Chapter 3

ARIMA Processes

The simplest time series model is *white noise*. A first generalisation of white noise is the *moving average* model.

Definition 3.1 (The MA(q) process). *The process $\{X_t,\, t \in \mathbb{Z}\}$ is said to be a* moving average of order q *if*

$$X_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}, \quad \{\epsilon_t\} \sim WN(0, \sigma^2), \tag{3.1}$$

where $\theta_1, \ldots, \theta_q$ are constants.

Definition 3.2 (The AR(p) process). *The process $\{X_t,\, t \in \mathbb{Z}\}$ is said to be an AR(p)* autoregressive process of order p *if it is stationary and if*

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = \epsilon_t, \quad \{\epsilon_t\} \sim WN(0, \sigma^2). \tag{3.2}$$

A process $\{X_t\}$ is an AR(p) process with mean μ if $\{X_t - \mu\}$ is an AR(p) process.

Definition 3.3 (The ARMA(p, q) process). *A process $\{X_t,\, t \in \mathbb{Z}\}$ is said to be an ARMA(p, q)* process *if it is stationary and*

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}, \tag{3.3}$$

where $\{Z_t\} \sim WN(0, \sigma^2)$. A process $\{X_t\}$ is an ARMA(p, q) process with mean μ if $\{X_t - \mu\}$ is an ARMA(p, q) process.

Clearly, an ARMA($0, q$) process is an MA(q) process, while an ARMA($p, 0$) process is an AR(p) process.

Generating Polynomials for the ARMA Process An important tool for analysis of ARMA processes is the so-called *generating polynomial*. Equations (3.3) can be written as

$$\phi(B) X_t = \theta(B) \epsilon_t, \quad t \in \mathbb{Z},$$

where

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p,$$

$$\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$$

and B is the backward shift operator. The polynomials $\phi(\cdot)$ and $\theta(\cdot)$ are called *generating polynomials.*

Causal Models An important property of a time series model is that X_t depends only on information available up to and including time t. A linear time series model that satisfies this is said to be *causal.*

Definition 3.4. *An ARMA(p,q) process defined by the equations*

$$\phi(B)X_t = \theta(B)\epsilon_t \quad \{Z_t\} \sim WN(0, \sigma^2)$$

is said to be causal *if there exists constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and*

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \quad t \in \mathbb{Z}. \tag{3.4}$$

Another way to express this is to require that

$$\mathrm{Cov}(X_t, \epsilon_{t+j}) = 0 \quad \text{for} \quad j = 1, 2, \ldots \tag{3.5}$$

The following theorem gives conditions under which an ARMA(p,q) process is causal.

Theorem 3.5. *Let $\{X_t\}$ be an ARMA(p,q) for which $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeros. Then $\{X_t\}$ is causal if and only if $\phi(z) \neq 0$ for all $|z| \leq 1$. The coefficients $\{\psi_j\}$ in Equation* (3.4) *are determined by the relation*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1.$$

Proof Assume that $\phi(z) \neq 0$ if $|z| \leq 1$. Then $\frac{1}{\phi(z)}$ is analytic within the unit disc and therefore there exists a $(\xi_j)_{j=0}^{\infty}$ such that $\sum_{j=0}^{\infty} |\xi_j| < +\infty$ such that

$$\frac{1}{\phi(z)} = \sum_{j=0}^{\infty} \xi_j z^j = \xi(z), \quad |z| \leq 1.$$

The operator $\xi(B)$ may be applied to both sides of the equation $\phi(B)X_t = \theta(B)\epsilon_t$ to give:

$$X_t = \xi(B)\theta(B)\epsilon_t,$$

which is well defined since $\sum |\xi_j| < +\infty$ and $\theta(z)$ is a polynomial of degree q.

Now assume that $\phi(z) = 0$ for some $|z| \leq 1$ and consider the power series expansion $\frac{1}{\phi(z)} = \sum \xi_j z^j$. The coefficients are not summable, hence X_t does not satisfy the definition of a linear time series model. $\square$

If $\phi(B)X_t = \theta(B)\epsilon_t$ and if $\phi(z) = 0$ for some z with $|z| = 1$ then there does not exist a stationary solution. Consider for example $X_t = X_{t-1} + \epsilon_t$; $\phi(z) = 1 - z$ so that $\phi(1) = 0$. For $\{\epsilon_t\} \sim WN(0, \sigma^2)$ and $X_0 = 0$, then $X_t = \sum_{j=1}^{t} \epsilon_j$ so that $\text{Var}(X_t) = \sigma^2 t$, which is clearly not stationary.

Example 3.1 (AR(1) process)**.**

Let $\{X_t\}$ be an AR(1) process:

$$X_t = \epsilon_t + \phi X_{t-1} \quad \text{or} \quad \phi(z) = 1 - \phi z. \tag{3.6}$$

Since $1 - \phi z = 0$ gives $z = 1/\phi$ it follows that X_t is causal if $|\phi| < 1$. For $|\phi| < 1$,

$$X_t = \epsilon_t + \phi X_{t-1} = \epsilon_t + \phi(\epsilon_{t-1} + \phi X_{t-2}) = \epsilon_t + \phi\epsilon_{t-1} + \phi^2 X_{t-2} = \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j}.$$

It now follows directly from Theorem 2.4 that

$$\gamma_X(h) = \sum_{j=0}^{\infty} \phi^{2j+|h|}\sigma^2 = \frac{\sigma^2 \phi^{|h|}}{1 - \phi^2}, \tag{3.7}$$

which corresponds to the computation made earlier.

If $|\phi| > 1$, Equation (3.6) may be rewritten as:

$$\phi^{-1} X_t = \phi^{-1}\epsilon_t + X_{t-1} \quad \text{or} \quad X_t = -\phi^{-1}\epsilon_{t+1} + \phi^{-1} X_{t+1}.$$

It follows that X_t has representation

$$X_t = -\sum_{j=1}^{\infty} \phi^{-j} \epsilon_{t+j}.$$

If $|\phi| = 1$ there does not exist a stationary solution. □

Definition 3.6. *An ARMA(p, q) process defined by the equations*

$$\phi(B)X_t = \theta(B)\epsilon_t, \quad \{\epsilon_t\} \sim WN(0, \sigma^2)$$

is said to be invertible *if there exists constants $\{\pi_j\}$ such that $\sum_{j=0}^{\infty} |\pi_j| < \infty$*
and

$$\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad t \in \mathbb{Z}. \tag{3.8}$$

Theorem 3.7. *Let $\{X_t\}$ be an ARMA(p, q) for which $\phi(\cdot)$ and $\theta(\cdot)$ have no common zeros. Then $\{X_t\}$ is invertible if and only if $\theta(z) \neq 0$ for all $|z| \leq 1$. The coefficients $\{\pi_j\}$ in Equation* (3.8) *solve the equation:*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1.$$

Proof The proof follows in the same way as the proof of Theorem 3.5. □

Example 3.2 (MA(1) process).

Let $\{X_t\}$ be an MA(1) process:

$$X_t = \epsilon_t + \theta\epsilon_{t-1} \quad \text{or} \quad \theta(z) = 1 + \theta z.$$

Since $1 + \theta z = 0$ gives $z = -1/\theta$ it follows that X_t is invertible if $|\theta| < 1$. In that case

$$\epsilon_t = X_t - \theta\epsilon_{t-1} = X_t - \theta(X_{t-1} - \theta\epsilon_{t-2}) = \sum_{j=0}^{\infty}(-1)^j\theta^j X_{t-j}.$$

From Equation (2.3), with $\psi_0 = 1$, $\psi_1 = \theta$ and $\psi_j = 0$ for $j \neq 0, 1$, it follows that

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma^2 & \text{if } h = 0, \\ \theta\sigma^2 & \text{if } |h| = 1, \\ 0 & \text{if } |h| > 1. \end{cases} \tag{3.9}$$

□

Example 3.3 (ARMA(1,1) process).

Let $\{X_t\}$ be an ARMA(1,1) process, i.e.

$$X_t - \phi X_{t-1} = \epsilon_t + \theta\epsilon_{t-1} \quad \text{or} \quad \phi(B)X_t = \theta(B)\epsilon_t,$$

where $\phi(z) = 1 - \phi z$ and $\theta(z) = 1 + \theta z$. Let $|\phi| < 1$ and $|\theta| < 1$ so that X_t is causal and invertible. Then $X_t = \psi(B)\epsilon_t$, where

$$\psi(z) = \frac{\theta(z)}{\phi(z)} = \frac{1+\theta z}{1-\phi z} = \sum_{j=0}^{\infty}(1+\theta z)\phi^j z^j = 1 + \sum_{j=1}^{\infty}(\phi+\theta)\phi^{j-1}z^j.$$

It follows from Equation (2.3) that

$$\begin{aligned} \gamma(0) &= \sigma^2\sum_{j=0}^{\infty}\psi_j^2 = \sigma^2\left(1 + \sum_{j=1}^{\infty}(\phi+\theta)^2\phi^{2(j-1)}\right) \\ &= \sigma^2\left(1 + (\phi+\theta)^2\sum_{j=0}^{\infty}\phi^{2j}\right) = \sigma^2\left(1 + \frac{(\phi+\theta)^2}{1-\phi^2}\right). \end{aligned}$$

For $h > 0$,

$$\begin{aligned} \gamma(h) &= \sigma^2\sum_{j=0}^{\infty}\psi_j\psi_{j+h} = \sigma^2\left((\phi+\theta)\phi^{h-1} + \sum_{j=1}^{\infty}(\phi+\theta)^2\phi^{2(j-1)+h}\right) \\ &= \sigma^2\phi^{h-1}\left(\phi+\theta+\phi\sum_{j=0}^{\infty}(\phi+\theta)^2\phi^{2j}\right) = \sigma^2\phi^{h-1}\left(\phi+\theta+\frac{\phi(\phi+\theta)^2}{1-\phi^2}\right). \end{aligned}$$

Using Equation (2.4), the spectral density may be computed as:

$$f_X(\lambda) = \frac{\sigma^2}{2\pi}\frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2} = \frac{\sigma^2}{2\pi}\frac{|1+\theta\cdot e^{-i\lambda}|^2}{|1-\phi\cdot e^{-i\lambda}|^2} = \frac{\sigma^2}{2\pi}\frac{1+\theta^2+2\theta\cos(\lambda)}{1+\phi^2-2\phi\cos(\lambda)}, \quad -\pi \le \lambda \le \pi. \tag{3.10}$$

□

3.1 ARMA Approximations

This section gives results indicating that any real valued stationary time series $\{X_t\}$ with a well defined spectral density f_X may be approximated arbitrarily closely by an MA(q) process, or by an AR(p) process. The result is the following:

Theorem 3.8. *Let $\{X_t\}$ be a real valued stationary time series with spectral density f_X, where f_X is symmetric and continuous. Then for each $\epsilon > 0$, there exist a $q < +\infty$ and an invertible MA(q) process $\{Y_t\}$, a $p < +\infty$ and a causal AR(p) process $\{U_t\}$ such that*

$$|f_Y(\lambda) - f_X(\lambda)| < \epsilon \quad \textit{for all } \lambda \in [-\pi, \pi]$$

and

$$|f_U(\lambda) - f_X(\lambda)| < \epsilon \quad \textit{for all } \lambda \in [-\pi, \pi]$$

There are no 'difficult' steps in this theorem, but the proof is rather long and involved and therefore omitted from the course. I have included it in the notes.

The proof requires the following crucial preliminary results.

Lemma 3.9. *Let f be a symmetric, continuous spectral density on $[-\pi, \pi]$. For each $\epsilon > 0$, there exists a non-negative integer p and a polynomial*

$$a(z) = \prod_{j=1}^{p}\left(1 - \frac{z}{\eta_j}\right) = 1 + a_1 z + \ldots + a_p z^p$$

with $|\eta_j| > 1$ for each $j \in \{1, \ldots, p\}$ and where $a_1, \ldots, a_p$ are real, such that

$$\left|A|a(e^{-i\lambda})|^2 - f(\lambda)\right| < \epsilon \qquad \forall \lambda \in [-\pi, \pi]$$

where

$$A = \frac{1}{2\pi(1 + a_1^2 + \ldots + a_p^2)}\int_{-\pi}^{\pi} f(\nu)d\nu.$$

Proof of Lemma 3.9 If $f(\lambda) \equiv 0$, then the result follows with $p = 0$. Assume that $M := \sup_{-\pi\le\lambda\le\pi} f(\lambda) > 0$. For any $\epsilon > 0$, set

$$\delta = \min\left\{M, \frac{\epsilon}{2 + \frac{4\pi M}{\int_{-\pi}^{\pi} f(\nu)d\nu}}\right\}$$

and set

$$f^{(\delta)}(\lambda) = \max\{f(\lambda), \delta\}.$$

The function $f^{(\delta)}$ is also a symmetric continuous spectral density function. It satisfies $f^{(\delta)} \ge \delta$ and

$$0 \le f^{(\delta)} - f \le \delta \qquad \forall \lambda \in [-\pi, \pi]$$

Now use the following standard (and obvious) result: let

$$S_n f(x) = \sum_{|j|\leq n} f_j e^{ijx} \qquad f_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ijx} f(x)dx$$

then

$$\frac{1}{n}(S_0 f + S_1 f + \ldots + S_n f) \longrightarrow f$$

uniformly on $[-\pi, \pi]$ as $n \to +\infty$. From this, there exists an integer r such that

$$\left| \frac{1}{r}\sum_{j=0}^{r-1} \sum_{|k|\leq j} g_k e^{ik\lambda} - f^{(\delta)}(\lambda) \right| < \delta \qquad \forall \lambda \in [-\pi, \pi]$$

where

$$g_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f^{(\delta)}(\nu) e^{i\nu k} d\nu.$$

By changing the order of summation and using the symmetry of $f^{(\delta)}$, it follows that

$$\frac{1}{r}\sum_{j=0}^{r-1}\sum_{|k|\leq j} g_k e^{-ik\lambda} = \sum_{|k|<r}\left(1 - \frac{|k|}{r}\right) g_k e^{ik\lambda}$$

This function is strictly positive for all λ. This follows since $f^{(\delta)} \geq \delta$.

Let

$$C(z) = \sum_{|k|<r}\left(1 - \frac{|k|}{r}\right) g_k z^k$$

and note that (by symmetry) $C(z) = 0 \Leftrightarrow C(z^{-1}) = 0$. Let $p = \max\{k : g_k \neq 0\}$, then

$$z^p C(z) = K_1 \prod_{j=1}^{p}\left(1 - \frac{z}{\eta_j}\right)(1 - z\eta_j)$$

for some K_1 and $\eta_1, \ldots, \eta_p$ with $|\eta_j| > 1$, $j = 1, \ldots, p$. This equation may be written as:

$$\begin{aligned} C(z) &= K_1 \prod_{j=1}^{p}\left(1 - \frac{z}{\eta_j}\right)\left(\frac{1}{z} - \eta_j\right) \\ &= (-1)^p K_1 (\prod_{j=1}^{p}\eta_j)\prod_{j=1}^{p}\left(1 - \frac{z}{\eta_j}\right)\left(1 - \frac{1}{\eta_j z}\right) \\ &= K_2 a(z) a(z^{-1}) \end{aligned}$$

where

$$a(z) = 1 + a_1 z + \ldots + a_p z^p = \prod_{j=1}^{p}\left(1 - \frac{z}{\eta_j}\right)$$

and

$$K_2 = (-1)^p \eta_1 \ldots \eta_p K_1.$$

Equating the coefficients of z^0 gives:

$$K_2 = \frac{b_0}{1 + a_1^2 + \ldots + a_p^2} = \frac{1}{(2\pi)(1 + a_1^2 + \ldots + a_p^2)}\int_{-\pi}^{\pi} f^{(\delta)}(\nu) d\nu.$$

Furthermore,

$$\left| K_2 |a(e^{-i\lambda})|^2 - f^{(\delta)}(\lambda) \right| < \delta \qquad \forall \lambda.$$

Also,

$$\frac{|a(e^{-i\lambda})|^2}{1+a_1^2+\ldots+a_p^2} \leq \frac{2\pi(f^{(\delta)}(\lambda)+\delta)}{\int_{-\pi}^{\pi} f^{(\delta)}(\nu)d\nu} \leq \frac{4\pi M}{\int_{-\pi}^{\pi} f(\nu)d\nu}.$$

Using A defined as in the statement of the theorem, it follows that

$$\begin{aligned}\left|K_2|a(e^{-i\lambda})|^2 - A|a(e^{-i\lambda})|^2\right| &\leq \frac{1}{2\pi}\left(\int_{-\pi}^{\pi}(f^{(\delta)}(\nu)-f(\nu))d\nu\right)\frac{4\pi M}{\int_{-\pi}^{\pi} f(\nu)d\nu} \\ &\leq \frac{4\pi M\delta}{\int_{-\pi}^{\pi} f(\nu)d\nu}\end{aligned}$$

Finally,

$$\left|A|a(e^{-i\lambda})|^2 - f(\lambda)\right| < \delta+\delta+\frac{4\pi M\delta}{\int_{-\pi}^{\pi} f(\nu)d\nu} < \epsilon$$

from the definition of δ. □

Proof of Theorem 3.8: AR process The aim is to prove that for a continuous symmetric spectral density f and an $\epsilon > 0$, there is a p and a causal AR(p) process

$$U_t - a_1U_{t-1} - \ldots - a_pU_{t-p} = \epsilon_t \qquad \{\epsilon_t\} \sim \text{WN}(0,\sigma^2)$$

such that

$$|f_X(\lambda) - f_U(\lambda)| < \epsilon \qquad \forall \lambda \in [-\pi,\pi]$$

Let $f^{(\epsilon)} = \max\{f(\lambda), \frac{\epsilon}{2}\}$, then $f^{(\epsilon)} \geq \frac{\epsilon}{2}$ and

$$0 \leq f^{(\epsilon)}(\lambda) - f(\lambda) \leq \frac{\epsilon}{2} \qquad \forall \lambda \in [-\pi,\pi].$$

Set

$$M = \max_{\lambda} f^{(\epsilon)}(\lambda) \qquad \delta = \min\left\{\frac{\epsilon}{4M^2}, \frac{1}{2M}\right\}$$

Applying Lemma 3.9 to the function $\frac{1}{f^{(\epsilon)}(\lambda)}$, which is a spectral density, gives a K and a polynomial such that

$$\left|K|a(e^{-i\lambda})|^2 - \frac{1}{f^{(\epsilon)}(\lambda)}\right| < \delta \qquad \forall \lambda \in [-\pi,\pi]$$

where the polynomial $a(z) = 1+a_1z+\ldots+a_pz^p$ is non zero for $|z| \leq 1$ and K is a positive constant. From the definition of δ,

$$\frac{1}{K|a(e^{-i\lambda})|^2} \leq \frac{f^{(\epsilon)}(\lambda)}{1-\delta f^{(\epsilon)}(\lambda)} \leq \frac{M}{1-M\delta} \leq 2M.$$

It follows that

$$\left|\frac{1}{K|a(e^{-i\lambda})|^2} - f^{(\epsilon)}(\lambda)\right| = \left|K|a(e^{-i\lambda})|^2 - \frac{1}{f^{(\epsilon)}(\lambda)}\right|\frac{f^{(\epsilon)}(\lambda)}{K|a(e^{-i\lambda})|^2} < 2M^2\delta \leq \frac{\epsilon}{2}.$$

The inequalities now give

$$\left|\frac{1}{K|a(e^{-i\lambda})|^2} - f(\lambda)\right| < \epsilon \qquad \forall \lambda \in [\pi,\pi].$$

The process

$$a(B)X_t = \epsilon_t \qquad \{\epsilon_t\} \sim \text{WN}(0, \frac{2\pi}{K})$$

has spectral density $\frac{1}{K|a(e^{-i\lambda})|^2}$ and the proof is complete. □

The proof of the result for the MA(q) process is left as an exercise. □

For ϵ small, q and p may be rather large. In practice, it is often possible to find an ARMA(p',q') process where $p'+q' \leq \min(p,q)$.

3.1.1 ACVF and causal invertible ARMA processes

The following results shows that if γ is the ACVF of an ARMA(p,q) process, then there is a *causal invertible* ARMA(p,q) with ACVF γ.

Theorem 3.10. *Let $\{X_t\}$ be a stationary ARMA(p,q) process with ACVF γ, then there is a causal invertible process $\{\widetilde{X}_t\}$ with ACVF γ.*

Proof. Let $a_1, \ldots, a_p$ be the roots of $\phi(z)$, so that $\phi(z) = \prod_{j=1}^p (1 - \frac{z}{a_j})$ and let $b_1, \ldots, b_q$ be the roots of $\theta(z)$, so that $\theta(z) = \prod_{j=1}^q (1 - \frac{z}{b_j})$. The spectral density is:

$$\begin{aligned} f(\lambda) &= \frac{\sigma^2}{2\pi} \frac{\prod_{j=1}^q (1 - \frac{e^{-i\lambda}}{b_j})}{\prod_{j=1}^p (1 - \frac{e^{-i\lambda}}{a_j})} = \frac{\sigma^2}{2\pi} \frac{\prod_{j:|b_j|>1} (1 - \frac{e^{-i\lambda}}{b_j}) \prod_{j:|b_j|<1} (1 - \frac{e^{-i\lambda}}{b_j})}{\prod_{j:|a_j|>1} (1 - \frac{e^{-i\lambda}}{a_j}) \prod_{j:|a_j|<1} (1 - \frac{e^{-i\lambda}}{a_j})} \\ &= \frac{\sigma^2}{2\pi} \frac{\prod_{j:|b_j|>1} (1 - \frac{e^{-i\lambda}}{b_j}) \prod_{j:|b_j|<1} \frac{e^{i\lambda}}{|b_j|^2} (1 - \overline{b_j} e^{-i\lambda})}{\prod_{j:|a_j|>1} (1 - \frac{e^{-i\lambda}}{a_j}) \prod_{j:|a_j|<1} \frac{e^{i\lambda j}}{|a_j|^2} (1 - \overline{a_j} e^{-i\lambda})}. \end{aligned}$$

Now we appeal to the fact that the spectral density is real and hence for each a_j (and b_j) not real, there is a root $\overline{a_j}$ of the AR polynomial (and $\overline{b_j}$ of the MA polynomial). This leads to cancellation of the unwanted $e^{i\lambda}$ terms. Hence this is the spectral density of a causal invertible ARMA(p,q) process where we replace roots $a_j : |a_j| < 1$ by roots $\frac{1}{\overline{a_j}}$ and replace WN(0,σ^2) by $WN\left(0, \sigma^2 \frac{\prod_{j:|a_j|>1} |a_j|^2}{\prod_{j:|b_j|>1} |b_j|^2}\right)$. The ACVF uniquely determines the spectral density and vice versa. □

3.2 The ARIMA Process

The ARIMA process is defined as follows:

Definition 3.11 (The ARIMA(p, d, q) process)**.** *Let d be a non-negative integer. The process $\{X_t, t \in \mathbb{Z}\}$ is said to be an ARIMA(p, d, q) process if $\nabla^d X_t$ is a causal ARMA(p, q) process.*

A causal ARIMA(p, d, q) process $\{X_t\}$ satisfies:

$$\phi(B) X_t = \phi^*(B)(1 - B)^d X_t = \theta(B)\epsilon_t, \quad \{\epsilon_t\} \sim \text{WN}(0, \sigma^2), \tag{3.11}$$

where $\phi^*(z) \neq 0$ for all $|z| \leq 1$. The process $Y_t := \nabla^d X_t = (I - B)^d X_t$ satisfies:

$$\phi^*(B) Y_t = \theta(B)\epsilon_t.$$

Example 3.4 (Random Walk)**.**

Consider the simple *random walk* process:

$$X_t = X_{t-1} + \epsilon_t \qquad \{\epsilon_t\} \sim \text{WN}(0, \sigma^2) \qquad 0 < \sigma^2 < +\infty.$$

This is not a stationary process; $\text{Var}(X_t) = t\sigma^2 \stackrel{t\to+\infty}{\longrightarrow} +\infty$; the central limit theorem gives that $\frac{X_t}{t^{1/2}} \stackrel{t\to+\infty}{\longrightarrow}_{(d)} N(0,\sigma^2)$. A stationary process may be obtained from X by differencing; let

$$Y_t = \nabla X_t = X_t - X_{t-1} = (I - B)X_t.$$

then Y_t is a stationary process;

$$Y_t = \epsilon_t \sim \text{WN}(0, \sigma^2).$$

It follows that the random walk $\{X_t : t \in \mathbb{Z}_+\}$ is an ARIMA(0,1,0) process. □

It is clear that, for $d \geq 1$ there are no *stationary solutions* of Equation (3.11). Furthermore, neither the mean nor ACVF of $\{X_t\}$ are determined by (3.11), since any process $X_t + Y_t$, where Y_t disappears by differencing d times, satisfies equation (3.11). For example, if Y is a random variable, then $\nabla(X_t+Y) = \nabla X_t$.

For $|\phi| < 1$, the process

$$X_t - \phi X_{t-1} = \epsilon_t \qquad \{\epsilon_t\} \sim \text{WN}(0, \sigma^2)$$

is a causal AR(1) process and is stationary, while for $\phi = 1$, the process is not stationary, but is an ARIMA(0,1,0) process.

Recall that a causal AR(1) process has autocorrelation function

$$\rho(h) = \phi^{|h|}, \quad |\phi| < 1.$$

and hence, for any h,

$$\lim_{|\phi|\uparrow 1} |\rho(h)| = 1.$$

Similarly it holds for any ARMA process that its ACVF decreases slowly if some of the roots of $\phi(z) = 0$ are near the unit circle. From a sample of finite length, it is difficult to distinguish between an ARIMA$(p,1,q)$ process and an ARMA$(p+1,q)$ where $\phi(z)$ has a root near the unit circle. An estimated ACVF that decreases slowly indicates that differencing may be advisable.

Suppose that $\{X_t\}$ is a causal and invertible ARMA(p,q) process:

$$\phi(B)X_t = \theta(B)Z_t, \quad \{\epsilon_t\} \sim \text{WN}(0, \sigma^2),$$

where $\theta(z) \neq 0$ for all $|z| \leq 1$ and $\phi(z)$ has no roots in the unit circle. Then

$$\phi(B)\nabla X_t = \phi(B)(1-B)X_t = \theta(B)(1-B)\epsilon_t, \quad \{\epsilon_t\} \sim \text{WN}(0, \sigma^2),$$

from which it follows that ∇X_t is a causal, but non-invertible ARMA$(p, q+1)$ process. A unit root in the moving average polynomial indicates that X_t has been *overdifferenced*.

3.2.1 Testing for Unit Roots

For given time series data, there are tests available to indicate whether or not there are unit roots present. One common test is the *Dickey Fuller* test, introduced by Dickey and Fuller (1979), which has been refined to produce the *Augmented Dickey Fuller Test* (abbrieviated to ADF). This is a relatively straightforward test. It assumes that $\{\epsilon_t\} \sim IIDN(0, \sigma^2)$ (independent, identically distributed *normal* random variables) and works on the principles of linear regression.

The disadvantage of this test is that presence of a unit root is the *null* hypothesis. In statistics, a *null* hypothesis is never *accepted*; the result of a hypothesis test is either 'reject the null hypothesis and accept the alternative hypothesis', or 'do not reject the null hypothesis'.

Failure to reject a null hypothesis does not imply that the hypothesis is true; it simply means that there is not enough evidence to establish the alternative.

There is a test, known as the KPSS test, which states the presence of a unit root as the *alternative* hypothesis; rejecting the *null* hypothesis of no unit root *establishes* that there is a unit root.

The Dickey Fuller Test Consider the AR(1) model:

$$X_t = \phi X_{t-1} + \epsilon_t \qquad \{\epsilon_t\} \sim \mathrm{WN}(0, \sigma^2).$$

Subtracting X_{t-1} from both sides gives:

$$\nabla X_t = (\phi - 1) X_{t-1} + \epsilon_t \Rightarrow \nabla X_t = \beta X_{t-1} + \epsilon_t \qquad \{\epsilon_t\} \sim \mathrm{WN}(0, \sigma^2).$$

The Dickey Fuller test simply takes a linear regression of $\{\nabla X_t\}$ against X_{t-1} and estimates the parameter β in the model, with error bounds. The test may also include a constant, and a deterministic drift; using linear regression, assuming $\{\epsilon_t\} \sim \mathrm{IID}N(0, \sigma^2)$, one tests whether the parameter β is significant in either

$$\nabla X_t = \alpha + \beta X_{t-1} + \epsilon_t$$

or

$$\nabla X_t = \alpha_0 + \alpha_1 t + \beta X_{t-1} + \epsilon_t.$$

While standard multiple linear regression techniques may be used, the approach by Dickey and Fuller represents a refinement where the estimates are made in a different way and the distribution of the test statistic $DF_\tau := \frac{\widehat{\beta}}{\mathrm{sd}(\widehat{\beta})}$ turns out not to be exactly t distributed. The distribution is known as the *Dickey–Fuller distribution*.

The tests have low statistical power; they cannot distinguish between a true unit-root ($\beta = 0$) and near unit-root (β close to zero). This is called the 'near observation equivalence' problem.

The Augmented Dickey Fuller Test The testing procedure for the ADF test is the same as for the Dickey–Fuller test but it is applied to the model

$$\nabla X_t = \alpha_0 + \alpha_1 t + \beta X_{t-1} + \delta_1 \nabla X_{t-1} + \cdots + \delta_{p-1} \nabla X_{t-p+1} + \epsilon_t \qquad \{\epsilon_t\} \sim \mathrm{IID}N(0, \sigma^2)$$

The lag length p is determined when applying the test, using standard model building techniques from multiple linear regression analysis. The unit root test is then carried out under the null hypothesis $\beta = 0$ against the alternative hypothesis $\beta < 0$. The test statistic

$$DF_\tau = \frac{\widehat{\beta}}{\text{sd}(\widehat{\beta})}$$

is computed it can be compared to the relevant critical value for the Dickey–Fuller test.

The KPSS Test The KPSS test was introduced by Kwiatkoski, Phillips, Schmidt and Shin in 1992 (Biometrics vol. 54 pp 159 - 178). It is based on the LM (Lagrange Multiplier) test in regression for omitted variables.

Assume that a time series $\{Y_t\}$ can be decomposed into a linear trend ξt, a random walk R_t and a stationary error process X_t:

$$\begin{cases} Y_t = \xi t + R_t + X_t \\ R_t = R_{t-1} + \epsilon_t \qquad \{\epsilon_t\} \sim IIDN(0, \sigma_Z^2) \end{cases}$$

R_0 is fixed. The hypothesis that $Y_t - \xi t$ is stationary is equivalent to the hypothesis that $\sigma^2 = 0$.

For the test statistic, it is assumed that $\{X_t\} \sim IIDN(0, \sigma_X^2)$. Let $(e_t)_{t\geq 1}$ denote the residuals from an OLS regression $Y_t = \beta_0 + \beta_1 t + \epsilon_t$, let $\widehat{\sigma}_e^2$ the estimate of the error variance from this regression and $S_t = \sum_{i=1}^t e_i$. The LM statistic for $Y_1, \ldots, Y_T$ is:

$$LM = \frac{\sum_{t=1}^T S_t^2}{\widehat{\sigma}_e^2}$$

Under the assumption that $\sigma_u^2 = 0$, the distribution (or at least the asymptotic distribution) of $\frac{1}{T^2} \sum_{t=1}^T S_t^2$ may be computed explicitly and $\widehat{\sigma}_e^2 \stackrel{T\to+\infty}{\longrightarrow} \sigma_X^2$.

Testing for unit roots using R The following gives a demonstration of a unit root test. Consider the log series of U.S. quarterly GDP from 1947.I to 2008.IV. The file is found in `q-gdp4708.txt` in the course directory. The data is plotted in Figure 3.1.
The following indicates that the unit root test cannot be rejected. The test used is the *KPSS* test.

```
> q.gdp4708 <- read.table(www, header=T)
> a = ur.kpss(q.gdp4708$xrate,type = "tau")
> summary(a)

#######################
# KPSS Unit Root Test #
#######################

Test is of type: tau with 3 lags.
```

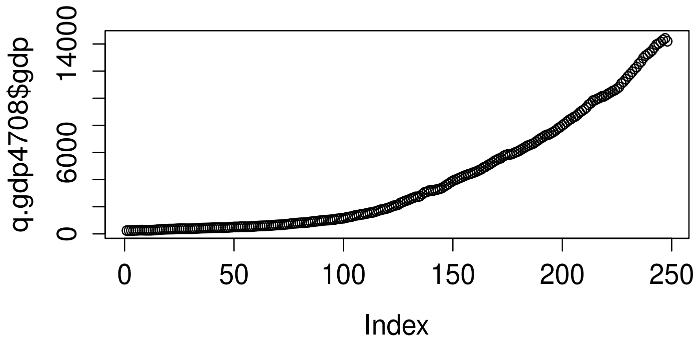

Figure 3.1: US quarterly GDP 1947 - 2008

```
Value of test-statistic is: 0.2444

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

The test statistic is larger than the 1% critical value. We may safely reject the null hypothesis of no unit root and accept the alternative of unit root. In fact, we can see that the unit root has multiplicity 2:

```
> library("forecast")
> auto.arima(q.gdp4708$gdp)
Series: q.gdp4708$gdp
ARIMA(0,2,1)

Coefficients:
         ma1
     -0.6438
s.e.   0.0685

sigma^2 estimated as 1361:  log likelihood=-1236.89
AIC=2477.79   AICc=2477.84   BIC=2484.8
```

3.3 SARIMA Processes

Seasonal series are characterised by a strong serial correlation at the seasonal lag and multiples thereof. Seasonal ARIMA models allow for randomness in the seasonal pattern from one cycle to the next.

Definition 3.12 (The SARIMA$(p,d,q)\times(P,D,Q)_s$ Process). *A process $\{X_t\}$ is said to be a Seasonal ARIMA(p,d,q)×(P,D,Q) process with period s if the differenced process*

$$Y_t := (1-B)^d(1-B^s)^D X_t$$

is a causal ARMA process,

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)\epsilon_t \qquad \{\epsilon_t\} \sim WN(0,\sigma^2)$$

where

$$\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p$$

$$\Phi(z) = 1 - \Phi_1 z^s - \ldots - \Phi_P z^{Ps}$$

$$\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$$

$$\Theta(z) = 1 + \Theta_1 z^s + \ldots + \Theta_Q z^{Qs}.$$

Note that the process $\{Y_t\}$ is causal if and only if both $\phi(z) \neq 0$ and $\Phi(z) \neq 0$ for all $|z| \leq 1$.

Note The SARMA process is a *stationary* process; the mean zero SARMA process satisfies $\mathbb{E}[X_t] \equiv 0$ for all t.

Therefore, the stationary SARMA process is not suitable for the situation where the process has a *deterministic* stationary component (so that $\mathbb{E}[X_t] = s_t$, where s_t is a deterministic periodic function). What is in view here is a process where the *autocovariance* is seasonal.

The SARMA process is therefore not suitable for modelling, for example, a situation where there is a 'January effect', when trade increases in January due to January sales.