Q1 ST3010 ASSIGNMENT 2 DAVID WETTBRECHT 12300644

1A ARIMA $(0,0,0)(0,0,1)_4$

Here we have a seasonal arima with $S=4$
Moving Average Model with $Q=1$

Usual assumptions regarding error terms:

$E[\varepsilon_t]=0$ $E[\varepsilon_t^2]=\sigma^2$ and $E[\varepsilon_{t_1}\varepsilon_{t_2}]=0$ if $t_1 \neq t_2$

- PACF shape: will be zero for lags which are not multiples of 4
- At lag 0 value will be 1
- For multiples of 4 $(4, 8, 12, 16 \ldots)$ will have values which create a shape similar to an exponential decrease or damped sine wave shape

proof: rewrite eq^n

$$y_t = c + \alpha\varepsilon_{t-4} + \varepsilon_t \qquad \varepsilon_{t-4} = y_{t-4} - c - \alpha\varepsilon_{t-8} \Rightarrow \text{Sub in}$$
$$= c + \alpha(y_{t-4} - c - \alpha\varepsilon_{t-8}) + \varepsilon_t \qquad \Rightarrow \text{multiply out}$$
$$= c + \alpha y_{t-4} - c\alpha - \alpha^2\varepsilon_{t-8} + \varepsilon_t \qquad \varepsilon_{t-8} = y_{t-8} - c - \alpha\varepsilon_{t-12}$$
$$= (1-\alpha+\alpha^2)c + \alpha y_{t-4} - \alpha^2 y_{t-8} + \alpha^3\varepsilon_{t-12} + \varepsilon_t$$
$$= (1-\alpha+\alpha^2-\alpha^3)c + \alpha y_{t-4} - \alpha^2 y_{t-8} + \alpha^3 y_{t-12} - \alpha^4\varepsilon_{t-16} + \varepsilon_t$$

We can keep expanding the eq^n in similar manner
α will be the PACF value at lag 4, $-\alpha^2$ is value at lag 8 and
α^3 is value at lag 12 and so on
Since $|\alpha|<1$ this ensures an exponential decrease or a
dampened sine wave if α is negative

1 B i. Expectation $E[y_t] = E[c+\alpha\varepsilon_{t-4}+\varepsilon_t]$

expectation is a linear operator $= E[c] + E[\alpha\varepsilon_{t-4}] + E[\varepsilon_t]$
$$= c + \alpha \underbrace{E[\varepsilon_{t-4}]}_{0} + \underbrace{E[\varepsilon_t]}_{0} = c$$

$$E[y_t] = c$$

ii. Variance $V[y_t] = E[(y_t - E[y_t])^2]$
$$= E[(c+\alpha\varepsilon_{t-4}+\varepsilon_t - c)^2]$$
$$= E[\alpha^2\varepsilon_{t-4}^2 + \varepsilon_t^2 + 2\alpha\varepsilon_{t-4}\varepsilon_t]$$

$$= \alpha^2 E[\varepsilon_{t-4}^2] + E[\varepsilon_t^2] + 2\alpha E[\varepsilon_{t-4}\varepsilon_t]$$
$$= \alpha^2\sigma^2 + \sigma^2 + 0$$
$$= (1+\alpha^2)\sigma^2 = Var[y_t]$$

Q1 DAVID WEITBRECHT 12300644

1 Biii $Cov[y_t, y_{t-k}] = \mathbb{E}[(y_t - \mathbb{E}[y_t])(y_{t-k} - \mathbb{E}[y_{t-k}])]$

$\mathbb{E}[y_t] = c \quad \mathbb{E}[y_{t-k}]$ also $= c$

$= \mathbb{E}[(c + \alpha\varepsilon_{t-4} + \varepsilon_t - c)(c + \alpha\varepsilon_{t-k-4} + \varepsilon_{t-k} - c)]$

$= \mathbb{E}[(\alpha\varepsilon_{t-4} + \varepsilon_t)(\alpha\varepsilon_{t-k-4} + \varepsilon_{t-k})]$

$= \mathbb{E}[\alpha^2\varepsilon_{t-4}\varepsilon_{t-k-4} + \alpha\varepsilon_t\varepsilon_{t-k-4} + \alpha\varepsilon_{t-4}\varepsilon_{t-k} + \varepsilon_t\varepsilon_{t-k}]$

$= \alpha^2\mathbb{E}[\varepsilon_{t-4}\varepsilon_{t-k-4}] + \alpha\mathbb{E}[\varepsilon_t\varepsilon_{t-k-4}] + \alpha\mathbb{E}[\varepsilon_{t-4}\varepsilon_{t-k}] + \mathbb{E}[\varepsilon_t\varepsilon_{t-k}]$

When $k=0$ $Cov = \alpha^2\mathbb{E}[\varepsilon_{t-4}\varepsilon_{t-4}] + \alpha\mathbb{E}[\varepsilon_t\varepsilon_{t-4}] + \alpha\mathbb{E}[\varepsilon_{t-4}\varepsilon_t] + \mathbb{E}[\varepsilon_t\varepsilon_t]$

$= (1+\alpha^2)\sigma^2$

When $k=-4$ $Cov = \alpha^2\mathbb{E}[\varepsilon_{t-4}\varepsilon_t] + \alpha\mathbb{E}[\varepsilon_t\varepsilon_t] + \alpha\mathbb{E}[\varepsilon_{t-4}\varepsilon_t] + \mathbb{E}[\varepsilon_t\varepsilon_{t+4}]$

$= \alpha\sigma^2$

At everyother value $Cov = 0$

1 Biv $Corr[y_t, y_{t-k}] = \dfrac{Cov[y_t, y_{t-k}]}{\sqrt{Var[y_t]\,Var[y_{t-k}]}}$

When $k=0$ $\Rightarrow \dfrac{(1+\alpha^2)\sigma^2}{\sqrt{((1+\alpha^2)\sigma^2)((1+\alpha^2)\sigma^2)}} = 1$

Note. $Var[y_t] = Var[y_{t-k}]$

When $k=-4$ $\dfrac{\alpha\sigma^2}{\sqrt{((1+\alpha^2)\sigma^2)\;((1+\alpha^2)\sigma^2)}} = \dfrac{\alpha\cancel{\sigma^2}}{(1+\alpha^2)\cancel{\sigma^2}} = \dfrac{\alpha}{1+\alpha^2}$ $k=-4$

Otherwise $\dfrac{0}{\sqrt{((1+\alpha^2)\sigma^2)((1+\alpha^2)\sigma^2)}} = 0$

ACF Shape: will show a spike at lag $k=-4$

Q1 DAVID WEITBRECHT 12300644

1 C i. Model $y_{n+k} = c + \alpha_{n+k-4} + \varepsilon_{n+k}$

The term ε_{n+k-4} can be expressed as a weighted sum of the past observations y_{n+k-4}, y_{n+k-8}, y_{n+k-12} etc as shown in 1A earlier. All these observations are available to compute for $k \in (1,2,3,4)$ for ε_{n+k-4}

Forecast is then: $\hat{y}_{n+k} = c + \alpha \varepsilon_{n+k-4}$

95% CI $\Rightarrow$ Forecast $\pm\ 2\sigma$

$= \hat{y}_{n+k} \pm 2\sigma = 95\%$ CI

ii. When $8 < k > 4$ the expectation is utilised as the forecast $\hat{y}_{n+k}$

$E[y_t] = c$ along with the error term

Error term will be $\alpha\, \varepsilon_{n+k-4} + \varepsilon_{n+k}$ which has a variance associated with it of $(1+\alpha^2)\sigma^2$

So the ~~AR~~ 95% CI will be $c \pm 2\sqrt{(1+\alpha^2)\sigma^2}$

iii. Forecast from $k = 9 \ldots 12$ will follow the same procedure as forecasts from $k = 4 \cdot 8$

Thus our 95% CI for $k = 9 \ldots 12$ will be

$c \pm 2\sqrt{(1+\alpha^2)\sigma^2}$

Q2 DAVID WETTBRECHT 12300844

2A. ARIMA $\overset{p\,d\,q}{(1,0,3)}\ \overset{P\,D\,Q\,s}{(1,2,0)_5}$

$(1-\phi_1 B)(1-\beta_1 B^5)(1-B^5)^2 = c + (1-w_1 B - w_2 B^2 - w_3 B^3)(1)(\varepsilon_t)$ = ARIMA $(1,0,3)(1,2,0)_5$

This is the ARIMA $(1,0,3)(1,2,0)_5$ model with backshift operator

B. Without Backshift operator:

Expand $(1-\phi_1 B)(1-\beta_1 B^5)(1-B^5)^2$

$\cdot(1-\beta_1 B^5 - \phi_1 B + \phi_1\beta_1 B^6)(1-2B^5+B^{10})$

$1-\beta_1 B^5 - \phi_1 B + \phi_1\beta_1 B^6 - 2B^5 + 2\beta_1 B^{10} + 2\phi_1 B^6 - 2\phi_1\beta_1 B^{11} + B^{10} - \beta_1 B^{15} - \phi_1 B^{11} + \phi_1\beta_1 B^{16}$

Applied to y_t LHS:

$1-\phi_1 y_t - \beta_1 y_{t-5} + \phi_1\beta_1 y_{t-6} - 2y_{t-5} + 2\phi_1 y_{t-6} + 2\beta_1 y_{t-10} - 2\phi_1\beta_1 y_{t-11} + y_{t-10} - \phi_1 y_{t-11} - \beta_1 y_{t-15} + \phi_1\beta_1 y_{t-16}$

is Left hand Side of ARIMA $(1,0,3)(1,2,0)_5$ model

Right hand Side: $c + (\varepsilon_t - w_1 B\varepsilon_t - w_2 B^2\varepsilon_t - w_3 B^3\varepsilon_t)$

$\Rightarrow c + \varepsilon_t - w_1\varepsilon_{t-1} - w_2\varepsilon_{t-2} - w_3\varepsilon_{t-3}$.

= RHS without backshift operator

1a $y_t = c + \alpha \varepsilon_{t-4} + \varepsilon_t$

PACF ⇒ The equation can be rewrote a

$$\varepsilon_{t-4} = (y_{t-4} + c + \alpha \varepsilon_{t-8})$$

$$y_t = c + y_{t-4} + c + \alpha$$

$$y_t = c + \alpha (y_{t-4} + c + \alpha \varepsilon_{t-8}) + \varepsilon_t$$
$$c + \alpha y_{t-4} + \alpha c + \alpha^2 \varepsilon_{t-8}$$

This can be extended further

$$\varepsilon_{t-8} = y_{t-8} + c + \alpha \varepsilon_{t-12}$$

$$\Rightarrow \quad = c + \alpha y_{t-4} + \alpha c + \alpha^2 (y_{t-8} + c + \alpha \varepsilon_{t-12})$$

can keep expanding

PACF is zero for lags that are not multiple of 4.
At lag 0, PACF is 1, at 4, 8, 12 will have an exponential decay or damped sine wave shape

b $E[y_t] = E[c + \alpha \varepsilon_{t-4} + \varepsilon_t]$

Expectation is a linear operator: Expectation of linear sum is the linear sum of expectations

$$E[c] + E[\alpha \varepsilon_{t-4}] + E[\varepsilon_t]$$

c is a constant so $E[c] = c$

$\alpha E[\varepsilon_{t-4}]$ α is a constant

$E[\varepsilon_{t-4}] = 0$ by our hypothesis that expectation of error is 0.

$E[\varepsilon_t]$ is also $= 0$

So $E[y_t] = c$

B. Var $= \cancel{E[x]^2 - E[x^2]}$

$E[x^2] - E[x]^2$

$\cancel{y_t^2 = (c + \alpha\varepsilon_{t-4} + \varepsilon_t)^2}$

$E[(y_t - E[y_t])^2]$

$E[(c + \alpha\varepsilon_{t-4} + \varepsilon_t - c)^2]$

$E[(\alpha\varepsilon_{t-4} + \varepsilon_t)^2]$

$E[\alpha^2\varepsilon_{t-4}^2 + 2\alpha\varepsilon_t\varepsilon_{t-4} + \varepsilon_t^2]$

$\alpha E[\varepsilon_{t-4}^2] + 2\alpha[\varepsilon_t\varepsilon_{t-4}] + E[\varepsilon_t]$

$= \alpha\sigma^2 + 0 + \sigma^2$

$+ (1+a)\sigma^2$

C. Covariance $[y_t, y_{t-k}]$ $\forall k$

$= E[(y_t - E[y_{t-k}])(y_{t-k} - E[y_t])]$

$E[(c + a\varepsilon_{t-4} + \varepsilon_t - c)(c + \alpha\varepsilon_{t-k} + \varepsilon_t)]$

$[(a\varepsilon_{t-4} + \varepsilon_t)(c + \alpha\varepsilon_{t-k-4} + \varepsilon_{t+k} - c)]$

$(a\varepsilon_{t-4} + \varepsilon_t)(\alpha\varepsilon_{t-k-4} + \varepsilon_{t+k})$

$E\ a^2\varepsilon_{t-4}\varepsilon_{t-k-4} + a\varepsilon_{t-4}\varepsilon_{t-k} + a\varepsilon_t\varepsilon_{t-k-4} + \varepsilon_t\varepsilon_{t+k}$

$\alpha E[\varepsilon_{t-4}\varepsilon_{t-k-4}] + \alpha E[\varepsilon_{t-4}\varepsilon_{t+k}] + \alpha E[\varepsilon_t\varepsilon_{t+k-4}] + E[\varepsilon_t\varepsilon_{t+k}]$

When $k=0$

$\Rightarrow \alpha E[\varepsilon_{t-4}\varepsilon_{t-4}] + \alpha E[\varepsilon_{t-4}\varepsilon_t] + \alpha E[\varepsilon_t\varepsilon_{t-4}] + E[\varepsilon_t\varepsilon_t]$

$\alpha\sigma^2 \quad 0 \quad 0 \quad = \sigma^2$

$k=0 \quad (1+a)\sigma^2$

When $k=4 \Rightarrow \alpha E[\varepsilon_{t-4}\varepsilon_{t-8}] + \alpha E[\varepsilon_{t-4}\varepsilon_{t+4}] + \alpha E[\varepsilon_t\varepsilon_{t+4}] + E[\varepsilon_t\varepsilon_{t+4}]$

$\alpha\sigma^2$

0 otherwise $\Rightarrow$ every other figure will be 0

$Cov(y_t, y_{t-k})$

$\mathbb{E}\left((y_t - \mathbb{E}(y_t))(y_{t-k} - \mathbb{E}(y_{t-k})) \right)$

$\mathbb{E}(y_t) = c \quad \mathbb{E}(y_{t-k}) = c$

$\mathbb{E}\left[(\alpha \varepsilon_{t-4} + \varepsilon_t)(\alpha \varepsilon_{t-k-4} + \varepsilon_{t-k}) \right]$

$\mathbb{E}\left[\alpha^2 \varepsilon_{t-4} \varepsilon_{t-k-4} + \alpha \varepsilon_{t-4} \varepsilon_{t-k} + \alpha \varepsilon_t \varepsilon_{t-k-4} + \varepsilon_t \varepsilon_{t-k} \right]$

$= \alpha^2 \mathbb{E}[\varepsilon_{t-4}\varepsilon_{t-k-4}] + \alpha \underbrace{\mathbb{E}[\varepsilon_{t-4}\varepsilon_{t-k}]}_{0} + \alpha \underbrace{\mathbb{E}[\varepsilon_t \varepsilon_{t-k-4}]}_{0} + \mathbb{E}[\varepsilon_t \varepsilon_{t-k}]$

$k=0 \quad (1+\alpha^2)\,\sigma^2$

at $k=-4 \quad \alpha\sigma^2$

everywhere else it is 0.

Correlation:

$$\frac{Cov}{var(y_t)\, var(y_{t-k})}$$

$$\text{at } k=0. \quad \frac{(1+\alpha^2)\sigma^2}{\sqrt{(1+\alpha^2)\sigma^2}}$$

Correlation of y_t, y_{t-4}

$$= \frac{Cov[y_t, y_{t+k}]}{\sqrt{Var[y_t]\ Var[y_{t+k}]}}$$

$k=0$ $$\frac{(1+\sigma)\sigma^2}{((1+a)\sigma^2)((1+a)\sigma^2))} = \frac{(1+\sigma)\sigma^2}{(1+\sigma)\sigma^2} = 1$$

$k=4$ $$\frac{a\sigma^2}{\sqrt{((1+a)\sigma^2)((1+a)\sigma^2))}} = \frac{a\cancel{\sigma^2}}{(1+a)\cancel{\sigma^2}} = \frac{a}{1+a}$$

$k \neq 0, 4$ $= 0$ otherwise

Given this information, there will only be one spike at lag 4 on the ACF

C $y_t = c + \alpha\, \varepsilon_{t-4} + \varepsilon_t$

$\hat{y}_{n+k}$ and 95% CI fo $k = 1..4$

Eqn for the four values is

$y_{n+k} = c + \alpha\, \varepsilon_{t-4+k} + \varepsilon_{n+k}$

ε_{t+k-4} can be re wrote in terms of a weighted sum of past observations such as y_{n+k-4}, y_{n+k-8} etc

All these past observations are available to compute ε_{n+k-4}

Hence forecasted value is $\hat{y}_{n+n} = c + \alpha\, \varepsilon_{n+k-4}$

with 95% CI using $\pm 2\sigma$

– Beyond

ARIMA(1,0,3)(1,2,0)$_5$:

$$(1-\phi_1 B)(1-\beta_1 B^5)(1)^0(1-B^5)^2 y_t = \cancel{c+}$$
$$= c + (1 - w_1 B - w_2 B^2 - w_3 B^3)(1)(\varepsilon_t)$$

$$(1-\phi_1 B)(1-\beta_1 B^5)(1-B^5)^2 y_t = c + (1 - w_1 B - w_2 B^2 - w_3 B^3)(\varepsilon_t)$$

Expand left hand side

$$(1-\phi_1 B)(1-\beta_1 B^5)(1-2B^5+B^{10})$$

$$(1-\beta_1 B^5 - \phi_1 B + \phi_1 \beta_1 B^6)(1-2B^5+B^{10})$$

$$\underbrace{1-\beta_1 B^5 - \phi B + \phi_1 \beta_1 B^6}\ \underbrace{-2B^5 + 2\beta_1 B^{10} + 2\phi_1 B^6 - 2\phi_1 \beta_1 B^{11}}\ \underbrace{+B^{10} + \beta_1 B^{15} + \phi_1 B^{11} - \phi_1 \beta_1 B^{16}}$$

apply this to y_t

$\Rightarrow$

$$1 - \beta_1 y_{t-5} - \phi y_t + \phi \beta_1 y_{t-6} - 2y_{t-5} + 2\beta_1 y_{t-10} + 2\phi_1 y_{t-6} - 2\phi_1 \beta_1 y_{t-11} + y_{t-10} + \beta_1 y_{t-15}$$
$$+ \phi_1 y_{t-11} - \phi_1 \beta_1 y_{t-16}$$

RHS:

$$c + 1 - w_1 \varepsilon_{t-1} - w_2 \varepsilon_{t-2} - w_3 \varepsilon_{t-3}$$