MATH3802
Chapter 3: Models for time series

Peter Thwaites

with thanks to Stuart Barber for previous years' notes and slides

February 2017

3.1 White noise

Definition 3.1 (White noise) The time series $\{X_t\}$ is called *white noise* if the X_t are i.i.d. with $\mathbb{E}(X_t) = 0$ for all t.

If $\{X_t\}$ is white noise, then

$$\mu(t) = \mathbb{E}(X_t) = 0;$$

$$\gamma_k = \text{cov}(X_t, X_{t+k}) = \begin{cases} \sigma^2 & \text{if } k = 0 \\ 0 & \text{else;} \end{cases}$$

$$\rho_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{else.} \end{cases}$$

Remarks:

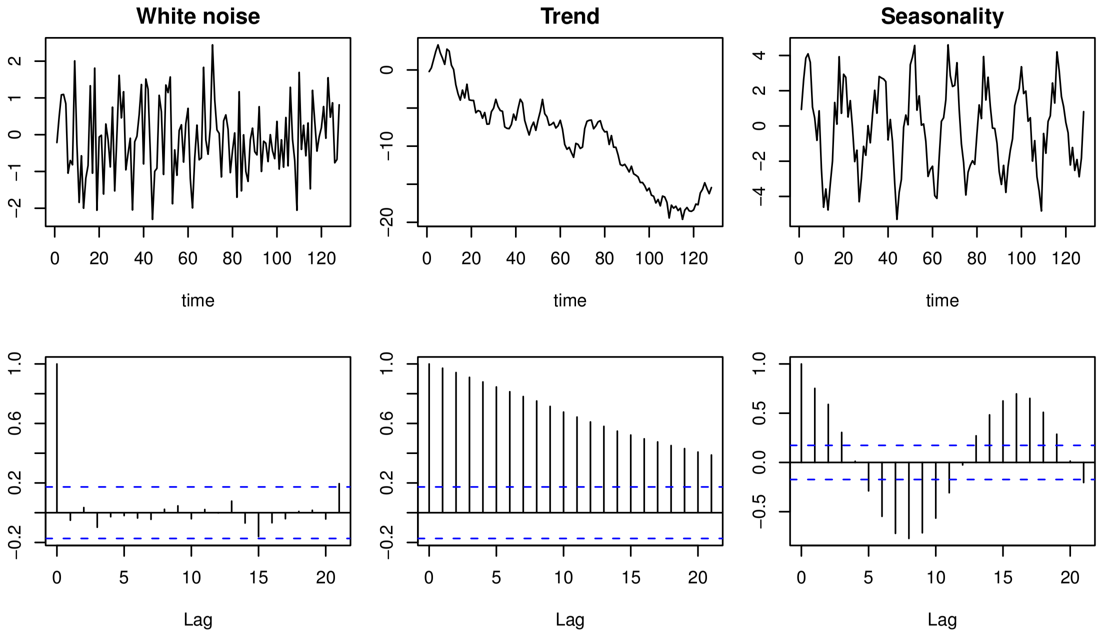
1. Often we assume $X_t \sim N(0, \sigma^2)$ for all t.
2. White noise is used to model the residuals of more complicated time series.
3. We will usually denote a white noise process as $\{\varepsilon_t\}$, with variance σ_ε^2.

We can use the correlogram to distinguish between white noise and processes with dependence.

Bartlett's theorem

Theorem (Bartlett, 1946) If $\{X_t\}$ is white noise, then for large n the distribution of $\widehat{\rho}_k, k \neq 0$, is approximately $N(0, 1/n)$.

One considers values $|\widehat{\rho}_k| > 1.96/\sqrt{n}$ significant at the 5% level.
But note that the $\widehat{\rho}_k$ are not independent of each other!



3.2 Moving average (MA) models

Definition 3.2 (Moving average processes) A stochastic process $\{X_t\}$ is called a *moving average process of order q* (or an $MA(q)$ process) if

$$X_t = \sum_{k=0}^{q} \beta_k \varepsilon_{t-k} = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q},$$

where $\beta_0, \ldots, \beta_q \in \mathbb{R}$ and $\{\varepsilon_t\}$ is white noise.

Remarks:

1. Without loss of generality, we can assume $\beta_0 = 1$ (since we can choose $\sigma_\varepsilon^2 = \mathrm{var}(\varepsilon_t)$).
2. Since $\{\varepsilon_t\}$ is stationary, $\{X_t\}$ is stationary.

ACF of MA

For an MA(q) process, we have

$$\begin{aligned}
\mu &= 0,\\
\gamma_0 &= (\beta_0^2 + \cdots + \beta_q^2)\sigma_\varepsilon^2,\\
\gamma_k &= \begin{cases} \displaystyle\sum_{i=0}^{q-k} \beta_i\beta_{i+k}\sigma_\varepsilon^2 & \text{if } 0 \leqslant k \leqslant q \\ 0 & \text{else.} \end{cases}
\end{aligned}$$

So the acf ‘cuts off’ at lag q. This is a special feature of MA processes and can be used to recognise data sets for which an MA process is suitable. More on this in chapter 4.

The R script sim.ma.R, available in the VLE, simulates a number of MA processes.

Example 3.1 (MA(0), MA(1))

1. MA(0) is white noise:

$$X_t = \beta_0\varepsilon_t, \ \text{iid } N(0, \beta_0^2\sigma_\varepsilon^2).$$

2. MA(1) has $X_t = \varepsilon_t + \beta\varepsilon_{t-1}$ (assuming $\beta_0 = 1$). Therefore,

$$\begin{aligned}
\gamma_0 &= (1+\beta^2)\sigma_\varepsilon^2 & \rho_0 &= 1\\
\gamma_1 &= \beta\sigma_\varepsilon^2 & \rho_1 &= \frac{\beta}{1+\beta^2}\\
\gamma_k &= 0 \ \ \forall k > 1 & \rho_k &= 0 \ \ \forall k > 1.
\end{aligned}$$

Invertibility

We have seen that MA(1) has

$$\rho_0 = 1 \qquad \rho_1 = \frac{\beta}{1+\beta^2} \qquad \rho_2 = \rho_3 = \cdots = 0.$$

Assume we have observed ρ_1 and want to determine β:

$$\beta_{1,2} = \frac{1 \pm \sqrt{1-4\rho_1^2}}{2\rho_1}.$$

There are two MA(1) processes with the same autocorrelation function! This problem is referred to as "invertibility"; we shall discuss this in more detail later in chapter 4.

3.3 Autoregressive (AR) models

Defintion 3.3 A stochastic process $\{X_t\}$ is an *autoregressive process of order p* (AR(p) process) if

$$X_t = \sum_{k=1}^{p} \alpha_k X_{t-k} + \varepsilon_t \qquad \forall t,$$

where $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$, $\{\varepsilon_t\}$ is white noise, and ε_t is independent of X_s for all $s < t$.

Remarks:

1. When constructing the process $\{X_t\}$, the first p values need to be specified as an initial condition.
2. We shall see that whether $\{X_t\}$ is stationary depends on $\alpha_1, \ldots, \alpha_p$ and on the initial conditions.

The R script sim.ar.R, available in the VLE, simulates a number of AR processes.

AR(1)

Example 3.2 The process $\{X_t\}$ given by

$$X_0 = 0, \quad X_t = X_{t-1} + \varepsilon_t \quad \forall t \in \mathbb{N},$$

with $\{\varepsilon_t\}$ a white noise process, is an AR(1) process with $\alpha_1 = 1$.

Such a process is called a *random walk*. We have $X_t = \sum_{k=1}^{t} \varepsilon_k$, so

$$\mathbb{E}(X_t) = 0 \qquad \text{var}(X_t) = t\sigma_\varepsilon^2.$$

Therefore $\{X_t\}$ is not stationary.

A general AR(1) process has $X_t = \alpha X_{t-1} + \varepsilon_t$, $\alpha \neq 0$, so

$$\mathbb{E}(X_t) = \alpha\mathbb{E}(X_{t-1}) = \cdots = \alpha^t\mathbb{E}(X_0)$$

$$\text{var}(X_t) = \alpha^2 \text{var}(X_{t-1}) + \sigma_\varepsilon^2.$$

Claim: Such a process is weakly stationary if and only if

1. $|\alpha| < 1$,
2. $\mathbb{E}(X_t) = 0 \ \forall \ t$, and
3. $\text{var}(X_t) = \sigma_\varepsilon^2/(1-\alpha^2) \ \forall \ t$ including $t = 0$.

Example 3.3 Consider the AR(1) process with $\alpha_1 = -0.8$. Note that $|\alpha_1| < 1$; necessary but not sufficient for weak stationarity.

- Let $X_t = -0.8X_{t-1} + \varepsilon_t$ with ε_t iid $N(0,1)$ and fix $X_0 = 0$. This is not stationary since $\mathrm{var}(X_0) = 0 \neq 1/0.36 = \mathrm{var}(X_t)$ for $t \gg 0$.
- Now let $X_0 \sim N(0, 1/0.36)$. Now $|\alpha_1| < 1$, $\mathbb{E}(X_t) = 0$ and $\mathrm{var}(X_t) = 1/0.36$ for all t; $\{X_t\}$ is weakly stationary.

Stationarity of AR(p)

Proposition 3.1 An AR(p) process is stationary if and only if all the roots $y_1, \ldots, y_p$ of the equation

$$\alpha(y) = 1 - \alpha_1 y - \cdots - \alpha_p y^p$$

are such that $|y_i| > 1$.

Remark: Note that for the AR(1) process $X_t = \alpha X_{t-1} + \varepsilon_t$, we have

$$\alpha(y) = 1 - \alpha_1 y = 0 \Leftrightarrow y = 1/\alpha$$

and

$$|y_1| > 1 \Leftrightarrow |\alpha| < 1$$

as we saw above.

Proof of propsition 3.1

We prove proposition 3.1 in three steps:

1. Write X_t in terms of $\alpha(\cdot)^{-1}$.
2. Write X_t in terms of $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$ — infinite recursion.
3. Examine coefficients $\{c_k\}$ of $\{\varepsilon_{t-k}\}$ to see when stationarity is possible.

Details of the proof appear in the Lecture Notes in the VLE. The definition and use of the backshift operator B is important, and will be examined.

Example 3.4 Consider the AR(2) process

$$X_t = \frac{1}{2}X_{t-1} + \frac{1}{2}X_{t-2} + \varepsilon_t.$$

The roots are $1, -2$ (exercise: check), so this process is *not* stationary.

Remarks

1. In case of stationarity, we have

$$\mathbb{E}(X_t) = \mathbb{E}\left(\alpha(B)^{-1}\varepsilon_t\right) = \mathbb{E}\left(\sum_{k=0}^{\infty} c_k B^k \varepsilon_t\right) = \sum_k c_k \mathbb{E}(\varepsilon_{t-k}) = 0.$$

2. The roots of $\alpha()$ can be complex. For example, consider $X_t = -X_{t-2} + \varepsilon_t$. This has $\alpha(y) = 1 - (-1)y^2 = 1 + y^2$, with roots $y_1 = i, y_2 = -i$. Hence $\{X_t\}$ is not stationary.

Stationarity of AR(2)

For AR(2) we have

$$\alpha(y) = 1 - \alpha_1 y - \alpha_2 y^2$$

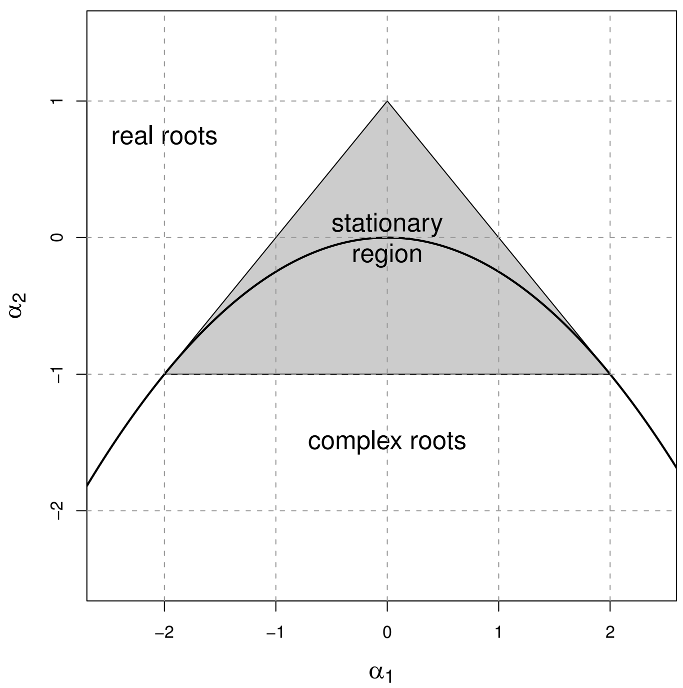
$$\Rightarrow \text{roots} \quad y_{1,2} = \frac{\alpha_1 \pm \sqrt{\alpha_1^2 + 4\alpha_2}}{-2\alpha_2}.$$

If $\alpha_1^2 + 4\alpha_2 > 0$ we have two real roots.
If $\alpha_1^2 + 4\alpha_2 < 0$ we have two complex roots. In this case,

$$y_{1,2} = \frac{\alpha_1 \pm i\sqrt{-4\alpha_2 - \alpha_1^2}}{-2\alpha_2}$$

$$\Rightarrow |y_{1,2}|^2 = \underbrace{\frac{\alpha_1^2}{4\alpha_2^2}}_{\text{Re}^2} + \underbrace{\frac{-4\alpha_2 - \alpha_1^2}{4\alpha_2^2}}_{\text{Im}^2}.$$

One can check that the process is stationary if $\alpha_2 > -1$, $\alpha_2 < 1 - \alpha_1$ and $\alpha_2 < 1 + \alpha_1$.



Stationarity of AR(p), $p > 2$

For AR(p), $p > 2$, we use a computer to find the roots.

Eg the *R* command `polyroot` which takes the argument $(1, -\alpha_1, \ldots, -\alpha_p)$. Using

```
> polyroot(c(1, -0.5, -0.5))
[1]  1-0i -2+0i
```

will verify example 3.4.

Autocovariance of AR(p)

Recall $\gamma_k = \text{cov}(X_t, X_{t+k})$

Since

$$X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \varepsilon_t$$
$$X_{t+k} = \alpha_1 X_{t+k-1} + \cdots + \alpha_p X_{t+k-p} + \varepsilon_{t+k},$$

we have

$$\gamma_k = \sum_{j=1}^{p} \alpha_j \gamma_{k-j} \qquad \forall k \geqslant 1.$$

Hence the autocorrelations $\rho_k = \gamma_k / \gamma_0$ are

$$\rho_k = \sum_{j=1}^{p} \alpha_j \rho_{k-j}.$$

AR(2)

Example 3.5 (AR(2)) We have

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} \qquad k = 1, 2, \ldots,$$

so

$$\begin{aligned}
\rho_0 &= 1 \\
\rho_1 &= \alpha_1 \rho_0 + \alpha_2 \rho_{-1} = \alpha_1 + \alpha_2 \rho_1 = \frac{\alpha_1}{1-\alpha_2} \\
\rho_2 &= \alpha_1 \rho_1 + \alpha_2 \rho_0 = \frac{\alpha_1^2}{1-\alpha_2} + \alpha_2 \\
&\vdots
\end{aligned}$$

Yule-Walker equations

We can determine $\alpha_1, \ldots, \alpha_p$ from $\rho_1, \ldots, \rho_p$ using

$$\left.\begin{array}{rcl}
\rho_1 &=& 1 \cdot \alpha_1 + \rho_1 \alpha_2 + \cdots + \rho_{p-1} \alpha_p \\
\rho_2 &=& \rho_1 \alpha_1 + 1 \cdot \alpha_2 + \cdots + \rho_{p-2} \alpha_p \\
&\vdots& \\
\rho_p &=& \rho_{p-1} \alpha_1 + \rho_{p-2} \alpha_2 + \cdots + 1 \cdot \alpha_p.
\end{array}\right\} \qquad (1)$$

The equations (1) are called the *Yule-Walker* equations. In matrix notation they are

$$\begin{bmatrix}
1 & \rho_1 & \cdots & \cdots & \rho_{p-1} \\
\rho_1 & 1 & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & 1 & \rho_1 \\
\rho_{p-1} & \cdots & \cdots & \rho_1 & 1
\end{bmatrix}
\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix}
=
\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_p \end{bmatrix}.$$

AR(2) Yule-Walker

Example 3.6(AR(2)) Consider an AR(2) process with $\rho_0 = 1, \rho_1 = 1/6, \rho_2 = -11/24$. We get the Yule-Walker equations

$$\begin{aligned} \alpha_1 + \frac{1}{6}\alpha_2 &= \frac{1}{6} \\ \frac{1}{6}\alpha_1 + \alpha_2 &= -\frac{11}{24}. \end{aligned}$$

Solving these equations for α_1 and α_2 yields

$$\alpha_1 = \frac{1}{4} \qquad \alpha_2 = -\frac{1}{2}$$

(exercise — check).

Run the R script yule-walker.R, available in the VLE.

AR(p) model-fitting summary

To fit an AR(p) model to data $\{X_t\} = (X_1, \ldots, X_n)$, we use the following steps.

1. Subtract trend and seasonal effects from $\{X_t\}$ to obtain residuals $\{Y_t\}$.
2. Estimate the a.c.f. of Y to obtain $\widehat{\rho}_1, \ldots, \widehat{\rho}_p$.
3. Solve the Yule-Walker equations to obtain $\widehat{\alpha}_1, \ldots, \widehat{\alpha}_p$.
4. Consider the residuals $Z_t = Y_t - \widehat{\alpha}_1 Y_{t-1} - \cdots - \widehat{\alpha}_p Y_{t-p}$. Use the Bartlett bands to check whether the $\{Z_t\}$ are (approximately) white noise (otherwise the model is not a good fit to the data).
5. Use the sample variance of the $\{Z_t\}$ to estimate σ_ε^2.
6. Add trend and seasonal effects back on to conclusions about $\{Y_t\}$ to get conclusions for $\{X_t\}$.

3.4 Mixed autoregressive moving average (ARMA) models

Definition 3.4 (ARMA(p, q) model) The ARMA(p, q) model satisfies

$$X_t = \sum_{i=1}^{p} \alpha_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}, \qquad (2)$$

with ε_t independent of $X_{t-1}, X_{t-2}, \ldots$.

Remarks:

1. Equation (2) can be written as

$$\alpha(B) X_t = \beta(B) \varepsilon_t,$$

where

$$\alpha(y) = 1 - \sum_{i=1}^{p} \alpha_i y^i, \qquad \beta(y) = 1 + \sum_{j=1}^{q} \beta_j y^j.$$

2. As for AR(p), we can write

$$X_t = \alpha(B)^{-1} \beta(B) \varepsilon_t.$$

The model is weakly stationary if

$$X_t = \sum_{k=0}^{\infty} \lambda_k \beta(B) \varepsilon_{t-k} = \sum_{k=0}^{\infty} \tilde{\lambda}_k \varepsilon_{t-k}$$

with $\sum_{k=0}^{\infty} \tilde{\lambda}_k^2 < \infty$.

This is again equivalent to the roots of $\alpha()$ lying outside the complex unit circle.

3. If stationary, we have $\mathbb{E}(X_t) = \sum_{k=0}^{\infty} \tilde{\lambda}_k \mathbb{E}(\varepsilon_{t-k}) = 0$.

4. We can reconstruct the noise as

$$\varepsilon_t = \beta(B)^{-1}\alpha(B)X_t = \sum_{k=0}^{\infty} \delta_k X_{t-k}.$$

If $\sum_{k=0}^{\infty} \delta_k^2 < \infty$, the process is called *invertible* (more on invertibility later in chapter 4).

In this case, the influence of X_{t_0} (t_0 fixed) on X_t gets smaller as $t \to \infty$.

Result $\{X_t\}$ is invertible iff the roots of β lie outside the complex unit circle.

Example

Example 3.7 Consider the ARMA(1,1) process

$$X_t = \alpha X_{t-1} + \varepsilon_t + \beta\varepsilon_{t-1}.$$

Stationary iff $|\alpha| < 1$.

Invertible iff roots of $1 + \beta y$ lie outside the unit circle, iff $|\beta| < 1$.

Auto-covariances of ARMA

$$\begin{aligned}
\gamma_k &= \operatorname{cov}(X_t, X_{t+k}) \\
\Rightarrow \gamma_k &= \sum_{i=1}^{p} \alpha_i \gamma_{k-i} \qquad \forall k > q \\
\Rightarrow \rho_k &= \sum_{i=1}^{p} \alpha_i \rho_{k-i} \qquad \forall k > q.
\end{aligned}$$

Reminder: AR(p) satisfies the same relationship (but for all $k > 0$). Hence the a.c.f. for AR and ARMA show the same behaviour for $k > q$.

Differencing

Recall the AR(1) process

$$X_t = X_{t-1} + \varepsilon_t = \sum_{s=0}^{t} \varepsilon_s,$$

where $\{\varepsilon_t\}$ is white noise, is not stationary.

Definition 3.5 (Difference operator) Let

$$\nabla X_t = X_t - X_{t-1} \quad \forall t.$$

The operator ∇ ("nabla") is called the *difference operator*.

Remarks

1. We can write $\nabla = 1 - B$ where B is the backshift operator:
2. If $\{X_t\}$ has stationary increments, then ∇X is stationary.
3. ∇ removes a constant mean:

$$\nabla(X_t + \mu) = \nabla X_t.$$

4. A linear trend is converted to a constant mean:

$$\nabla(X_t + \alpha + \beta t) = \nabla X_t + \beta.$$

Autoregressive Integrated Moving Average process

Definition 3.6 (ARIMA(p, d, q)) $\{X_t\}$ is an ARIMA(p, d, q) process if $\nabla^d X$ is a stationary ARMA(p, q) process.

Remark An ARIMA(p, d, q) process $\{X_t\}$ can be written as an ARMA($p + d, q$) process which has a unit root and hence is non-stationary for $d > 0$.

Example 3.8 Let $\varepsilon_t \sim N(0, \sigma^2)$ and $X_t = \sum_{s=1}^{t} \varepsilon_s$ (random walk).

Then $\nabla X_t = X_t - X_{t-1} = \varepsilon_t$ is white noise, i.e. $\{\varepsilon_t\}$ is an ARMA(0,0) process. Hence $\{X_t\}$ is ARIMA(0,1,0).