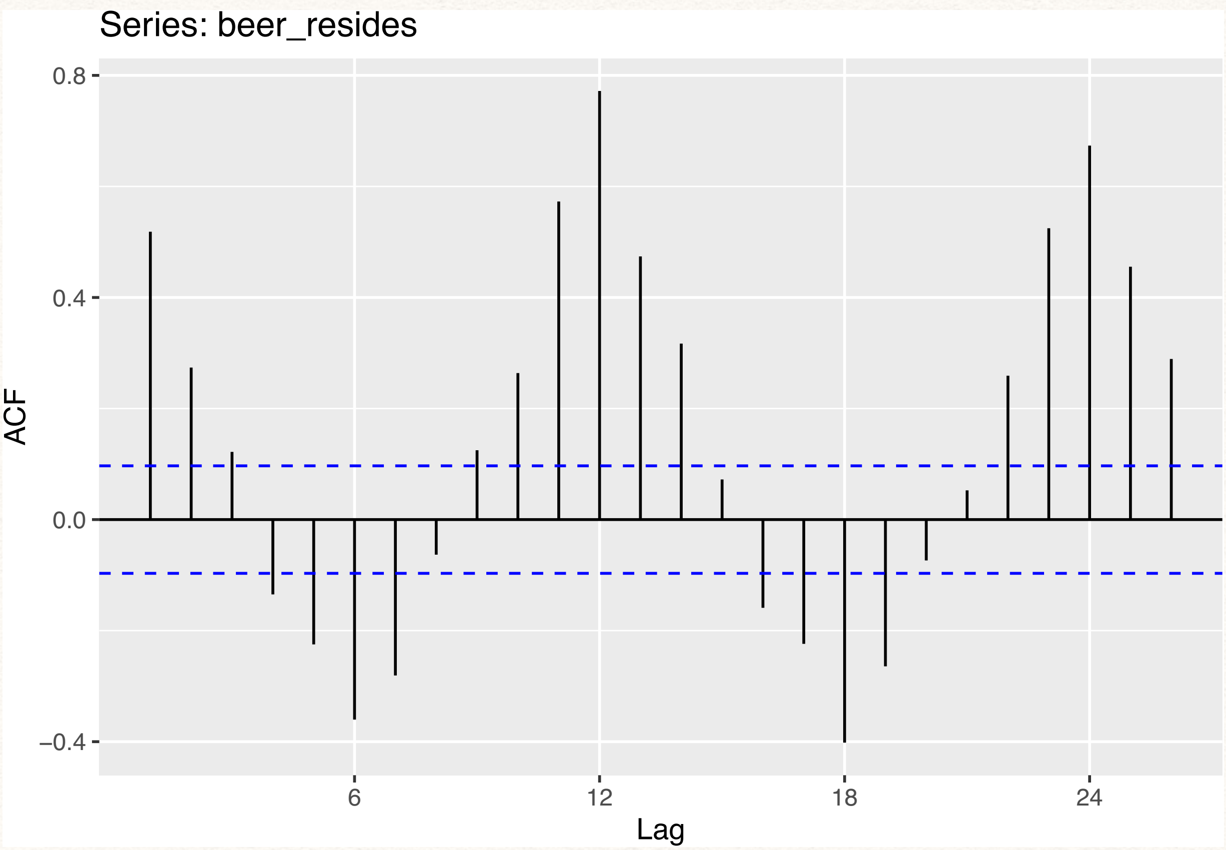
The residual plots by removing these the quadratic trends respectively from the dataset :

```r
beer_resides<-cbind(
                    res_quad=beer-fitted(beer_quadratic))

autoplot(beer_resides,facet=TRUE)
```

```
## Warning: Ignoring unknown parameters: facet
```



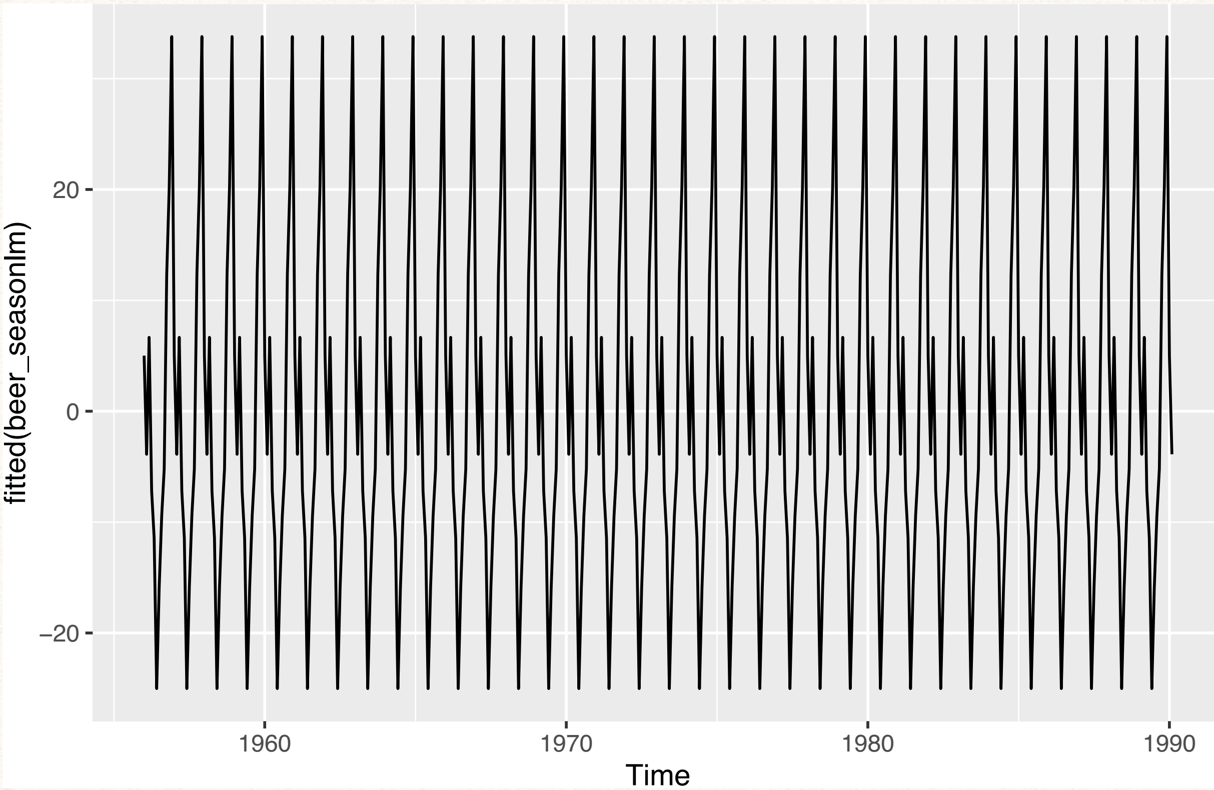

```r
#autoplot(sales_resides,facet=TRUE)
ggAcf(beer_resides)
```



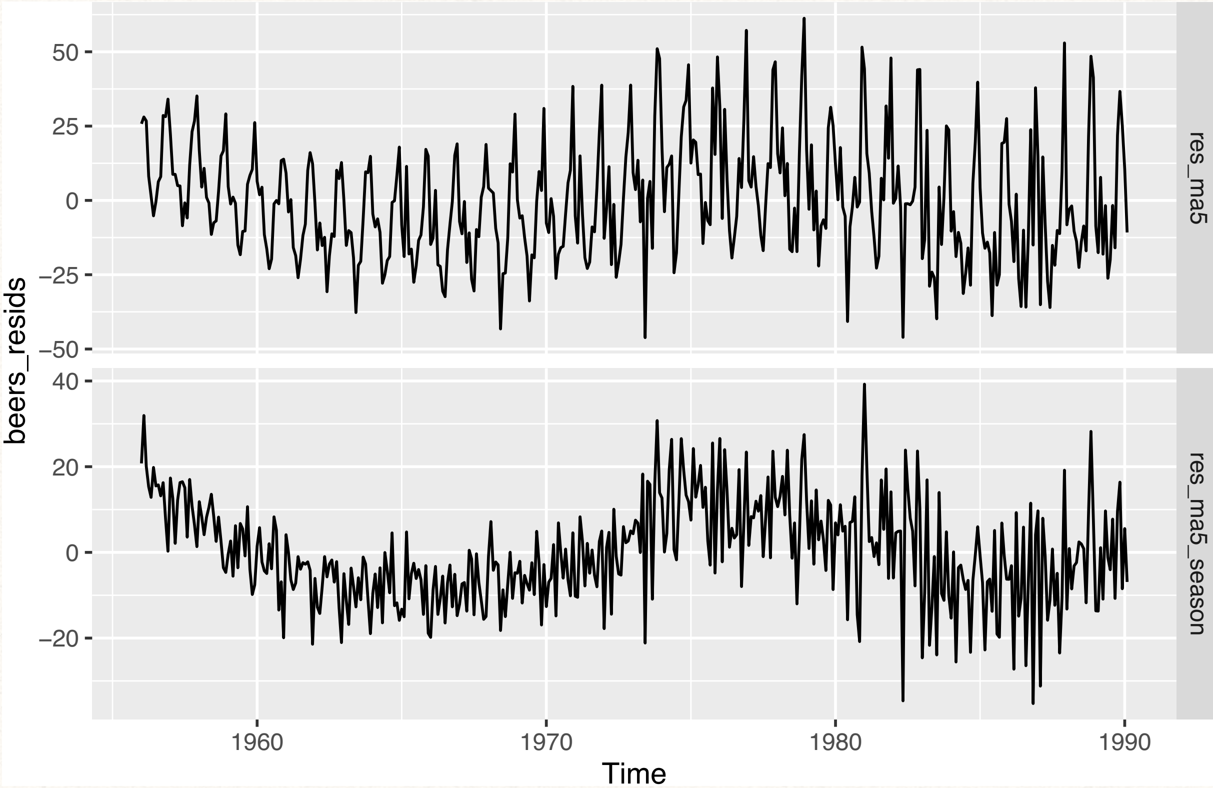
```r
frequency(beer_resides)
```

```
## [1] 12
```

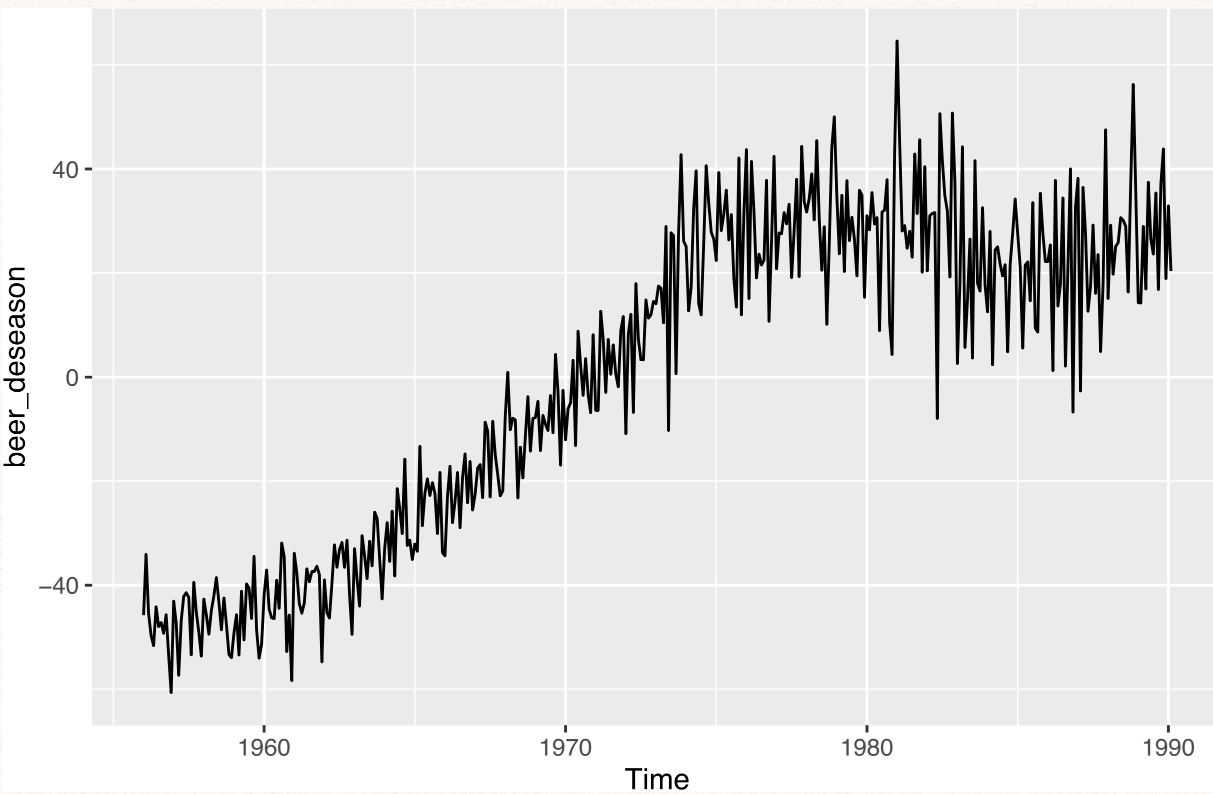
```r
beer_seasonlm<-tslm(beer_resides~season)
autoplot(fitted(beer_seasonlm))
```



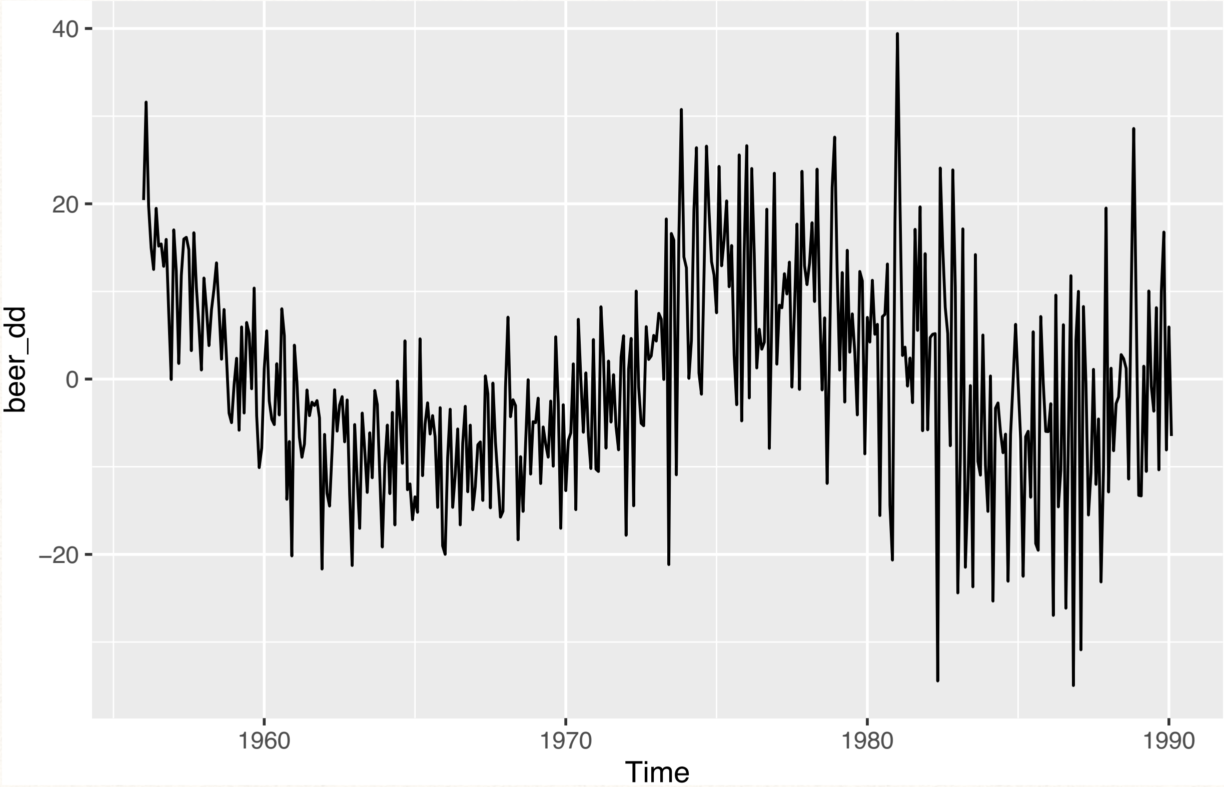

```r
beers_resids<-cbind(res_ma5=beer_resides,
                    res_ma5_season=beer_resides-fitted(beer_seasonlm))
autoplot(beers_resids,facet=TRUE)
```



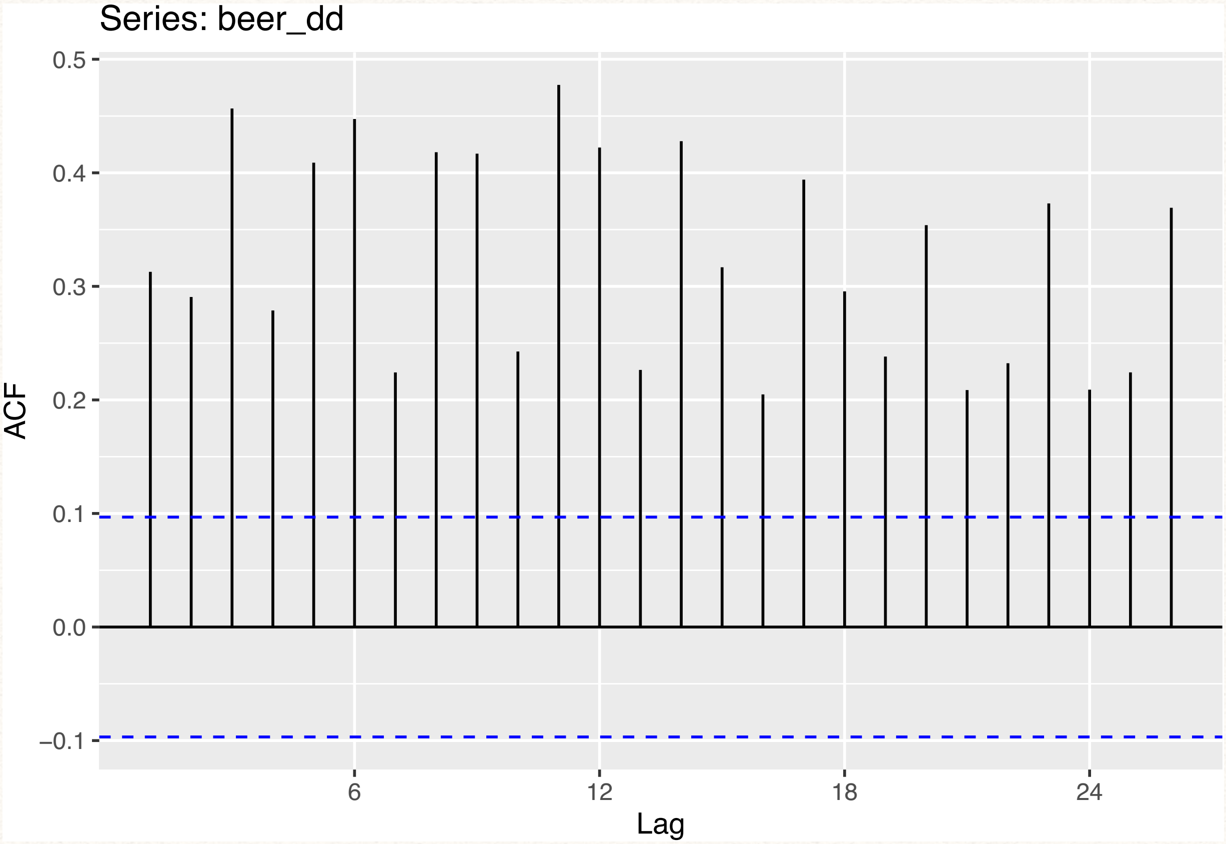

```
beer_deseason<-beer-mean(beer)-fitted(beer_seasonlm)
autoplot(beer_deseason)
```



```
beer_qua<- tslm(beer_deseason~trend+I(trend^2))
beer_dd<-beer_deseason-fitted(beer_qua)
autoplot(beer_dd)
```



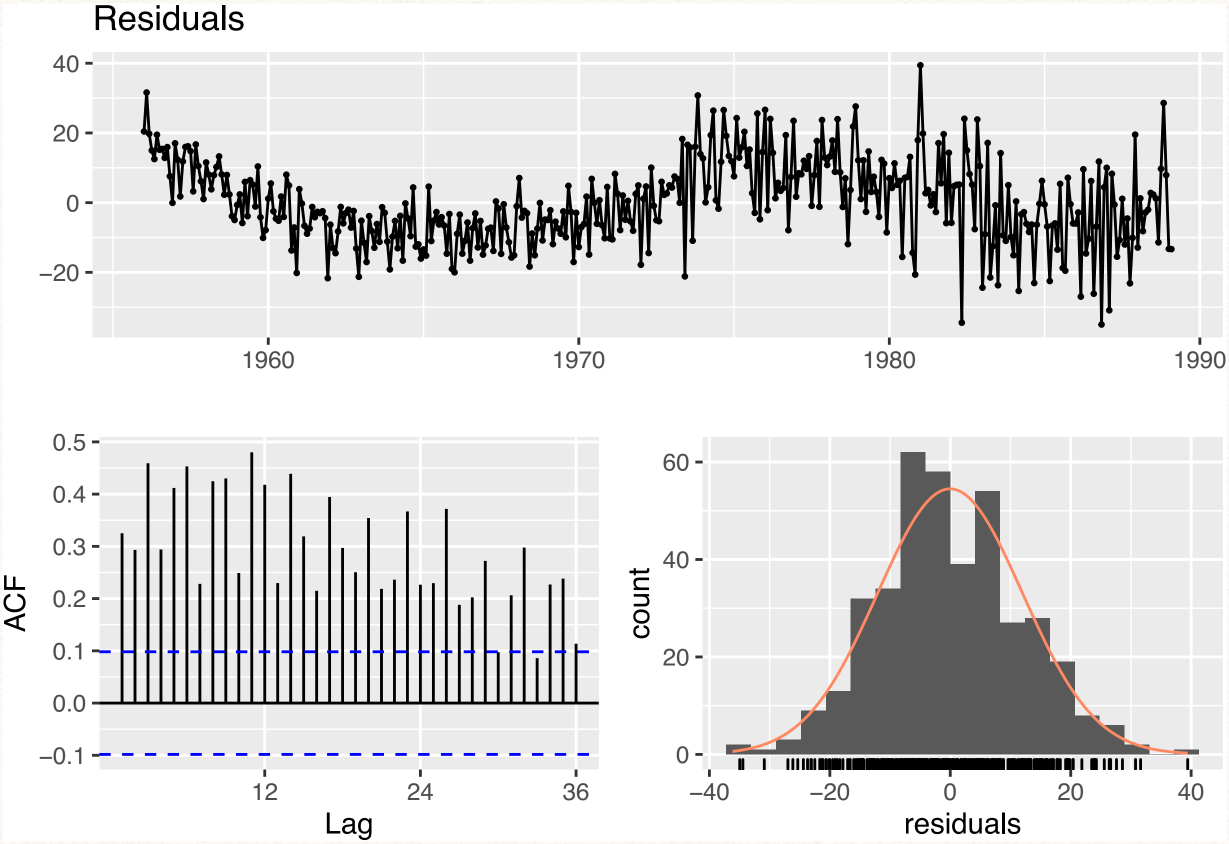

```
ggAcf(beer_dd)
```



By the autoplot and residual plot of the data, it shows that the data may possibly be a white noise process, with low correlation.

```
beer_ddd<-head(beer_dd,-12)
checkresiduals(beer_ddd)
```

```
## Warning in modeldf.default(object): Could not find appropriate degrees of
## freedom for this model.
```



Try to do the ARIMA model again:

```
pro<-auto.arima(beer_ddd,stepwise=FALSE, seasonal=FALSE, ic='aic',trace=TRUE)
```

```
##
##  Fitting models using approximations to speed things up...
##
##  ARIMA(0,1,0)                                : 3215.129
##  ARIMA(0,1,0)            with drift          : 3217.114
##  ARIMA(0,1,1)                                : 2946.174
##  ARIMA(0,1,1)            with drift          : 2947.363
##  ARIMA(0,1,2)                                : 2939.357
##  ARIMA(0,1,2)            with drift          : 2940.609
##  ARIMA(0,1,3)                                : 2924.046
##  ARIMA(0,1,3)            with drift          : 2925.451
##  ARIMA(0,1,4)                                : 2925.712
##  ARIMA(0,1,4)            with drift          : 2927.091
##  ARIMA(0,1,5)                                : 2915.388
##  ARIMA(0,1,5)            with drift          : 2916.907
##  ARIMA(1,1,0)                                : 3111.811
##  ARIMA(1,1,0)            with drift          : 3113.748
##  ARIMA(1,1,1)                                : 2952.66
##  ARIMA(1,1,1)            with drift          : 2953.785
##  ARIMA(1,1,2)                                : 2953.807
##  ARIMA(1,1,2)            with drift          : 2954.976
##  ARIMA(1,1,3)                                : 2933.006
##  ARIMA(1,1,3)            with drift          : 2934.259
##  ARIMA(1,1,4)                                : 2933.831
```

```
##  ARIMA(1,1,4)            with drift         : 2935.139
##  ARIMA(2,1,0)                               : 3005.483
##  ARIMA(2,1,0)            with drift         : 3007.368
##  ARIMA(2,1,1)                               : 2929.203
##  ARIMA(2,1,1)            with drift         : 2930.578
##  ARIMA(2,1,2)                               : 2928.436
##  ARIMA(2,1,2)            with drift         : 2929.725
##  ARIMA(2,1,3)                               : 2910.104
##  ARIMA(2,1,3)            with drift         : 2912.099
##  ARIMA(3,1,0)                               : 2998.079
##  ARIMA(3,1,0)            with drift         : 2999.973
##  ARIMA(3,1,1)                               : 2928.147
##  ARIMA(3,1,1)            with drift         : 2929.609
##  ARIMA(3,1,2)                               : 2927.022
##  ARIMA(3,1,2)            with drift         : 2928.434
##  ARIMA(4,1,0)                               : 2959.86
##  ARIMA(4,1,0)            with drift         : 2961.755
##  ARIMA(4,1,1)                               : 2915.703
##  ARIMA(4,1,1)            with drift         : 2917.263
##  ARIMA(5,1,0)                               : 2934.629
##  ARIMA(5,1,0)            with drift         : 2936.461
##
##  Now re-fitting the best model(s) without approximations...
##
##
##
##
##  Best model: ARIMA(2,1,3)
```

```r
summary(pro)
```

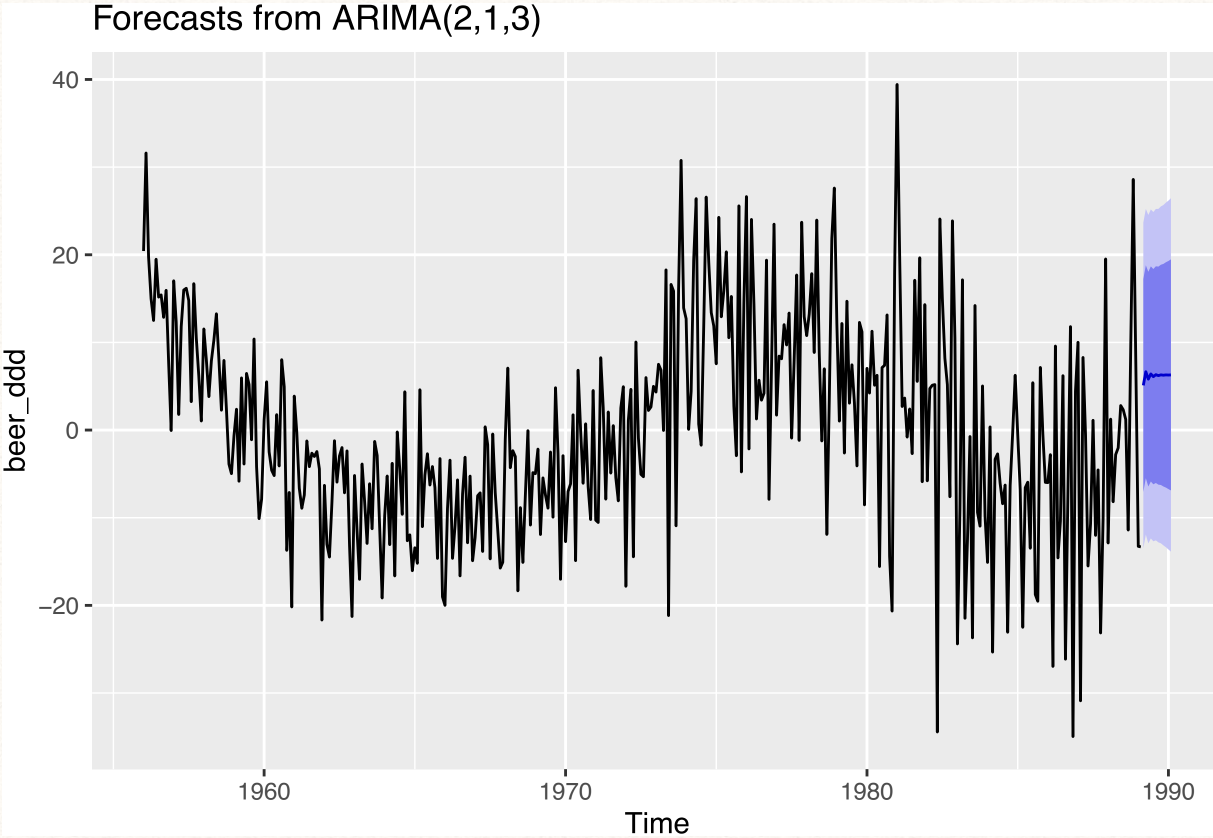
```
## Series: beer_ddd
## ARIMA(2,1,3)
##
## Coefficients:
##          ar1     ar2      ma1      ma2     ma3
##       0.0996  0.4361  -1.1228  -0.4402  0.6487
## s.e.  0.1378  0.1212   0.1208   0.2098  0.1093
##
## sigma^2 estimated as 89.82:  log likelihood=-1455.05
## AIC=2922.11   AICc=2922.33   BIC=2946.01
##
## Training set error measures:
##                     ME     RMSE     MAE      MPE    MAPE      MASE         ACF1
## Training set -0.202667 9.405641 7.28973 50.03155 318.183 0.7778955 -0.03251424
```

```r
confint(pro)
```

```
##          2.5 %      97.5 %
## ar1 -0.1705255  0.36979818
## ar2  0.1985049  0.67371573
## ma1 -1.3595752 -0.88611656
## ma2 -0.8514258 -0.02906865
## ma3  0.4344729  0.86291610
```

Here the 95% confidence interval shows the bounds od parameter.

```r
# forcast the 12-step data
foo<-forecast(pro,h=12)
autoplot(foo)
```



Now show the result by table

```r
#b1<-log(tail(beer_dd,12))
b1<-tail(beer_dd,12)
e1<-b1-foo$mean
d1<- data_frame(e1,foo$lower,foo$upper,foo$mean)
```

```
## Warning: `data_frame()` is deprecated, use `tibble()`.
## This warning is displayed once per session.
```

```r
colnames(d1)<-c('error','2.5%','97.5%','prediction')
d1
```

```
## # A tibble: 12 x 4
##      error `2.5%`[,"80%"] [,"95%"] `97.5%`[,"80%"] [,"95%"] prediction
##      <dbl>          <dbl>    <dbl>           <dbl>    <dbl>      <dbl>
##  1  -3.62           -7.05    -13.5            17.2     23.7       5.10
##  2 -17.2            -5.49    -11.9            18.8     25.2       6.65
##  3   4.24           -6.44    -12.9            18.1     24.5       5.81
##  4  -7.21           -5.87    -12.4            18.7     25.2       6.40
##  5  -9.76           -6.19    -12.7            18.4     24.9       6.09
##  6   1.83           -6.04    -12.6            18.7     25.2       6.32
##  7 -16.6            -6.23    -12.8            18.6     25.2       6.21
##  8   3.69           -6.27    -12.9            18.9     25.5       6.30
##  9  10.5            -6.44    -13.2            19.0     25.7       6.26
```

```
## 10 -14.4              -6.56    -13.4             19.1     26.0        6.29
## 11  -0.328            -6.73    -13.6             19.3     26.2        6.28
## 12 -12.8              -6.89    -13.9             19.5     26.5        6.29
```

It is obvious that comparing with part b, the error od prediction become much smaller.