FORECASTING
STAT0010

Alexandros Beskos

a.beskos@ucl.ac.uk

'Lecture 5' Outline

1. Non-stationary processes
2. ARIMA models
3. Non-zero means
4. Seasonal models
5. Multiplicative seasonal ARMA models

Example 1

If an ARMA process $\{Y_t\}$ *has* d*-many AR roots* $= 1$*, (and all other roots outside unit circle) then*

$$\phi(B)(1-B)^d Y_t = \theta(B)\epsilon_t \,.$$

Now define $Z_t := (1-B)^d Y_t = \nabla^d Y_t$ *(i.e. take* d*th difference of* Y_t*). Then* Z_t *is stationary.*

Remark 2 (In practice, if a process looks non-stationary...)

- *1st differencing often adequate to make a series stationary*
- *sometimes, 2nd differencing is required*
- *don't take too many differences!*

CAVEAT! Over-differencing will cause:

- an increase in variance
- an increase in the order of MA — any ACF cut-off lag will increase
- non-invertibility (MA poly. has unit root)

We can now extend our interest from stationary *ARMA* models to any non-stationary model that can be 'transformed' into a stationary *ARMA* model by differencing.

Definition 3 (Integrated autoregressive, moving average $ARIMA(p, d, q)$ process)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0, \sigma^2)$ and let the process $\{(1 - B)^d Y_t\}$ be an $ARMA(p, q)$ process of order (p, q). Then, $\{Y_t\}$ is an integrated autoregressive, moving average process of order (p, d, q), written $ARIMA(p, d, q)$, with model equation:

$$\phi(B)(1 - B)^d Y_t = \theta(B)\epsilon_t \,.$$

Remark 4

If $\{Y_t\}$ has to be differenced d-many times before it is a stationary $ARMA(p, q)$ process, then $\{Y_t\}$ is an $ARIMA(p, d, q)$ process.

Example 5 ($ARIMA(0, 1, 0)$ or $I(1)$)

Recall random walk: $\nabla Y_t = \epsilon_t$, with $\epsilon_t \sim \mathcal{WN}(0, \sigma^2)$.

Example 6 ($ARIMA(0,1,1)$ or $IMA(1,1)$)

Find autocorrelation of Y_t, where $(1-B)Y_t = (1-\theta_1 B)\epsilon_t$. I.e.

$$Y_t = Y_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}\,.$$

Assuming $Y_0 = 0$,

$$\begin{aligned}
Y_1 &= \epsilon_1 - \theta_1\,\epsilon_0 \\
Y_2 &= Y_1 + \epsilon_2 - \theta_1\,\epsilon_1 = \epsilon_2 - \theta_1\,\epsilon_0 + (1-\theta_1)\,\epsilon_1 \\
Y_3 &= Y_2 + \epsilon_3 - \theta_1\,\epsilon_2 = \epsilon_3 - \theta_1\,\epsilon_0 + (1-\theta_1)\,\epsilon_2 + (1-\theta_1)\,\epsilon_1 \\
&\vdots \\
Y_t &= \epsilon_t - \theta_1\epsilon_0 + (1-\theta_1)\sum_{j=1}^{t-1}\epsilon_j
\end{aligned}$$

Now $\mathbb{E}(Y_t) = 0$. Hence $\mathrm{cov}(Y_t, Y_{t+k}) = \mathbb{E}(Y_t Y_{t+k}) =$

$$\mathbb{E}\left(\left(\epsilon_t - \theta_1\epsilon_0 + (1-\theta_1)\sum_{j=1}^{t-1}\epsilon_j\right)\left(\epsilon_{t+k} - \theta_1\epsilon_0 + (1-\theta_1)\sum_{j=1}^{t+k-1}\epsilon_j\right)\right)$$

For $k \geq 0$:

$$\mathbb{E}\left(\left(\epsilon_t - \theta_1\epsilon_0 + (1-\theta_1)\ \sum_{j=1}^{t-1}\epsilon_j\right)\left(\epsilon_{t+k} - \theta_1\epsilon_0 + (1-\theta_1)\ \sum_{j=1}^{t+k-1}\epsilon_j\right)\right)$$

$$= \mathbb{E}(\epsilon_t\epsilon_{t+k}) + \theta_1^2\mathbb{E}(\epsilon_0^2) + (1-\theta_1)^2\sum_{j=1}^{t-1}\sum_{\ell=1}^{t+k-1}\mathbb{E}(\epsilon_j\epsilon_\ell) + (1-\theta_1)\mathbb{E}\left(\epsilon_t\sum_{\ell=1}^{t+k-1}\epsilon_\ell\right)$$

$$= \sigma^2\left(\delta_{0,k} + \theta_1^2 + (1-\theta_1)^2\sum_{j=1}^{t-1}1\right) + (1-\theta_1)\mathbb{E}\left(\epsilon_t\sum_{\ell=1}^{t+k-1}\epsilon_\ell\right)$$

$$= \sigma^2(\delta_{0,k} + \theta_1^2 + (1-\theta_1)^2(t-1)) + (1-\theta_1)\mathbb{E}\left(\epsilon_t\sum_{\ell=1}^{t+k-1}\epsilon_\ell\right)$$

$$= \begin{cases} \sigma^2\left(1 + \theta_1^2 + (1-\theta_1)^2(t-1)\right), & k = 0 \\ \sigma^2\left(\theta_1^2 + (1-\theta_1)^2(t-1) + 1 - \theta_1\right), & k > 0 \end{cases}$$

I.e., for $k > 0$:

$$\mathrm{cov}(Y_t, Y_{t+k}) = \mathrm{var}(Y_t) - \theta_1\sigma^2\,.$$

$\mathrm{cov}(Y_t, Y_{t+k}) = \mathrm{var}(Y_t) - \theta_1\sigma^2$. Now,

$$\mathrm{corr}(Y_t, Y_{t+k}) = \frac{\mathrm{cov}(Y_t, Y_{t+k})}{\sqrt{\mathrm{var}(Y_t)\,\mathrm{var}(Y_{t+k})}} = \frac{\mathrm{var}(Y_t) - \theta_1\sigma^2}{\sqrt{\mathrm{var}(Y_t)\,\mathrm{var}(Y_{t+k})}}.$$

Remark 7

Note

- $\mathrm{var}(Y_t)$ *is a function of* $t \Rightarrow \{Y_t\}$ *is non-stationary.*
- *for* t *large and* k *small,* $\mathrm{corr}(Y_t, Y_{t+k}) \approx 1$. *(ACF very slow decay)*

Example 8 ($ARIMA(1,1,0)$ or $ARI(1,1)$)

Find autocorrelation of Y_t*, where* $(1-\phi_1 B)(1-B)Y_t = \epsilon_t$. *See notes (page31) — more straightforward than previous example(!).*

Remark 9 (non-zero means)

Recall a stationary $ARMA(p,q)$ process given by

$$Y_t = \sum_{j=1}^{p} \phi_j Y_{t-j} + \epsilon_t - \sum_{j=1}^{q} \theta_j \epsilon_{t-j}\,,$$

implies $\mathbb{E}(Y_t) = 0$.

Take expectations of both sides: $\mathbb{E}(Y_t) = \sum_{j=1}^{p} \phi_j \mathbb{E}(Y_{t-j})$. Define $\mu := \mathbb{E}(Y_t)$. Then, since $\{Y_t\}$ stationary: $\mu = \mu \sum_{j=1}^{p} \phi_j$

$$\text{i.e.} \quad \mu\left(1 - \sum_{j=1}^{p} \phi_j\right) = 0\,.$$

But if $1 - \sum_{j=1}^{p} \phi_j = 0$ then the $AR(p)$ polynomial

$$\phi(x) = 1 - \sum_{j=1}^{p} \phi_j x^j\,,$$

has a root (at $x = 1$) which is not outside unit circle. Therefore, we must have that $\mu = 0$.

To allow $\{Y_t\}$ to have (possibly) non-zero mean...

Definition 10 (*ARMA*(p, q) process with constant)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0, \sigma^2)$. *Then,* $\{Y_t\}$, *with* $\mu := \mathbb{E}(Y_t)$, *is an* *autoregressive, moving average process of order* (p, q), *written* *ARMA*(p, q), *if it contains* p*-many* *AR* *terms and* q*-many* *MA* *terms:*

$$Y_t - \mu = \sum_{j=1}^{p} \phi_j (Y_{t-j} - \mu) + \epsilon_t - \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \, .$$

Remark 11

I.e., for $\mathbb{E}(Y_t) = \mu$, *the* *ARMA* *model can be written*

$$\phi(B)(Y_t - \mu) = \theta(B)\epsilon_t \, ,$$

and the *ARIMA* *model is*

$$\phi(B)(1 - B)^d (Y_t - \mu) = \theta(B)\epsilon_t \, ,$$

Proposition 12

Let $\{X_t\}$ *be a stationary process. Let*

$$Y_t := \mu(t; d) + X_t ,$$

where

$$\mu(t; d) := \sum_{j=0}^{d} m_j t^j , \quad m_j \in \mathbb{R} .$$

(i.e. the mean of Y_t *is a deterministic polynomial of degree* d*).*
Then,

$$\nabla^d Y_t = d! m_d + \nabla^d X_t ,$$

and $\mathbb{E}(\nabla^d Y_t) = d! m_d$ *(a constant).*

Proof Omitted.

Corollary 13

Let the d*-differenced process* $\{Z_t := \nabla^d Y_t\}$ *be a stationary* $ARMA(p, q)$ *process with non-zero, constant, mean. Then,* $\{Y_t\}$ *contains a polynomial trend term of degree* d*.*

Example 14

Recall the process

$$Y_t = m_0 + m_1 t\,, \quad m_0, m_1 \in \mathbb{R}\,.$$

In this case Y_t *is not stationary, since (the 'mean function')*

$$\mathbb{E}Y_t = m_0 + m_1 t\,,$$

which depends on t*.*

But, we can 'transform' the original non-stationary process to a stationary one by differencing. Consider:

$$\begin{aligned} Z_t &:= Y_t - Y_{t-1} \\ &= \cancel{m_0} + \cancel{m_1 t} - (\cancel{m_0} + m_1(\cancel{t} - 1)) \\ &= m_1\,, \end{aligned}$$

which is now stationary.

Example 15

Consider the process $Y_t = m_0 + m_1 t + m_2 t^2$*. Show* $\nabla^2 Y_t = 2m_2$*.*

Recall (lecture 1) that time series can comprise: **trend terms**, **seasonal behaviour**, **short term correlations**, and **'noise'**.

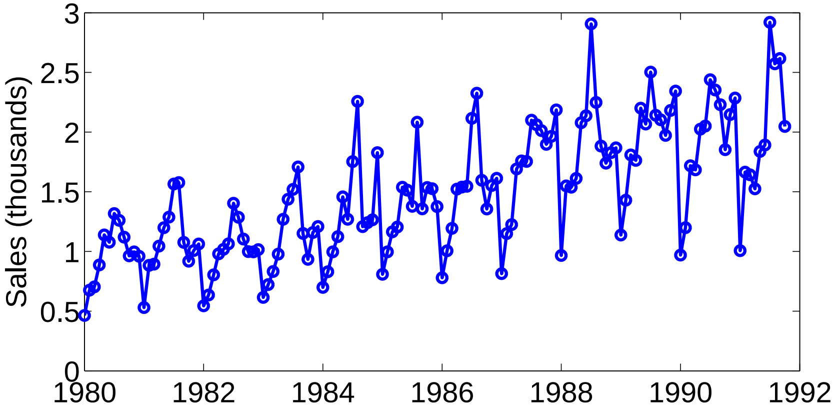

Figure: Australian red wine sales, Jan. '80 — Oct. '91

Often, time series contain seasonal (or periodic) components. For example, consider the following 'economic' data

Quarterly 'economic' data

Quarter	1	2	3	4
Year 1	y_1	y_2	y_3	y_4
Year 2	y_5	y_6	y_7	y_8
Year 3	y_9	y_{10}	y_{11}	y_{12}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

It might be reasonable to assume that correlations exist between quarters 1 and 5; 2 and 6; and so on. Namely

$$Y_t = \Phi_1 Y_{t-4} + \epsilon_t\,, \quad \epsilon_t \sim \mathcal{WN}(0, \sigma^2)\,.$$

This is a seasonal *AR* model.

If the data had been sampled at monthly intervals, we **might** expect:

$$Y_t = \Phi_1 Y_{t-12} + \epsilon_t\,, \quad \epsilon_t \sim \mathcal{WN}(0, \sigma^2)\,.$$

Likewise, the noise terms could contain perdiodic components.

Definition 16 (seasonal moving average $SMA(1)_s$ model)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0, \sigma^2)$. Then $\{Y_t\}$ is a seasonal moving average process of order 1, with period s, written as $SMA(1)_s$, if

$$Y_t = \epsilon_t - \Theta_1 \epsilon_{t-s}\,, \quad s \geq 2$$

Note $\mathbb{E}(Y_t) = 0$. Multiply both sides by Y_{t-k}, take $\mathbb{E}$:

$$\begin{aligned}
\gamma(k) &= \mathbb{E}(Y_t Y_{t-k}) = \mathbb{E}\big((\epsilon_t - \Theta_1\epsilon_{t-s})(\epsilon_{t-k} - \Theta_1\epsilon_{t-k-s})\big) \\
&= \mathbb{E}(\epsilon_t\epsilon_{t-k}) + \Theta_1^2\mathbb{E}(\epsilon_{t-s}\epsilon_{t-k-s}) - \Theta_1\mathbb{E}(\epsilon_t\epsilon_{t-k-s}) - \Theta_1\mathbb{E}(\epsilon_{t-s}\epsilon_{t-k}) \\
&= \sigma^2\delta_{0,k} + \sigma^2\Theta_1^2\delta_{0,k} - \sigma^2\Theta_1\delta_{-s,k} - \sigma^2\Theta_1\delta_{s,k} \\
&= \sigma^2\big(\delta_{0,k}(1+\Theta_1^2) - \Theta_1\delta_{s,|k|}\big) \\
&= \begin{cases} \sigma^2(1+\Theta_1^2)\,, & k = 0 \\ -\sigma^2\Theta_1\,, & |k| = s \\ 0\,, & \text{oth.} \end{cases}
\end{aligned}$$

$$\gamma(k) = \begin{cases} \sigma^2(1+\Theta_1^2)\,, & k=0 \\ -\sigma^2\Theta_1\,, & |k|=s \\ 0\,, & \text{oth.} \end{cases}$$

And, $\rho(k) = \gamma(k)/\gamma(0) \Rightarrow$

$$\rho(k) = \begin{cases} 1\,, & k=0 \\ \dfrac{-\Theta_1}{1+\Theta_1^2}\,, & |k|=s \\ 0\,, & \text{oth.} \end{cases}$$

Remark 17

The $SMA(1)_s$ process $Y_t = \epsilon_t - \Theta_1\epsilon_{t-s}$ can be written, in backshift notation, as $Y_t = (1-\Theta_1 B^s)\epsilon_t$.

Definition 18 (seasonal moving average $SMA(Q)_s$ model)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0,\sigma^2)$. Then $\{Y_t\}$ is a seasonal moving average process of order Q, with period s, written as $SMA(Q)_s$, if

$$Y_t = \epsilon_t - \Theta_1\epsilon_{t-s} - \Theta_2\epsilon_{t-2s} - \ldots - \Theta_Q\epsilon_{t-Qs}\,, \quad s \geq 2$$

Remark 19

The $SMA(Q)_s$ *process* $Y_t = \epsilon_t - \sum_{j=1}^{Q} \Theta_j \epsilon_{t-js}$ *can be written, in backshift notation, as* $Y_t = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \ldots - \Theta_Q B^{Qs})\epsilon_t$ *. I.e.*

$$Y_t = \left(1 - \sum_{j=1}^{Q} \Theta_j B^{js}\right)\epsilon_t =: \Theta(B)\epsilon_t$$

Definition 20

$\Theta(B)$ *is called the seasonal MA characteristic polynomial.*

Remark 21

A $SMA(Q)_s$ *process is always stationary. It is invertible iff roots of* $\Theta(B)$ *lie outside unit circle. (c.f.* $MA(q)$ *process.)*

Remark 22

$$\rho(ks) = \begin{cases} \frac{-\Theta_k + \sum_{j=1}^{Q-k} \Theta_j \Theta_{j+k}}{1 + \sum_{j=1}^{Q} \Theta_j^2}, & k = 1, 2, \ldots, Q \\ 0, & oth. \end{cases}$$

I.e. ACF is only non-zero at $k = 0, s, 2s, \ldots, Qs$.

Definition 23 (seasonal autoregressive $SAR(1)_s$ model)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0, \sigma^2)$. Then $\{Y_t\}$ is a seasonal autoregressive process of order 1, with period s, written as $SAR(1)_s$, if

$$Y_t = \Phi_1 Y_{t-s} + \epsilon_t$$

Note $\mathbb{E}(Y_t) = 0$. For $k \geq 1$, multiply both sides by Y_{t-k}, take $\mathbb{E}$:

$$\begin{aligned} \gamma(k) &= \mathbb{E}(Y_t Y_{t-k}) = \mathbb{E}(Y_{t-s} Y_{t-k}) + \underbrace{\cancel{\mathbb{E}(\epsilon_t Y_{t-k})}}_{0,\, k \geq 1} \\ &= \Phi_1 \gamma(k-s)\,, \end{aligned}$$

Now, divide both sides by $\gamma(0)$:

$$\rho(k) = \Phi_1 \rho(k-s), \quad k \geq 1\,.$$

$$\begin{aligned} \text{At } k = s: \quad \rho(s) &= \Phi_1 \rho(0) = \Phi_1 \\ \text{At } k = 2s: \quad \rho(2s) &= \Phi_1 \rho(s) = \Phi_1^2 \\ \vdots \qquad\qquad & \vdots \\ \text{At } k = \ell s: \quad \rho(\ell s) &= \Phi_1 \rho((\ell-1)s) = \Phi_1^{\ell}\,, \ell \in \mathbb{N} \end{aligned}$$

Recall

$$\rho(k) = \Phi_1 \rho(k - s), \quad k \geq 1 .$$

At $k = s - m$, where $1 \leq m \leq s - 1$:

$$\rho(s - m) = \Phi_1 \rho(-m) = \Phi_1 \rho(m) . \tag{1}$$

But at $k = m$:

$$\rho(m) = \Phi_1 \rho(m - s) = \Phi_1 \rho(s - m) = \Phi_1^2 \rho(m), \quad \text{[from (1)]}$$

But, $\{Y_t\}$ stationary $\Rightarrow |\Phi_1| < 1$, i.e. $\Phi_1 \neq 1 \Rightarrow \rho(m) = 0$ for $m = 1, \ldots, s - 1$.

Similarly, $\rho(2s - m) = \Phi_1 \rho(s - m) = 0$ [from (1)]. Likewise, in general: $\rho(\ell s - m) = 0 \, \forall \ell \in \mathbb{Z}$, and

$$\rho(k) = \begin{cases} \Phi_1^k , & k = 0, \pm s, \pm 2s, \pm 3s, \ldots \\ 0 , & \text{oth.} \end{cases}$$

I.e., the non-zeros at lags $k = 0, \pm s, \pm 2s, \ldots$ decay exponentially.

Remark 24

The $SAR(1)_s$ *process* $Y_t = \Phi_1 Y_{t-s} + \epsilon_t$ *can be written, in backshift notation, as* $(1 - \Phi_1 B^s) Y_t = \epsilon_t$.

Definition 25 (seasonal autoregressive $SAR(P)_s$ model)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0, \sigma^2)$. *Then* $\{Y_t\}$ *is a* seasonal autoregressive process *of order* P, *with period* s, *written as* $SAR(P)_s$, *if*

$$Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \ldots + \Phi_P Y_{t-Ps} + \epsilon_t$$

Remark 26

The $SAR(P)_s$ *process* $Y_t = \epsilon_t + \sum_{j=1}^{P} \Phi_j Y_{t-js}$ *can be written, in backshift notation, as*

$$\Phi(B) Y_t := \left(1 - \sum_{j=0}^{P} \Phi_j B^{js}\right) Y_t = \epsilon_t$$

$SAR(P)_s:\ Y_t = \epsilon_t + \sum_{j=1}^{P} \Phi_j Y_{t-js}$, can be written as

$$\Phi(B)Y_t := \left(1 - \sum_{j=0}^{P} \Phi_j B^{js}\right) Y_t = \epsilon_t$$

Definition 27

$\Phi(B)$ is called the seasonal AR characteristic polynomial.

Remark 28

A $SAR(P)_s$ process is stationary iff all roots of $\Phi(B)$ lie outside unit circle. (c.f. $AR(p)$ process.)

Generally, processes can have short-term correlations (like *ARMA* processes) as well as seasonal correlations (like *SMA* and *SAR*).

Definition 29 (multiplicative seasonal *ARMA* models)

Let $\{\epsilon_t\} \sim \mathcal{WN}(0, \sigma^2)$. *Then* $\{Y_t\}$ *is a multiplicative seasonal autoregressive, moving average process of order* $(p, q) \times (P, Q)_s$ *, with period* s*, written as* $SARMA(p, q) \times (P, Q)_s$ *, if*

$$\phi(B)\,\Phi(B)\,Y_t = \theta(B)\,\Theta(B)\,\epsilon_t$$

where

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

Example 30

$SARMA(0, 1) \times (0, 1)_{12}$, $SARMA(0, 1) \times (1, 0)_4$, *and* $SARMA(1, 1) \times (2, 1)_s$, *pages 35–37 in the course notes.*