Time series Analysis
Chapter 8-Seasonality

Muyi LI

Xiamen University

Outline:

- Seasonal ARMA(P, Q) Models;
- Multiplicative Seasonal ARIMA$(p, d, q) \times (P, D, Q)$;
- Illustration of Real Data Analysis

Quarterly earnings per share of the Coca-Cola Company

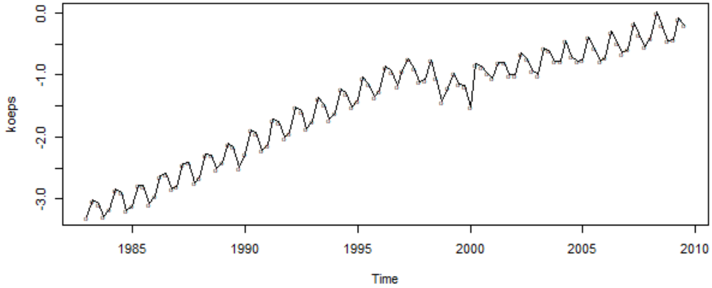


Figure: Quarterly earnings per share of the Coca-Cola from 1983.I to 2009 III

Quarterly earnings per share of the Coca-Cola Company

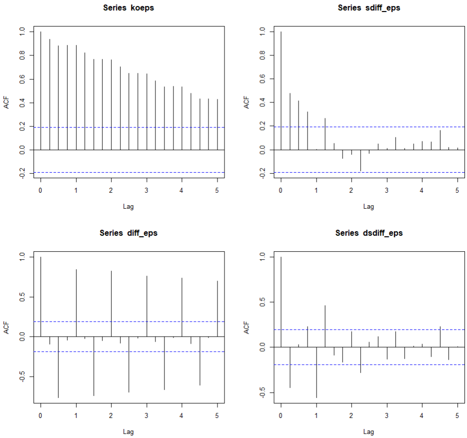


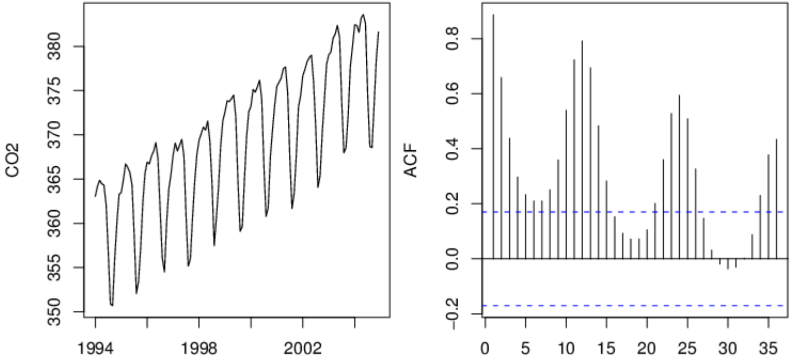
Figure: Sample ACF of the log-series of quarterly earnings per share of the Coca-Cola Company from 1983.I to 2009 III, where diffeps is the first differenced series, sdiffeps is the seasonally differenced series, and dsdiffeps denotes both first and seasonal differencing.

Time series with Seasonal Effect

- **Definition of Seasonality:** A seasonal pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week). Seasonality is always of a fixed and known period. Hence, seasonal time series are sometimes called periodic time series.
- More examples.

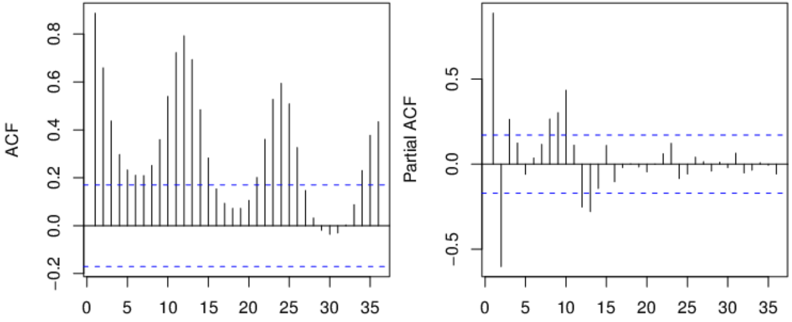
Time Series with Seasonal Effect

Consider the monthly CO2 levels at Alert, Northwest Territories, Canada, from January 1994 to December 2004.



library(TSA)
data("CO2")
plot(co2,ylab='CO2')

Its sample ACF and sample PACF are as follows.



Stationary Seasonal ARMA models

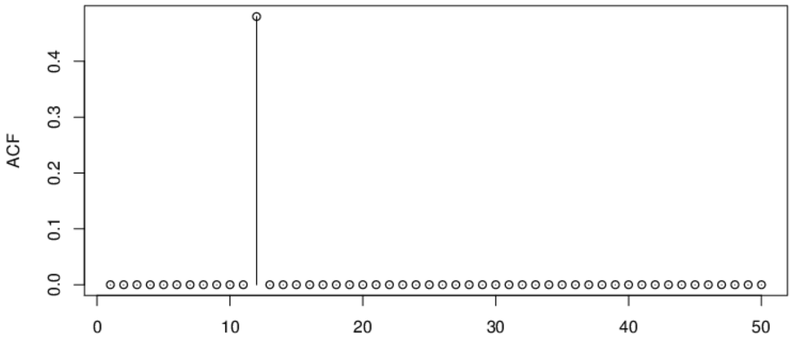
Consider the seasonal MA(1) model,

$$y_t = e_t - \Theta e_{t-12},$$

where $\{e_t\}$ is i.i.d. with mean 0 and variance σ_e^2.

- Note that this is just a special MA model. It is always **stationary** just like the common MA models.
- The variance is $\gamma_0 = (1 + \Theta^2)\sigma_a^2$.
- Its ACF has the property that $\rho_{12} = -\Theta/(1 + \Theta^2)$ and $\rho_k = 0$ as $k \neq 12$.

ACF of seasonal MA(1) process



seasonal MA(Q) model

- The seasonal MA(Q) model of order P and seasonal period s is defined as

$$y_t = e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \cdots - \Theta_Q e_{t-Qs},$$

  where $\{e_t\} \sim i.i.d.(0, \sigma_e^2)$.
- The seasonal MA characteristic polynomial is defined as

$$\Theta(x) = 1 - \Theta_1 x^s - \Theta_2 x^{2s} - \cdots - \Theta_Q x^{Qs}.$$

- It is a special case of an ordinary MA model of order $q = Qs$ but with all coefficients equal to zero except at the seasonal lags $s, 2s, \cdots, Qs$.
- It is always stationary just like the common MA models.
- It is invertible if all the roots of $\Theta(x) = 0$ are outside the unit circle.

Properties of ACF and PACF in Seasonal MA(Q) with period s

- Its ACFs will be nonzero ONLY at the seasonal lags of $s, 2s, \cdots, Qs$. In particular, for $k = 1, 2, \dot{;} Q$,

$$\rho_{k*s} = \frac{-\Theta_k + \Theta_1\Theta_{k+1} + \Theta_2\Theta_{k+2} + \cdots + \Theta_{Q-k}\Theta_Q}{1 + \Theta_1^2 + \Theta_2^2 + \cdots + \Theta_Q^2}.$$

- Its PACFs are nonzero only at lags $s, 2s, \cdots$, where it behaves like a combination of decaying exponentials and damped sine functions.

Seasonal AR model

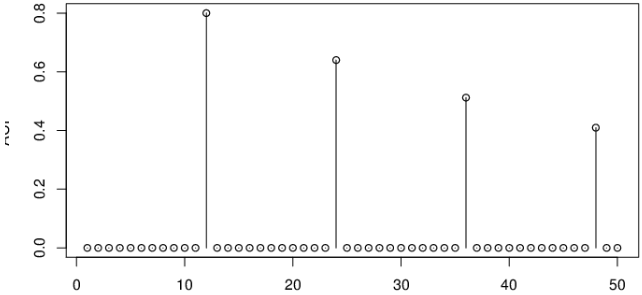
Consider the seasonal AR(1) model with period 12,

$$y_t = \Phi y_{t-12} + e_t,$$

where $\{e_t\} \sim i.i.d.(0, \sigma_a^2)$.

- It is just a special AR model. It is always invertible.
- Its AR characteristic polynomial has the form of $\Phi(x) = 1 - \Phi x^{12}$, and the stationarity condition is $|\Phi| < 1$.
- Its ACF has the property that $\rho_{12*k} = \Phi^k$ for $k = 1, 2, \cdots$, and the values of the ACF at other lags are all zero.

ACF of Seasonal AR(1) model with period 12



seasonal AR(P) model

- The seasonal AR(P) model with period s is defined as

$$y_t = \Phi_1 y_{t-s} + \Phi_2 y_{t-2*s} + \cdots + \Phi_P y_{t-P*s} + e_t,$$

where $\{e_t\} \sim i.i.d.(0, \sigma_e^2)$.

- The seasonal AR characteristic polynomial is

$$\Phi(x) = 1 - \Phi_1 x^s - \Phi_2 x^{2s} - \dot{-} \Phi_P x^{Ps}.$$

- It is a special case of ordinary AR(p) model of order $p = Ps$ with nonzero coefficients only at the seasonal lags $s, 2s, \cdots, Ps$.
- It is stationary if all the roots of $\Phi(x) = 0$ are outside the unit circle.
- It is always invertible.
- Its ACFs are nonzero only at lags $s, 2s, \cdots,$ where it behaves like a combination of decaying exponentials and damped sine functions.
- Its PACFs are nonzero only at lags $s, 2s, \cdots, Ps$, and *cuts off after P seasonal periods*.

Multiplicative Seasonal ARMA$(p, q) \times (P, Q)_s$ Model

- Consider the non-seasonal and seasonal effect together, we define a multiplicative seasonal ARMA$(p, q) \times (P, Q)_s$ model

$$\phi(B)\Phi(B)Y_t = \theta(B)\Theta(B)a_t$$

- It can contain a constant term μ:

$$\phi(B)\Phi(B)Y_t = \mu + \theta(B)\Theta(B)a_t$$

or

$$\phi(B)\Phi(B)(Y_t - c) = \theta(B)\Theta(B)a_t$$

- where

$$\begin{aligned}
\phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \\
\Phi(B) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}, \\
\theta(B) &= 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q, \\
\Theta(B) &= 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs}.
\end{aligned}$$

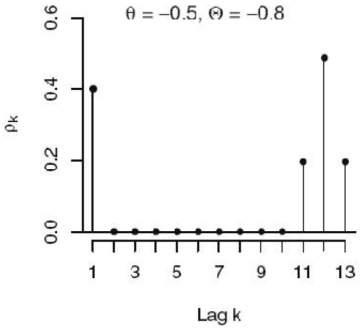
ARMA$(p, q) \times (P, Q)_s$ Model

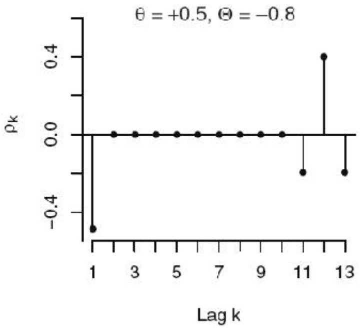
- It is a special case of ordinary ARMA model with AR order $p + Ps$ and MA order $q + Qs$,
- The number of nonzero coefficients is only $p + P + q + Q$, which is considerably smaller than $p + Ps + q + Qs$, and hence allow a much more parsimonious model.

Example: ARMA$(0,1)(0,1)_{12}$

- Consider $y_t = e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13}$.
- It is actually a seasonal ARMA$(0,1)(0,1)_{12}$ model and can be written as:
$$y_t = (1-\theta B)(1-\Theta B^{12})e_t.$$
- $$\gamma_0 = (1+\theta^2)(1+\Theta^2)\sigma_a^2,$$
$$\rho_1 = -\frac{\theta}{1+\theta^2}, \quad \rho_{12} = -\frac{\Theta}{1+\Theta^2},$$
and
$$\rho_{11} = \rho_{13} = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)}.$$
Autocorrelations at all other lags are zero.

ACF of ARMA$(0,1)(0,1)_{12}$





Example: ARMA$(0,1)(1,0)_{12}$ Model

- $y_t = \Phi y_{t-12} + e_t - \theta e_{t-1}$ or $(1 - \Phi B^{12}) y_t = (1 - \theta B) e_t$.
- $$\gamma_1 = \Phi \gamma_{11} - \theta \sigma_a^2 \quad \text{and} \quad \gamma_k = \Phi \gamma_{k-12} \text{ for } k \geq 2.$$
- $$\begin{aligned} \gamma_0 &= \left(\frac{1+\theta^2}{1-\Phi^2} \right) \sigma_a^2, \\ \rho_{12k} &= \Phi^k \text{ for } k \geq 1, \\ \rho_{12k-1} = \rho_{12k+1} &= \left(-\frac{\theta}{1+\theta^2} \Phi^k \right) \text{ for } k = 0, 1, 2, \cdots, \end{aligned}$$

ACF of ARMA$(0,1)(1,0)_{12}$

Quarterly earnings per share of the Coca-Cola Company

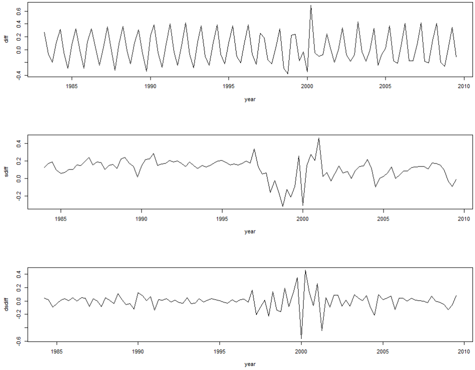


Figure: The top panel is for the first differenced series, the middle panel is for the seasonally differenced series and the lower panel is for both the first and seasonally differenced series.

Quarterly earnings per share of the Coca-Cola Company

- Fitting the data by the ARIMA(0,1,1)(0,1,1) model, then we have

$$(1-L)(1-L^4)x_t = (1-0.4096_{(0.0866)}L)(1-0.8203_{(0.0743)}L^4)a_t,$$

  with $\hat{\sigma}_a^2 = 0.00724$.

- ```
  m1=arima(koeps,order=c(0,1,1),seasonal=list(order=c(0,1,1),period=4)
  ```
  ```

Quarterly earnings per share of the Coca-Cola Company

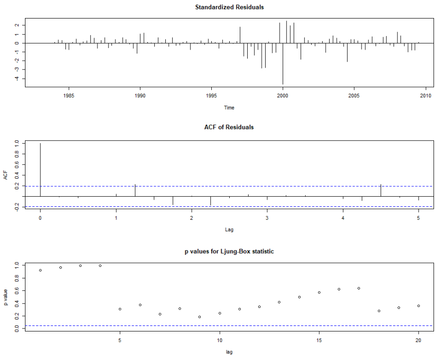


Figure: Model checking.(a) Time plot of the standardized residuals;(b)ACF of the standardized residuals;(3)plot of $p-$ values of the Ljung-Box statistics of the standardized residuals.

```
tsdiag(m1,gof=20)
Box.test(m1$residuals,lag=12,type='Ljung')
```

Quarterly earnings per share of the Coca-Cola Company

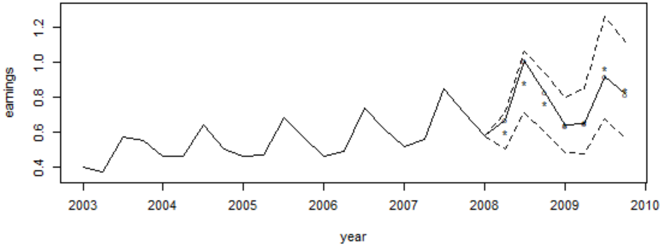


Figure: Out-of-sample point and interval forecasts.The forecast origin is the fourth quarter of 2007.Cirles represent the actual earnings in the forecasting period, asterisks represent point forecasts, and dashed lines represents a 95% interval forecasts

R code for Quarterly earnings per share of the Coca-Cola Company

```
da=read.table(D:/data/q-ko-earns8309.txt, head=T)
head(da)
eps=log(da$value)
koeps=ts(eps,frequency=4, start=c(1983,1))
plot(koeps,type=l)
diffeps=diff(koeps)
sdiffeps=diff(koeps,4)
dsdiffeps=diff(sdiffeps)
par(mfcol=c(2,2))
acf(koeps,lag=20)
acf(diffeps,lag=20)
acf(sdiffeps,lag=20)
acf(dsdiffeps,lag=20)
```

Prediction

```
pm1=predict(m1,7)
names(pm1)
pred=pm1$pred
se=pm1$se
```

R code: Cont'd

Anti-log transformation

```
ko=da$value
fore=exp(pred+se^2/2)
v1=exp(2*pred+se^2)*(exp(se^2)-1)
s1=sqrt(v1)
eps=ko[80:107]
tdx=(c(1:28)+3)/4+2002
upp=c(ko[100], fore+2*s1)
low=c(ko[100], fore-2*s1)
plot(tdx,eps,xlab='year',ylab='earnings',type='l',ylim=c(0.35,1.3))
points(tdx[22:28],ko[101:107],pch='o',cex=0.7)
points(tdx[22:28],fore,pch='*')
lines(tdx[21:28],upp,lty=2)
lines(tdx[21:28],low,lty=2)
```

Identify a Seasonal Model

1. **Step 1**: Do a time series plot of the data. Examine it for features such as trend and seasonality.
2. **Step 2**: Do any necessary differencing. The general guidelines are:
   - If there is seasonality and no trend, then take a difference of lag S (period).
   - If there is linear trend and no obvious seasonality, then take a first difference. If there is a curved trend, consider a transformation of the data before differencing.
   - If there is both trend and seasonality, apply a seasonal difference to the data and then re-evaluate the trend. If a trend remains, then take first differences.
   - If there is neither obvious trend nor seasonality, dont take any differences.

Identify a Seasonal Model: Cont'd

- **Step 3:** Examine the ACF and PACF of the differenced data (if differencing is necessary).
  - We are using this information to determine possible models. This can be tricky going involving some (educated) guessing. Some basic guidance:
    - Non-seasonal terms: Examine the early lags (1, 2, 3,) to judge non-seasonal terms. Spikes in the ACF (at low lags) indicate non-seasonal MA terms. Spikes in the PACF (at low lags) indicated possible non-seasonal AR terms.
    - Seasonal terms: Examine the patterns across lags that are multiples of S. For example, for monthly data, look at lags 12, 24, 36, and so on (probably wont need to look at much more than the first two or three seasonal multiples). Judge the ACF and PACF at the seasonal lags in the same way you do for the earlier lags.

Identify a Seasonal Model: Cont'd

- **Step 4:** Estimate the model(s) that might be reasonable on the basis of Step 3. Dont forget to include any differencing that you did before looking at the ACF and PACF. In the software, specify the original series as the data and then indicate the desired differencing when specifying parameters in the arima command that youre using.
- **Step 5:** Examine the residuals (with ACF, Box-Pierce, and any other means) to see if the model seems good. Compare AIC or BIC values if you tried several models.
  If things dont look good here, its back to Step 3 (or maybe even Step 2).

Homework of Seasonal ARMA model

1. Assume y_t is an $ARMA(1,1) \times (1,0)_{12}$,

$$(1-\phi_1 B)(1-\Phi_1 B^{12})y_t = (1-\theta_1 B)e_t.$$

   - Express y_t by an ordinary ARMA representation.
   - What's the properties of ACF and PACF of this model?
   - Simulate a sample sequence with 1000 data points from this model, plot sample ACFs and PACFs for the first 36 lags.

2. Consider an AR(p) model, $Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + a_t$, and a seasonal AR($p$) model, $Z_t = \phi_1 Z_{t-s} + \cdots + \phi_p Z_{t-ps} + a_t$, where $\{a_t\} \sim WN(0, \sigma_a^2)$. Show that the AR($p$) model is stationary if and only if the seasonal AR(p) model is stationary.