Ch 6 : Seasonal ARIMA model

Contents

April 19, 2017

1

Trend with Seasonality

[ToC]

1.1 Trend with Seasonality period s

[ToC]

Suppose Y_t is the observation. The Model says:

$$Y_t = m_t + S_t + X_t$$

$$\begin{aligned} m_t &: \quad \text{Trend Component} \\ S_t &: \quad \text{Seasonal Component} \\ X_t &: \quad \text{Stationary Time Series} \end{aligned}$$

where $EY_t = 0$.

Condition on Seasonal component: $S_{t+s} = S_t$ and $\sum_{j=1}^{s} S_j = 0$.

Example: Accident Data

Accidental deaths in USA., 1973 to 1978 from Brockwell

```
acf1 <- acf; library(TSA); acf <- acf1  #- Load TSA package

D  <- read.csv("http://gozips.uakron.edu/~nmimoto/pages/datasets/acci.txt", header=T)

D1 <- ts(D, start=c(1973,1), freq=12)  #- Turn D into ts object with frequency

plot(D1, type='o', ylab="num of accidents")
D1

#--- plot with month ---
plot(D1, type="l", ylab="num of accidents")
points(y=D1, x=time(D1), pch=as.vector(season(D1)))
```

Example: Oil Filter Sales Data

Oil Filter Sales Data (Cryer p7) inside TSA package.

```
acf1 <- acf; library(TSA); acf <- acf1  #- Load TSA package

data(oilfilters)
D2 <- oilfilters

is.ts(D2) # is data in TS format?
D2        # look inside

#--- plot with month ---
plot(D2, type="l",ylab="Sales")
points(y=D2, x=time(D2), pch=as.vector(season(D2)))
```

Seasonality

1. s is the seasonality frequency. (e.g. $s = 12$ for monthly seasonality).

2. Seasonality repeats every s observation.

$$S_{t+s} = S_t.$$

3. Seasonality is not a trend. It's a temporary deviation from overall trend.

$$\sum_{j=1}^{s} S_j = 0.$$

1.2 Removing Seasonality

[ToC]

---

1. MA filter
2. Seasonal Average
3. Seaonal Differencing

1.3 Method 1: MA filter

[ToC]

If s is odd, let it be $2q+1$. Then use linear MA filter.

$$\hat{m}_t = \frac{1}{(2q+1)} \sum_{i=-q}^{q} X_{t-j}$$

If s is even, let it be $2q$. Then use

$$\hat{m}_t = \frac{1}{2q}\Big(.5X_{t-q} + \sum_{i=-q-1}^{q-1} X_{t-j} + .5X_{t+q} \Big)$$

Because seasonality should sum up to zero for each seasonal cycle,
we have estimated trend:

$$\hat{m}_t = \frac{1}{(2q+1)} \sum_{i=-q}^{q} m_{t-j} + \underbrace{\frac{1}{(2q+1)} \sum_{i=-q}^{q} S_{t-j}}_{0} + \underbrace{\frac{1}{(2q+1)} \sum_{i=-q}^{q} Y_{t-j}}_{\text{small}}.$$

Then estimate for $(S_t + Y_t)$ is

$$w_k = X_t - \hat{m}_t \qquad q < t < n - q.$$

Use this to estimate the seasonal part,

$$\hat{S}_k = w_k - \frac{1}{s}\sum_{i=1}^{s} w_i.$$

Now we have deseasonalized data $(m_t + Y_t)$

$$d_t = X_t - \hat{S}_t \qquad t = 1 \cdots n$$

Now go back and re-estimate trend using d_t.

1.4 Method 2: Seasonal Average

[ToC]

Suppose we have monthly seasonality, $s = 12$. Then take average for each month. For example, average for January will be

$$\bar{S}_1 = \sum_{j=0} X_{1+12j}$$

and July average will be

$$\bar{S}_7 = \sum_{j=0} X_{7+12j}.$$

Note that this seasonal average will take out the trend part as well.

Example: Accidents

```
plot(D1, type="o")

#--- Take Monthly Averages

Mav1  <- aggregate(c(D1), list(month=cycle(D1)), mean)$x            #- 1yr long Mtly Av $
M.av1 <- ts(Mav1[cycle(D1)], start=start(D1), freq=frequency(D1))    #- Mtly Av as long as D1
Ds.D1 <- D1-M.av1                                                    #- Subtract from original

plot(M.av1, type="o")

plot(Ds.D1, type="o")


layout(matrix(1:3, 3,1))
plot(D1,    type="o")
plot(M.av1, type="o")
plot(Ds.D1, type="o"); abline(h=0)

Stationarity.tests(Ds.D1)
```

Example: Oil Filter Sales

```
plot(D2, type="o")

#--- Take Monthly Averages

Mav2  <- aggregate(c(D2), list(month=cycle(D2)), mean)$x            #- 1yr long Mtly Av $
M.av2  <- ts(Mav2[cycle(D2)], start=start(D2), freq=frequency(D2))   #- Mtly Av as long as D2
Ds.D2 <- D2-M.av2                                                   #- Subtract from original

plot(M.av2, type="o")

plot(Ds.D2, type="o")


layout(matrix(1:3, 3,1))
plot(D2, type="o")
plot(M.av2, type="o")
plot(Ds.D2, type="o"); abline(h=0)

Stationarity.tests(Ds.D2)
```

Fit deseasonalized series with ARMA

```
library(forecast)
source("http://gozips.uakron.edu/~nmimoto/477/TS_R-90.txt")

auto.arima(Ds.D1)
auto.arima(Ds.D2)

Fit1 <- auto.arima(Ds.D1, d=0)
Fit1

Fit2 <- auto.arima(Ds.D2, d=0)
Fit2

plot(M.av2, type="o")

plot(Ds.D2, type="o")


layout(matrix(1:3, 3,1))
plot(D2, type="o")
plot(M.av2, type="o")
plot(Ds.D2, type="o"); abline(h=0)
```

1.5 Method 3: Seasonal Differencing (Box-Jenkins Method)

[ToC]

For guessed seasonality s, define

$$\nabla_s = (1 - B^s)$$

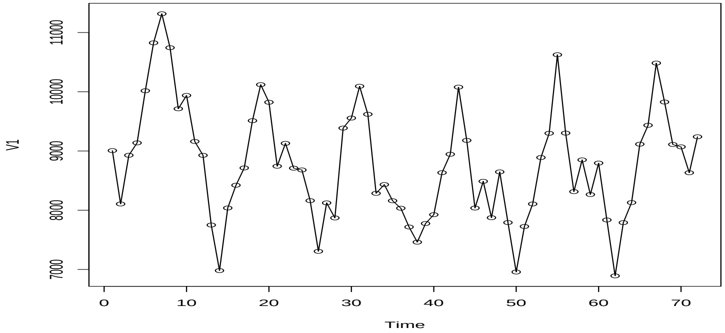
(Don't confuse this with ∇^s). Then we have

$$\begin{aligned} \nabla_s X_t &= (1 - B^s)X_t = X_t - X_{t-s} \\ &= m_t - m_{t-s} + \underbrace{S_t - S_{t-s}}_{0} + Y_t - Y_{t-s}. \end{aligned}$$

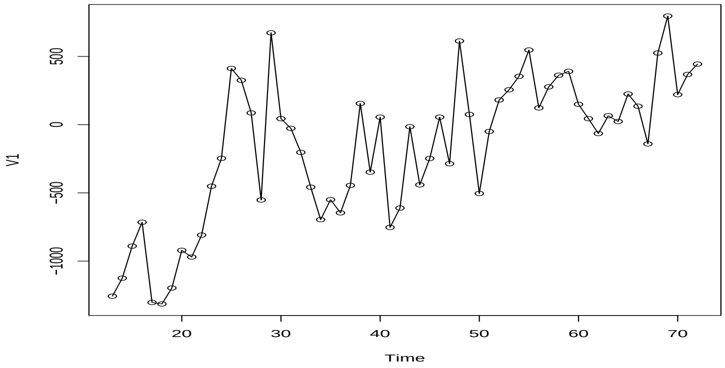
since $S_t = S_{t-s}$, this eliminates the seasonality. Now fit ARIMA to $\nabla_s X_t$.

Example: Accidents

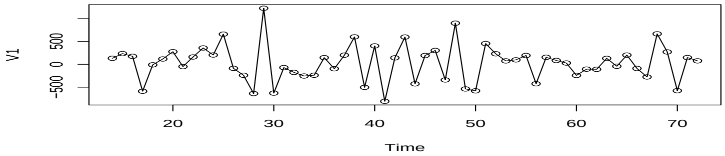
Monthly Accidental Death in USA 1973-1978 From Brockwell and Davis (2002)

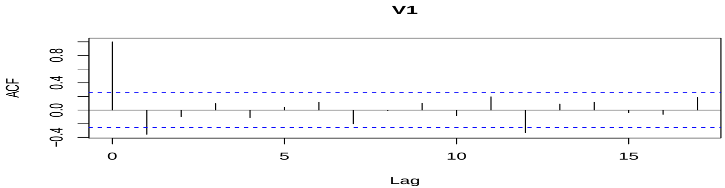


Plot of $\bigtriangledown_{12} X_t$



Plot of $\bigtriangledown \bigtriangledown_{12} X_t$ and its ACF





Example: Accident Data

Accidental deaths in USA., 1973 to 1978 from Brockwell

```
D  <- read.csv("http://gozips.uakron.edu/~nmimoto/pages/datasets/acci.txt", header=T)
D1 <- ts(D, start=c(1973,1), freq=12)  #- Turn D into ts object with frequency

plot(D1, type='o', ylab="num of accidents")
D1

plot(diff(D1, 12))                # take seasonal difference (Del_12)

plot(diff(diff(D1, 12)))          # take (Del)(Del_12)

Stationarity.tests(diff(diff(D1, 12)))

auto.arima(diff(diff(D1, 12)))

Fit1 <- Arima(diff(diff(D1, 12)), order=c(0,0,1) )

Randomness.tests(Fit1$resid)      #$
```

Seasonal ARIMA:

We took ∇_{12}, then ∇, then fit MA(1). Therefore our model is

$$\begin{aligned} \nabla \nabla_{12} Y_t &= X_t \\ X_t &= e_t - \theta_1 e_{t-1} \end{aligned}$$

This is called

$$\text{sARIMA}(p, d, q)(P, D, Q)_{12} \quad \text{model}$$

$$\text{sARIMA}(0, 1, 1)(0, 1, 0)_{12} \quad \text{model}$$

Example: Oil Filter Sales Data

Oil Filter Sales Data (Cryer p7) inside TSA package.

```
acf1 <- acf; library(TSA); acf <- acf1
data(oilfilters);  D2 <- oilfilters
D2

plot(diff(D2, 12))                # take seasonal difference (Del_12)

Stationarity.tests(diff(D2, 12))

auto.arima(diff(D2, 12))
```

Seasonal ARIMA for Oil Filter:

We took $\bigtriangledown_{12}$, then it looks like a WN. Therefore our model is

$$\begin{aligned} \bigtriangledown_{12} X_t &= X_t \\ X_t &= e_t \end{aligned}$$

This is called

$$\text{sARIMA}(p, d, q)(P, D, Q)_{12} \quad \text{model}$$

$$\text{sARIMA}(0, 0, 0)(0, 1, 0)_{12} \quad \text{model}$$

2

SARIMA model

[ToC]

2.1 SARIMA model

[ToC]

- What if there's an annual effeect on monthly data?
- i.e. Aug 2015 will depend on Aug 2014, Aug 2013, and so on...
- Jan 2015 will depend on Jan 2014, Jan 2013, and so on...
- That's autocorrelation with lag 12.
- Consider MA(12) with only one coefficient, $\theta_{12} = \Theta$

$$X_t = e_t - \Theta e_{t-12}.$$

Calculate ACVF,

$$
\begin{aligned}
\gamma(1) &= \mathrm{Cov}(X_{t+1}, X_t) = \mathrm{Cov}\Big(e_{t+1} - \Theta e_{t-11},\ \ e_t - \Theta e_{t-12}\Big) = 0 \\
\gamma(12) &= \mathrm{Cov}(X_{t+12}, X_t) = \mathrm{Cov}\Big(e_{t+12} - \Theta e_t,\ \ e_t - \Theta e_{t-12}\Big) = -\Theta\sigma^2.
\end{aligned}
$$

There's correlation only at lag 12.

Seasonal MA(Q) with $s = 12$

$$X_t = e_t - \Theta_1 e_{t-12} \qquad sMA(Q = 1) \quad \text{(s=12)}$$

$$X_t = e_t - \Theta_1 e_{t-12} - \Theta_2 e_{t-24} \qquad sMA(Q = 2) \quad \text{(s=12)}$$

$\gamma(h)$ will be zero except at las s, $2s$, $3s$, up to Qs.

Representation with B

- sMA(1) s=12

$$\begin{aligned} X_t &= e_t - \Theta_1 e_{t-12} \\ &= \underbrace{(1 - \Theta_1 B^{12})}_{\text{seasonal characteristic polynomial}} e_t \end{aligned}$$

- sMA(2) s=12

$$\begin{aligned} X_t &= e_t - \Theta_1 e_{t-12} - \Theta_2 e_{t-24} \\ &= \underbrace{(1 - \Theta_1 B^{12} - \Theta_2 B^{24})}_{\text{seasonal characteristic polynomial}} e_t \end{aligned}$$

Seasonal AR(P) with period s

$$X_t = \Phi_1 X_{t-s} + e_t \qquad \text{sAR(P=1)}$$

$$X_t = \Phi_1 X_{t-s} + \Phi_2 X_{t-2s} + e_t \qquad \text{sAR(P=2)}$$

sAR(1) period 12

$$X_t = \Phi_1 X_{t-12} + e_t$$

$$\begin{aligned}
\gamma(0) &= E(X_t, X_t) = E[X_t \cdot (\Phi_1 X_{t-12} + e_t)] \\
&= \Phi_1 \gamma(12) + E(X_t, e_t) \\
&= \Phi_1 \gamma(12) + E[(\Phi_1 X_{t-12} + e_t) \cdot e_t] \\
&= \Phi_1 \gamma(12) + \sigma^2
\end{aligned}$$

$k \geq 1$,

$$\begin{aligned}\gamma(k) &= E(X_{t-k}\, X_t) = E[X_{t-k}\, (\Phi_1 X_{t-12} + e_t)] \\ &= \Phi_1 \gamma(k-12) + 0\end{aligned}$$

$$\begin{cases}\gamma(0) = \Phi_1 \gamma(12) + \sigma^2 & \qquad\qquad - \quad (2) \\ \gamma(k) = \Phi_1 \gamma(k-12) \quad \text{if } k \geq 1 & \qquad\qquad - \quad (1)\end{cases}$$

- Take Eqn (1), $\gamma(k) = \Phi_1 \gamma(k-12)$. Letting $k = 1$, we get

$$\gamma(1) = \Phi_1 \gamma(-11) = \Phi_1 \gamma(11).$$

- Take (1), $k = 11$

$$\gamma(11) = \Phi_1 \gamma(-1) = \Phi_1 \gamma(1).$$

which can only happen if

$$\gamma(1) = 0, \text{ and } \gamma(11) = 0.$$

(1) says $\gamma(k) = \Phi_1\gamma(k-12)$. Repeating for other k,

$$
\begin{aligned}
\gamma(2) &= \Phi_1\gamma(10) = \Phi_1^2\gamma(2) &\Rightarrow& \quad \gamma(2) = 0, \gamma(10) = 0 \\
\gamma(3) &= \Phi_1\gamma(9) = \Phi_1^2\gamma(3) &\Rightarrow& \quad \gamma(3) = 0, \gamma(9) = 0 \\
&\vdots \\
\gamma(5) &= \Phi_1\gamma(7) = \Phi_1^2\gamma(5) &\Rightarrow& \quad \gamma(5) = 0, \gamma(7) = 0 \\
\gamma(6) &= \Phi_1\gamma(6) = \Phi_1^2\gamma(6) &\Rightarrow& \quad \gamma(6) = 0, \gamma(6) = 0
\end{aligned}
$$

Now take Eqn (2), and Eqn (1) with $k = 12$,

$$\begin{aligned}\gamma(0) &= \Phi_1 \gamma(12) + \sigma^2, \\ \gamma(12) &= \Phi_1 \gamma(0)\end{aligned}$$

Substituting in,

$$\begin{aligned}\gamma(0) &= \Phi_1 \Big(\Phi_1 \gamma(0) \Big) + \sigma^2 \\ &= \Phi_1^2 \gamma(0) + \sigma^2\end{aligned}$$

Solving, we get

$$\gamma(0) = \frac{\sigma^2}{1 - \Phi_1^2}, \qquad \gamma(12) = \Phi_1 \gamma(0).$$

ACVF for sAR(1) with s=12

$$\begin{cases} \gamma(0) = \frac{\sigma^2}{1-\Phi_1^2} \\ \gamma(k12) = \Phi_1^k \gamma(0) \\ \text{ otherwise } \gamma(h) = 0 \end{cases}$$

2.2 sARIMA(p,d,q) x (P,D,Q)[s] model

[ToC]

- Consider sMA(1) with $s = 12$ with innovation ε_t, where $\varepsilon_t \sim$ MA(1), i.e.

$$\begin{aligned} X_t &= \varepsilon_t - \Theta_1 \varepsilon_{t-12} \\ \varepsilon_t &= e_t - \theta_1 e_{t-1} \end{aligned}$$

- That means we have

$$\begin{aligned} X_t &= (1 - \Theta_1 B^{12})\varepsilon_t \\ &= \underbrace{(1 - \Theta_1 B^{12})(1 - \theta_1 B)}_{\text{characteristic polynomial}} e_t \end{aligned}$$

- This is called ARMA$(0,1) \times (0,1)_{12}$ model.

$\mathbf{ARMA}(p,q) \times (P,Q)_s$

- If $P = 1, Q = 1$

$$(1 - \Phi_1 B^s)\tilde{Y}_t = (1 - \Theta_1 B^s)\tilde{e}_t$$

where $\tilde{Y}_t$ and $\tilde{e}_t$ is

$$(1 - \Phi_1 B^s)\,\Phi(B)Y_t = (1 - \Theta_1 B^s)\,\Theta(B)e_t$$

Causality and Invertibility

Causality :

$$(1 - \Phi_1 z^s)(1 - \phi_1 z - \cdots - \phi_p z^p) = 0$$

must have all the roots outside of the unit circle.

$$\begin{aligned}(1 - \Phi_1 z^s) &= 0 \\ (1 - \phi_1 z - \cdots - \phi_p z^p) &= 0\end{aligned}$$

Similar in invertibility.

Seasonal ARIMA model

$$X_t \quad \sim \quad \text{ARIMA}(p,d,q) \times (P,D,Q)_s$$

means

$$\nabla^d \nabla_s^D X_t = Y_t \quad \sim \quad \text{ARMA}(p,q) \times (P,Q)_s$$

$Y_t \sim \text{ARMA}(p,q) \times (1,1)_{12}$ means

$$\begin{aligned}\left(1 - \phi_1 B - \cdots - \phi_p B^p\right)\left(1 - \Phi_1 B^{12}\right) Y_t \\ = \left(1 - \theta_1 B - \cdots - \theta_q B^q\right)\left(1 - \Theta_1 B^{12}\right) e_t\end{aligned}$$

Theoretical ACF

Suppose you have ARMA$(1,1)\times(0,1)_{12}$ model

$$\big(1-\phi_1 B\big)Y_t = \big(1-\theta_1 B\big)\big(1-\Theta_1 B^{12}\big)e_t$$

Then this is same as

$$\big(1-\phi_1 B\big)Y_t = \big(1-\theta_1 B-\Theta_1 B^{12}+\theta_1\Theta_1 B^{13}\big)e_t$$

$$\big(1-.7B\big)Y_t = \big(1+.7B+.5B^{12}+(.7)(.5)B^{13}\big)e_t$$

```
rho <- ARMAacf(ar=c(.7), ma=c( .7,0,0,0,0,0,0,0,0,0,0,.5,.35), lag.max=40)
plot(rho, col="red")
```

2.3 Example: CO2 data

[ToC]

Monthly CO2 levels at NW territories CANADA from Jan 1959 through Dec 1997.

```
data(co2)  #- No need to load package for this data

plot(co2, type="o", ylab='CO2')
acf(co2,lag.max=36)

library(forecast)
source("http://gozips.uakron.edu/~nmimoto/477/TS_R-90.txt")

#- take Del -
layout(matrix(c(1,1,2,3), 2, 2, byrow=T))
plot(diff(co2),ylab='First Difference of CO2',xlab='Time')
acf(diff(co2))
pacf(diff(co2))


#- take Del_12 -
layout(matrix(c(1,1,2,3), 2, 2, byrow=T))
plot(diff(co2, 12),ylab='First Difference of CO2',xlab='Time')
acf(diff(co2, 12))
```

```
pacf(diff(co2, 12))


#- Take Del and Del_12 -
layout(matrix(c(1,1,2,3), 2, 2, byrow=T))
plot( diff(diff(co2, 12)) )
acf(  diff(diff(co2, 12)), lag.max=50)
pacf( diff(diff(co2, 12)), lag.max=50)




Fit1 = Arima(co2, order=c(0,1,1), seasonal=list(order=c(0,1,1), period=12))
Fit1
Randomness.tests(Fit1$residuals)



Fit2 = auto.arima(co2)
Fit2
Randomness.tests(Fit2$residuals)

#- Fit2 have some non-significant parameters.  Fit1 is better model.
```

```
#--- New co2 data (inside TSA package) ---

  acf1 <- acf
  library(TSA)
  acf <- acf1

  data(co2)

  plot(co2,ylab='CO2')
```

3

Effect of Seasnal Differencing on Trend

[ToC]

3.1 Effect of $\bigtriangledown_s$ on Deterministic Trend

[ToC]

- Suppose $s = 12$, and your observation Y_t has linear trend, as well as the seasonality with period $s = 12$.

$$Y_t = a + bt + S_t + X_t.$$

Then

$$\begin{aligned}
\bigtriangledown_s Y_t &= Y_{t+12} - Y_t &= a + b(t+12) + S_{t+12} + X_{t+12} \\
& & -(a + b(t) \quad + S_t \quad + X_t) \\
&= b(12) + X_{t+12} - X_t
\end{aligned}$$

This is statonary series with constant $b(12)$. (called `drift` by `auto.arima`)

- If your observation Y_t has quadratic trend, as well as the seasonality.

$$Y_t = a + bt + ct^2 + S_t + X_t.$$

$$\begin{aligned} Y_{t+12} - Y_t &= a + b(t+12) + c(t+12)^2 + S_{t+12} + X_{t+12} \\ &\quad -(a + b(t) + c(t^2) \qquad\qquad + S_t \quad + X_t) \\ &= \qquad b(12) + c(24t) + c(12^2) + X_{t+12} - X_t \end{aligned}$$

- This is statonary series with linear trend $b(12) + c(12^2) + c(24)t$.
- If you take $\bigtriangledown$ again, who will be left?

Linear Trend plus WN

[ToC]

- Suppose $s = 12$, and your observation Y_t has linear trend,

$$Y_t = a + bt + e_t \qquad \text{where } e_t \sim N(0,1)$$

- Taking $\triangledown_{12}$ yields,

$$\begin{aligned} Y_{t+12} - Y_t &= a + b(t+12) + S_{t+12} + e_{t+12} \\ &\quad -(a + b(t) \qquad\quad + S_t \quad + e_t) \\ &= \qquad b(12) + e_{t+12} - e_t \end{aligned}$$

- What kind of process is this?

$$K_t = e_t - e_{t-12}$$

- What is the better approach to take?

Example

```
source("http://gozips.uakron.edu/~nmimoto/477/TS_R-90.txt")

#--- Linear trend only
  X <- 2+5*(1:200); plot(X, type="o")
  plot(diff(X, 12), type="o")

#--- Linear trend + WN
  X <- 2+.5*(1:200) + rnorm(200,0,5)
  plot(X, type="o")
  plot(diff(X, 12), type="o")

  auto.arima(diff(X, 12))    #- they won't pick up the seasonal term

  acf(diff(X, 12))


  X <- ts(2+.5*(1:200) + rnorm(200,0,5), start=c(1,1), freq=12)   # <- set freq=12
  plot(X, type="o")
  plot(diff(X, 12), type="o")
  Fit1 <- auto.arima(diff(X, 12))    #- Now they do pick up the seasonal
  Fit1

  Randomness.tests(Fit1$resid)
```

```
  Arima(diff(X, 12), order=c(0,0,0), seasonal=c(0,0,1))
```

```
#--- Quadratic trend only
  X <- 2+3*(1:200)^2; plot(X, type="o")
  plot(diff(X, 12), type="o")
  plot(diff(diff(X, 12)), type="o")
```

3.2 Effect of $\bigtriangledown_s$ on Random Trend

[ToC]

- Suppose $s = 12$, and your observation Y_t has random walk without drift as trend, together with the seasonality.

$$Y_t = W_t + S_t + X_t.$$

Where

$$W_t = \sum_{i=1}^{t} e_i \qquad e_i \sim_{iid} N(0, \sigma^2).$$

- If you take seasonal difference,

$$\begin{aligned} \bigtriangledown_{12} Y_t &= \bigtriangledown_{12} W_t + \bigtriangledown_{12} S_t + \bigtriangledown_{12} X_t \\ &= \bigtriangledown_{12} W_t \quad + \quad \bigtriangledown_{12} X_t. \end{aligned}$$

- $$
\begin{aligned}
\nabla_{12} W_t &= W_t - W_{t-12} = \sum_{i=0}^{t-1} e_{t-i} - \sum_{i=12}^{t-1} e_{t-i} \\
&= \sum_{i=0}^{11} e_{t-i} \qquad\qquad \sim N(0, 12\sigma^2)
\end{aligned}
$$

  Since each e_i is iid Normal.

- Note that this is MA(12) with unit root. (non-invertible)

$$
\begin{aligned}
\nabla_{12} W_t &= \sum_{i=0}^{11} e_{t-i} \\
\nabla_{12} W_{t-1} &= \sum_{i=1}^{12} e_{t-i}
\end{aligned}
$$

Example

```
source("http://gozips.uakron.edu/~nmimoto/477/TS_R-90.txt")

#--- RW without drift
X <- cumsum(rnorm(200, 0, 1)); plot(X, type="l")
plot(diff(X, 12), type="o")


X1 <- ts(X, start=c(1,1), freq=12)
auto.arima(diff(X1, 12))

acf(diff(X1,12))


#--- RW with drift
X <- ts(cumsum(rnorm(200, .5, 1)), start=c(1,1), freq=12)
plot(X, type="l")
plot(diff(X, 12), type="o")

auto.arima(diff(X, 12) )
                                   #- picks up the drift (.5)*(12)

Arima(diff(X, 12),  )
```

3.3 Example: CO2 data

[ToC]

Monthly CO2 levels at NW territories CANADA from Jan 1959 through Dec 1997.

```
data(co2)  #- No need to load package for this data
source("http://gozips.uakron.edu/~nmimoto/477/TS_R-90.txt")

plot(co2, type="o", ylab='CO2')
acf(co2,lag.max=36)
pacf(co2,lag.max=36)


plot(diff(co2),ylab='First Difference of CO2',xlab='Time')   #- take Del -

plot(diff(co2, 12),ylab='First Difference of CO2',xlab='Time')  #- take Del_12 -


#- Take Del and Del_12 -
layout(matrix(c(1,1,2,3), 2, 2, byrow=T))
plot( diff(diff(co2, 12)) )
acf(  diff(diff(co2, 12)), lag.max=50)
pacf( diff(diff(co2, 12)), lag.max=50)
```

```
  Fit1 = auto.arima(co2);    Fit1
                                                  #- Fit1 have some non-significant parameters.

  Fit2 = Arima(co2, order=c(1,1,1), seasonal=list(order=c(1,1,2), period=12)); Fit2



  Randomness.tests(Fit2$resid)             #-$-








#--- New co2 data (inside TSA package) ---

  acf1 <- acf
  library(TSA)
  acf <- acf1

  data(co2)

  plot(co2,ylab='New CO2 data')
```

sARIMA$(p, d, q) \times (P, D, Q)_s$ for CO2 data

- $s = 12$ because `freq=12` was set in ts object.
- Took $\bigtriangledown_{12}$ then $\bigtriangledown$.

$$\bigtriangledown \bigtriangledown_{12} Y_t = X_t$$

- Checked stationarity of $\bigtriangledown \bigtriangledown_{12} Y_t$.
- Modeled X_t with MA(1) and sMA(1).

$$X_t = (1 - \theta_1)(1 - \Theta_1 B^{12})\, e_t$$

where $\hat{\theta}_1 = .35$, $\hat{\Theta}_1 = .85$ and

$$e_t \sim N(0, \hat{\sigma} = .29)$$

4

Forecasting sARMA

[ToC]

4.1 Forecsting

Suppose we have ARIMA$(1,0,1) \times (1,1,0)_{12}$ model,

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

This is same as modeling $\bigtriangledown_{12} Y_t$ with ARMA(13,1),

$$\begin{aligned} X_t &= \bigtriangledown_{12} Y_t \\ \left(1-\phi_1 B-\Phi_1 B^{12}+\phi_1\Phi_1 B^{13}\right) X_t &= (1-\theta_1 B)e_t \end{aligned}$$

We know how to get a prediction $\hat{X}(h)$ for ARMA(p, q), Our predictor for Y_t will be

$$\begin{aligned}
Y_{n+1} &= Y_{n+1-12} + \hat{X}(1) \\
Y_{n+2} &= Y_{n+1-12} + \hat{X}(2) \\
&\vdots
\end{aligned}$$

4.2 Tests for Seasonality

[ToC]

How do you check if you need to take a seasonal difference, instead of regular difference?

1. OCSB test (H_0: Seasonal Unit Root Exists)
   - Default method in `auto.arima()`
   - Osborn-Chui-Smith-Birchenhall (1988)

2. CH test (H_0: Deterministic Seasonality Exists)
   - Canova-Hansen (1995)

http://robjhyndman.com/hyndsight/forecast3/

OCSB test

Osborn-Chui-Smith-Birchenhall (1988) test for

- Default choice in `auto.arima()` for choosing value of D.

$$\begin{cases} H_0: \text{ Seasonal Unit Root Exists} \\ H_A: H_0 \text{ is false} \end{cases}$$

- Small p-value means ”Take Seasonal Difference”.

```
data(co2) #- No need to load package for this data
source("http://gozips.uakron.edu/~nmimoto/477/TS_R-90.txt")
library(forecast)
```

```
plot(co2, type="o")
nsdiffs(X, m=12, test="ocsb")
```

```
auto.arima(co2)

Fit1 <- Arima(co2, order=c(0,1,1), seasonal=c(0,1,1))
Fit1

plot(forecast(Fit1))
forecast(Fit1)


plot(diff(co2, 12))
Stationarity.tests(diff(co2, 12))

auto.arima(co2, d=0)

Fit2 <- Arima(co2, order=c(1,0,0), seasonal=c(2,1,1))
```