MONASH University

MONASH BUSINESS SCHOOL

ETF3231/5231 Business forecasting

Week 9: ARIMA models

https://bf.numbat.space/

AACSB ACCREDITED EQUIS ACCREDITED ASSOCIATION OF MBAS AMBA ACCREDITED

SUMMARY

$$\phi(B)(1-B)^d y_t = c + \theta(B) + \varepsilon_t$$

$(p, q) \rightarrow$ short-run

$(c, d) \rightarrow$ long-run

constant

$c = 0$	$d = 0$	$\hat{y}_{T+\infty} \rightarrow 0$	ARMA stat around	$E(y_t) = 0$
$c \neq 0$	$d = 0$	$\hat{y}_{T+\infty} \rightarrow E(y_t)$	" " "	$E(y_t) \neq 0$
$c = 0$	$d = 1$	$\hat{y}_{T+\infty} \rightarrow$ const	RW + ARMA	• diff of RW is stat and the ARMA part will converge to const.

linear trend

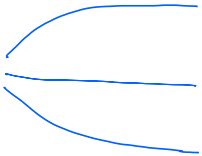
$c \neq 0$	$d = 1$		RW + drift + ARMA
		$\hat{y}_{T+\infty} \rightarrow t$	
$c = 0$	$d = 2$		Two unit roots

quadratic trend $\;[\; c \neq 0 \quad d = 2$

DO IT AT YOUR OWN RISK.

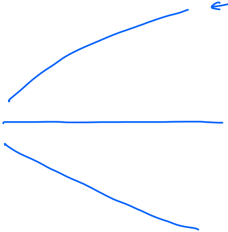
$d = 0$

constant (in sample st. dev).

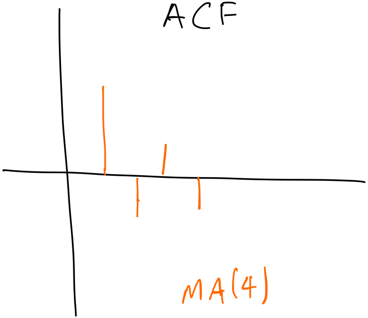


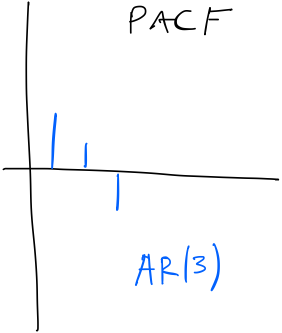
$d = 1$

keep increasing



EXTRA JUST TO RECAP *





* from theory / reading ACF / PACF we can tell either AR or MA orders

Outline

Outline

Modelling procedure with `ARIMA()`

1. Plot the data. Identify any unusual observations. Always
2. If necessary, transform the data (using a Box-Cox transformation) to stabilize the variance. remove heteroscedasticity, make variance stationary
3. If the data are non-stationary: take first differences of the data until the data are stationary. make stationary in the mean
4. Examine the ACF/PACF: Is an AR(p) or MA(q) model appropriate? - select model(s). - Most likely more than one
5. Try your chosen model(s), and use the AICc to search for a better model.
6. Check the residuals from your chosen model by plotting the ACF of the residuals, and doing a portmanteau test of the residuals. If they do not look like white noise, try a modified model. Talk about IA4. Remember models are DGP approximations. You can only do your best (95% v 70% coverage)
7. Once the residuals look like white noise, calculate forecasts.

ARIMA()

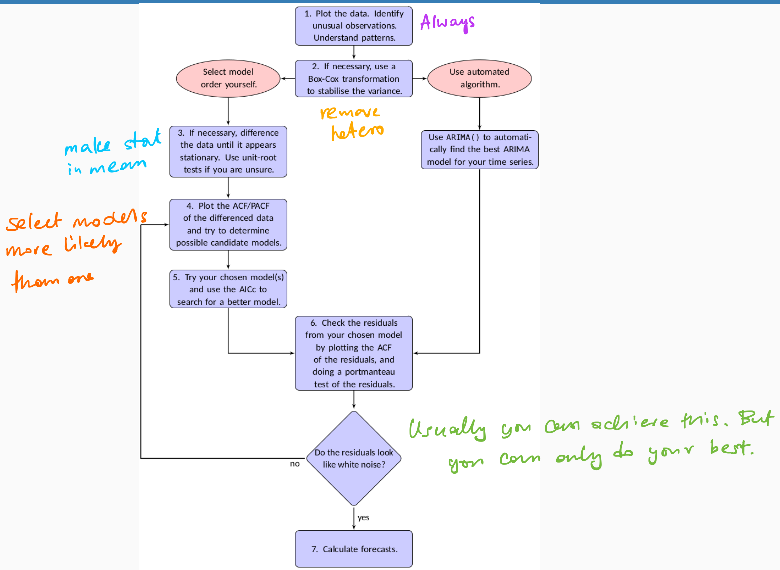
Automatic modelling procedure with `ARIMA()`

1. Plot the data. Identify any unusual observations.
2. If necessary, transform the data (using a Box-Cox transformation) to stabilize the variance.

3. Use `ARIMA()` to automatically select a model.

6. Check the residuals from your chosen model by plotting the ACF of the residuals, and doing a portmanteau test of the residuals. If they do not look like white noise, try a modified model.
7. Once the residuals look like white noise, calculate forecasts.

Modelling procedure



Outline

Point forecasts

1. Rearrange ARIMA equation so y_t is on LHS.
2. Rewrite equation by replacing t by $T + h$.
3. On RHS, replace future observations by their forecasts, future errors by zero, and past errors by corresponding residuals.

Start with $h = 1$. Repeat for $h = 2, 3, \ldots$.

Point forecasts

ARIMA(3,1,1) forecasts: Step 1

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B) y_t = (1 + \theta_1 B)\varepsilon_t,$$

Point forecasts

ARIMA(3,1,1) forecasts: Step 1

$$(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)(1-B)y_t=(1+\theta_1 B)\varepsilon_t,$$

$$\begin{aligned}\left[1-(1+\phi_1)B+(\phi_1-\phi_2)B^2+(\phi_2-\phi_3)B^3+\phi_3 B^4\right]y_t \\ =(1+\theta_1 B)\varepsilon_t,\end{aligned}$$

$$\begin{aligned}y_t-(1+\phi_1)y_{t-1}+(\phi_1-\phi_2)y_{t-2}+(\phi_2-\phi_3)y_{t-3} \\ +\phi_3 y_{t-4}=\varepsilon_t+\theta_1\varepsilon_{t-1}.\end{aligned}$$

$$\begin{aligned}y_t=(1+\phi_1)y_{t-1}-(\phi_1-\phi_2)y_{t-2}-(\phi_2-\phi_3)y_{t-3} \\ -\phi_3 y_{t-4}+\varepsilon_t+\theta_1\varepsilon_{t-1}.\end{aligned}$$

Point forecasts (h=1)

$$\begin{aligned} y_t = (1+\phi_1)y_{t-1} - (\phi_1 - \phi_2)y_{t-2} - (\phi_2 - \phi_3)y_{t-3} \\ - \phi_3 y_{t-4} + \varepsilon_t + \theta_1\varepsilon_{t-1}. \end{aligned}$$

Point forecasts (h=1)

$$y_t = (1+\phi_1)y_{t-1} - (\phi_1 - \phi_2)y_{t-2} - (\phi_2 - \phi_3)y_{t-3} - \phi_3 y_{t-4} + \varepsilon_t + \theta_1\varepsilon_{t-1}.$$

ARIMA(3,1,1) forecasts: Step 2

$E(\varepsilon_{T|T}) = e_T$ $\quad var(\varepsilon_{T|T}) = 0$

$E(\varepsilon_{T+1|T}) = 0$ $\quad var(\varepsilon_{T+1|T}) = \sigma^2$

$$y_{T+1} = (1+\phi_1)y_T - (\phi_1 - \phi_2)y_{T-1} - (\phi_2 - \phi_3)y_{T-2} - \phi_3 y_{T-3} + \varepsilon_{T+1} + \theta_1\varepsilon_T.$$

ARIMA(3,1,1) forecasts: Step 3

$$\hat{y}_{T+1|T} = (1+\phi_1)y_T - (\phi_1 - \phi_2)y_{T-1} - (\phi_2 - \phi_3)y_{T-2} - \phi_3 y_{T-3} + \theta_1 e_T.$$

Point forecasts (h=1)

$$y_t = (1+\phi_1)y_{t-1} - (\phi_1 - \phi_2)y_{t-2} - (\phi_2 - \phi_3)y_{t-3} - \phi_3 y_{t-4} + \varepsilon_t + \theta_1\varepsilon_{t-1}.$$

ARIMA(3,1,1) forecasts: Step 2

$E(\varepsilon_{T|T}) = e_T$ $\quad var(\varepsilon_{T|T}) = 0$

$E(\varepsilon_{T+1|T}) = 0$ $\quad var(\varepsilon_{T+1|T}) = \sigma^2$

$$E(y_{T+1}) = (1+\hat\phi_1)y_T - (\hat\phi_1 - \hat\phi_2)y_{T-1} - (\hat\phi_2 - \hat\phi_3)y_{T-2} - \hat\phi_3 y_{T-3} + E(\varepsilon_{T+1}) + \theta_1 E(\varepsilon_T).$$

ARIMA(3,1,1) forecasts: Step 3

$$\hat{y}_{T+1|T} = (1+\hat\phi_1)y_T - (\hat\phi_1 - \hat\phi_2)y_{T-1} - (\hat\phi_2 - \hat\phi_3)y_{T-2} - \hat\phi_3 y_{T-3} + \hat\theta_1 e_T.$$

Point forecasts (h=2)

$$
\begin{aligned}
y_t = (1+\phi_1)y_{t-1} - (\phi_1 - \phi_2)y_{t-2} - (\phi_2 - \phi_3)y_{t-3} \\
- \phi_3 y_{t-4} + \varepsilon_t + \theta_1 \varepsilon_{t-1}.
\end{aligned}
$$

Point forecasts (h=2)

$$y_t = (1+\phi_1)y_{t-1} - (\phi_1 - \phi_2)y_{t-2} - (\phi_2 - \phi_3)y_{t-3} - \phi_3 y_{t-4} + \varepsilon_t + \theta_1\varepsilon_{t-1}.$$

ARIMA(3,1,1) forecasts: Step 2

$$y_{T+2} = (1+\phi_1)y_{T+1} - (\phi_1 - \phi_2)y_{T} - (\phi_2 - \phi_3)y_{T-1} - \phi_3 y_{T-2} + \varepsilon_{T+2} + \theta_1\varepsilon_{T+1}.$$

ARIMA(3,1,1) forecasts: Step 3

$$\hat{y}_{T+2|T} = (1+\hat{\phi}_1)\hat{y}_{T+1|T} - (\hat{\phi}_1 - \hat{\phi}_2)y_{T} - (\hat{\phi}_2 - \hat{\phi}_3)y_{T-1} - \hat{\phi}_3 y_{T-2}.$$

Prediction intervals

95% prediction interval

$$\hat{y}_{T+h|T} \pm 1.96\sqrt{v_{T+h|T}}$$

where $v_{T+h|T}$ is estimated forecast variance.

Prediction intervals

95% prediction interval

$$\hat{y}_{T+h|T} \pm 1.96\sqrt{v_{T+h|T}}$$

where $v_{T+h|T}$ is estimated forecast variance.

true for any time series model

- $\boxed{v_{T+1|T} = \hat{\sigma}^2}$ for all ARIMA models regardless of parameters and orders.
- Multi-step prediction intervals for ARIMA(0,0,q):

$$y_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}.$$

very easy to do (P.T.O.)

$$v_{T|T+h} = \hat{\sigma}^2 \left[1 + \sum_{i=1}^{h-1} \theta_i^2 \right], \qquad \text{for } h = 2, 3, \ldots.$$

* AR(1) → MA(∞)

* more complex beyond our scope

$MA(q)$ $\quad y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$

- $\operatorname{var}(\theta \varepsilon_t) = \theta^2 \operatorname{var}(\varepsilon_t) = \theta^2 \sigma^2$ for $t > T$

$$y_{T+1} = c + \varepsilon_{T+1} + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \dots + \theta_q \varepsilon_{T-q+1}$$

$$\operatorname{var}(y_{T+1|T}) = \operatorname{var}(c + \varepsilon_{T+1} + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1} + \dots + \theta_q \varepsilon_{T-q+1})$$

$$= \operatorname{var}(\varepsilon_{T+1|T}) = \sigma^2$$

(Everything else has been observed hence $\operatorname{var}(\varepsilon_t) = 0$ $\quad t \leq T$)

$$\operatorname{var}(y_{T+2|T}) = \operatorname{var}(c + \varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} + \theta_2 \varepsilon_T + \theta_3 \varepsilon_{T-1} + \dots + \theta_q \varepsilon_{T-q+2})$$

$$= \operatorname{var}(\varepsilon_{T+2|T}) + \theta_1^2 \operatorname{var}(\varepsilon_{T+1|T})$$

$$= \sigma^2 + \theta_1^2 \sigma^2 = (1 + \theta_1^2)\sigma^2$$

and so on .... $\quad (1 + \theta_1^2 + \theta_2^2)\sigma^2 \dots$

Prediction intervals

* remember the role of d ?

- Prediction intervals increase in size with forecast horizon.
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are uncorrelated and normally distributed.

Prediction intervals

* remember the role of d

- Prediction intervals increase in size with forecast horizon.
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are uncorrelated and normally distributed.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors.

+ correct model to start with

Outline

- not much more to learn.
- Basically implement you knowledge to now include the seasonal frequency.

Seasonal ARIMA models

ARIMA	$\underbrace{(p,d,q)}$	$\underbrace{(P,D,Q)_m}$
	$\uparrow$ Non-seasonal part of the model	$\uparrow$ Seasonal part of of the model

• annual seas. a problem for
- weekly
- daily

- monthly
- quarterly

where m = number of observations per year.

m = number of observations per week - daily (exam 2022) (Hint)

Seasonal ARIMA models

QUARTERLY EXAMPLE

E.g., ARIMA(1, 1, 1)(1, 1, 1)$_4$ model (without constant)

$$\underbrace{(1-\phi_1 B)}_{\text{Non-seasonal AR(1)}}\underbrace{(1-\Phi_1 B^4)}_{\text{Seasonal AR(1)}}\underbrace{(1-B)}_{\text{Non-seasonal difference}}\underbrace{(1-B^4)}_{\text{Seasonal difference}}y_t = \underbrace{(1+\theta_1 B)}_{\text{Non-seasonal MA(1)}}\underbrace{(1+\Theta_1 B^4)}_{\text{Seasonal MA(1)}}\varepsilon_t.$$

Seasonal ARIMA models

E.g., ARIMA(1, 1, 1)(1, 1, 1)$_4$ model (without constant)

$$(1-\phi_1 B)(1-\Phi_1 B^4)(1-B)(1-B^4)y_t = (1+\theta_1 B)(1+\Theta_1 B^4)\varepsilon_t.$$

All the factors can be multiplied out and the general model written as follows:

$$\begin{aligned} y_t = {} & (1+\phi_1)y_{t-1} - \phi_1 y_{t-2} + (1+\Phi_1)y_{t-4} \\ & - (1+\phi_1+\Phi_1+\phi_1\Phi_1)y_{t-5} + (\phi_1+\phi_1\Phi_1)y_{t-6} \\ & - \Phi_1 y_{t-8} + (\Phi_1+\phi_1\Phi_1)y_{t-9} - \phi_1\Phi_1 y_{t-10} \\ & + \varepsilon_t + \theta_1\varepsilon_{t-1} + \Theta_1\varepsilon_{t-4} + \theta_1\Theta_1\varepsilon_{t-5}. \end{aligned}$$

* very messy
* unintuitive

Seasonal ARIMA models

The seasonal part of an AR or MA model will be seen in the seasonal lags of the PACF and ACF.

ARIMA(0,0,0)(0,0,1)$_{12}$ will show:

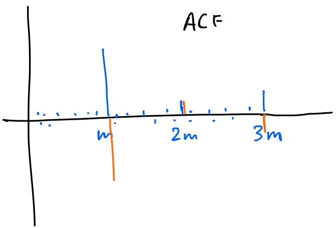
- a spike at lag 12 in the ACF but no other significant spikes.
- The PACF will show exponential decay in the seasonal lags; that is, at lags 12, 24, 36, ...

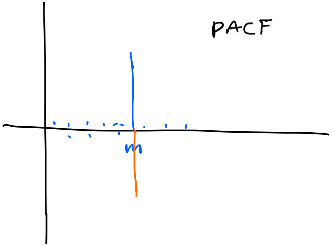
* not many more than 1 or 2.

ARIMA(0,0,0)(1,0,0)$_{12}$ will show:

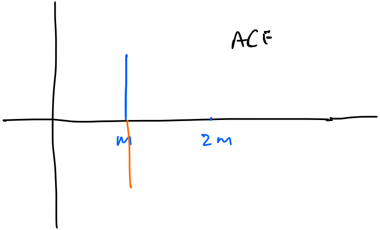
- exponential decay in the seasonal lags of the ACF
- a single significant spike at lag 12 in the PACF.

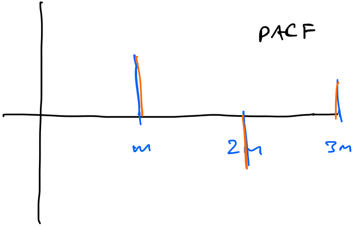
Seasonal AR(1)





Seasonal MA(1)



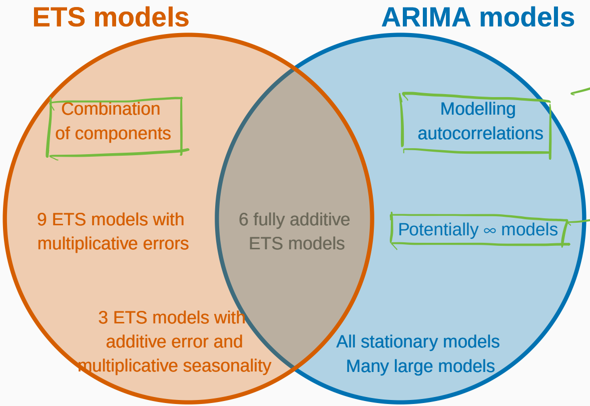


Outline

ARIMA vs ETS

- Myth that ARIMA models are more general than exponential smoothing.
- Linear exponential smoothing models all special cases of ARIMA models.
- Non-linear exponential smoothing models have no equivalent ARIMA counterparts.
- Many ARIMA models have no exponential smoothing counterparts.
- ETS models are all non-stationary. Models with seasonality or non-damped trend (or both) have two unit roots; all other models have one unit root.

ARIMA vs ETS



A different philosophy

but you don't really explore these.

Equivalences

ETS model	ARIMA model		Parameters
ETS(A,N,N)	ARIMA(0,1,1)	very easy to show	$\theta_1 = \alpha - 1$
ETS(A,A,N)	ARIMA(0,2,2)		$\theta_1 = \alpha + \beta - 2$
			$\theta_2 = 1 - \alpha$
ETS(A,A$_d$,N)	ARIMA(1,1,2)		$\phi_1 = \phi$
			$\theta_1 = \alpha + \phi\beta - 1 - \phi$
			$\theta_2 = (1 - \alpha)\phi$
ETS(A,N,A)	ARIMA$(0,0,m)(0,1,0)_m$		
ETS(A,A,A)	ARIMA$(0,1,m+1)(0,1,0)_m$		These are more complex
ETS(A,A$_d$,A)	ARIMA$(1,0,m+1)(0,1,0)_m$		

ETS (A,N,N) as ARIMA (0,1,1) $\theta_1 = \alpha - 1$

$y_t = l_{t-1} + \varepsilon_t$

$l_t = l_{t-1} + \alpha \varepsilon_t$

Diff obs equation $y_t - y_{t-1} = l_{t-1} - l_{t-2} + \varepsilon_t - \varepsilon_{t-1}$

from level equation $l_{t-1} = l_{t-2} + \alpha \varepsilon_{t-1} \Rightarrow l_{t-1} - l_{t-2} = \alpha \varepsilon_{t-1}$

$\Rightarrow \quad y_t - y_{t-1} = \alpha \varepsilon_{t-1} + \varepsilon_t - \varepsilon_{t-1}$

$\Rightarrow \quad y_t = y_{t-1} + \varepsilon_t + (\alpha - 1) \varepsilon_{t-1}$

$\Rightarrow \quad y_t = y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ where $\theta_1 = \alpha - 1$