4.3 ARMA Processes

Why ARMA models?

- *Wold's decomposition theorem*. Every stationary time series Y can be written as infinite moving average

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i},$$

where μ is deterministic, $\epsilon \sim WN(0, \sigma^2)$ and $\sum_{i=0}^{\infty} \psi_i^2 < \infty$.

- It is impractical to attempt to estimate infinitely many weights $\{\psi_i\}$.
- Thus, the infinite moving average representation of stationary TS is useless in practice, except for certain special cases:
  - **Case 1** - finite order MA models: only finitely many $\{\psi_i\} \neq 0$.
  - **Case 2** - finite order AR models: $\{\psi_i\}$ are generated by finite parameters.
  - **Case 3** - Mixture of finite order AR and MA models (ARMA):
    - it may not always be true that ψ_i has an exponential pattern. In this case, we can increase the order of AR models from AR(1) to AR(p).
    - In some case, it is easier to add an MA part instead of increasing AR order[4].

[4] Refer to Chapter 5 for modeling with ARMA; also see later.

Definition 4.5 (The ARMA(p, q) process)

The process $\{X_t, t \in \mathbb{Z}\}$ is said to be mixed autoregressive moving average process of order (p, q), or ARMA(p, q), if

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q},$$

where $\{Z_t\} \sim WN(0, \sigma^2)$. In operator notation: $\Phi(B)X_t = \Theta(B)Z_t$.

Theorem 4.4

- ARMA(p, q) is *causal* iff all roots of $\Phi(\cdot)$ lie *outside* the unit circle.
- ARMA(p, q) is *invertible* iff all roots of $\Theta(\cdot)$ lie *outside* the unit circle.
- ARMA(p, q) is *causal* and *invertible* iff there are no common factors of $\Phi(y)$ and $\Theta(y)$, and all the roots of $\Phi(\cdot)$ and $\Theta(\cdot)$ lie *outside* the unit circle.

Example 4.5

Consider an ARMA(2,1) process

$$X_t - 0.75X_{t-1} + 0.5625X_{t-2} = Z_t + 1.25Z_{t-1}, \quad \{Z_t\} \sim WN(0, \sigma^2).$$

Verify the causality and invertibility.

- The AR polynomial

$$\Phi(y) = 1 - 0.75y + 0.5625y^2$$

has zeros at $y = \frac{2(1 \pm i\sqrt{3})}{3}$, which lie outside the unit circle. The process is causal.

- The MA polynomial $\Theta(y) = 1 + 1.25y$ has a zero at $y = -0.8$ and hence the process is not invertible.

Parsimonious Convention (1/3)

- The parsimonious convention in ARMA modeling refers to the principle of selecting the simplest model that adequately explains the structure of a time series while avoiding unnecessary complexity and overfitting.
- A more complex ARMA model (with higher-order AR or MA terms) may fit historical data better but at the cost of increased complexity and potential overfitting. A parsimonious model aims to achieve a balance between goodness-of-fit and model simplicity, ensuring better generalizability to new data.

Parsimonious Convention (2/3)

- Akaike Information Criterion (AIC):

$$AIC = -2\ln(L) + 2k,$$

where L is the likelihood of the model and k is the number of estimated parameters. Lower AIC values indicate a better trade-off between fit and complexity.

- Bayesian Information Criterion (BIC):

$$BIC = -2\ln(L) + k\ln(n),$$

where n is the number of observations. Like AIC, lower BIC values indicate a more parsimonious model.

Both AIC and BIC penalize excessive model complexity, guiding the selection of the simplest model that retains predictive power.

Example.
- ARMA(2,1) has an AIC of 150 and a BIC of 160.
- ARMA(3,2) has an AIC of 149 and a BIC of 165.

Parsimonious Convention (3/3)

- Box and Jenkins (1976) argue that an ARMA(p, q) model with small values of p and q will do as well at explaining a process as a high order AR(p) or MA(q) process.
- Experience suggest that ARMA(p, q) process with $p + q \leq 2$ constitute an extremely useful class of parsimonious models for stationary time series data.
- We summarize main features of AR(2), MA(2), and ARMA(1,1) processes.

4.4 Low order models

Low order models - AR(2) (1/5)

AR(2) process: $$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + Z_t$$
$$\Rightarrow \Phi(B)X_t = Z_t, \quad \Phi(z) = 1 - \alpha_1 z - \alpha_2 z^2$$

Case 1: $\boldsymbol{\alpha_1^2 + 4\alpha_2 \geq 0}$.

- $$\Phi(z) = 0 \quad \Leftrightarrow \quad z_1, z_2 = \frac{-\alpha_1 \pm \sqrt{\alpha_1^2 + 4\alpha_2}}{2\alpha_2}.$$
- Then (refer to exercises in tutorial)

$$-2 < \alpha_1 < 2, \quad -1 < \alpha_2 < 1$$

and

$$\alpha_1 + \alpha_2 < 1, \quad \alpha_2 - \alpha_1 < 1.$$

Case 2: $\alpha_1^2 + 4\alpha_2 < 0$.

- the complex roots are given by

$$z_1, z_2 = \frac{-\alpha_1 \pm i\sqrt{|\alpha_1^2 + 4\alpha_2|}}{2\alpha_2}.$$

The radius of the roots is

$$r^2 = \left(\frac{\alpha_1}{2\alpha_2}\right)^2 + \left(\frac{\sqrt{|\alpha_1^2 + 4\alpha_2|}}{2\alpha_2}\right)^2 = -\frac{1}{\alpha_2}.$$

- Let $\cos(\theta) = \frac{\alpha_1}{2\sqrt{-\alpha_2}}$, then

$$z_1 = re^{i\theta}, \qquad z_2 = re^{-i\theta}.$$

- Causality is equivalent to $r > 1 \Leftrightarrow -1 < \alpha_2 < 0$.

Calculations of ACFs for causal AR(2).

- Multiplying X_{t-k} on both sides of $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + Z_t$, and dividing γ_0, we get Yule-Walker equations

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2), \quad k = 1, 2, \ldots. \tag{3}$$

- We know $\rho(0) = 1$, $\rho(1) = \frac{\alpha_1}{1-\alpha_2}$
- Equation (3) is called *linear difference equations*.

Low order models - AR(2) (4/5)

Recall the linear difference equaton

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2), \quad k = 1, 2, \ldots.$$

Its characteristic equation of (3) is $\lambda^2 - \alpha_1 \lambda - \alpha_2 = 0$. Let λ_1, λ_2 be the roots.

Case 1: $\alpha_1^2 + 4\alpha_2 > 0$. In this case, λ_1 and λ_2 are two distinct real roots. The general solution of (3) is given by

$$\rho(k) = A_1 \lambda_1^k + A_2 \lambda_2^k, \quad k = 0, \pm 1, \cdots$$

Case 2: $\alpha_1^2 + 4\alpha_2 < 0$. In this case, λ_1 and λ_2 are two distinct complex roots. The general solution of (3) is given by

$$\rho(k) = r^k \Big(A_1 \cos(\theta k) + A_2 \sin(\theta k) \Big), \quad k = 0, \pm 1, \cdots$$

Case 3: $\alpha_1^2 + 4\alpha_2 = 0$. In this case, $\lambda_1 = \lambda_2 := \lambda$ are two common real roots. The general solution of (3) is given by

$$\rho(k) = (A_1 + A_2 k) \lambda^k, \quad k = 0, \pm 1, \cdots$$

Summary of solutions of Yule-Walker equations.

1. $\alpha_1^2 + 4\alpha_2 > 0$: two distinct real solutions of the characterisitc equation $\lambda_1 \neq \lambda_2$. The general solution of Yule-Walker equations is

$$\rho(k) = A_1\lambda_1^k + A_2\lambda_2^k, \quad k = 0, \pm 1, \cdots$$

2. $\alpha_1^2 + 4\alpha_2 < 0$: two distinct complex solutions of the characteristic equation $\lambda_1, \lambda_2 = re^{\pm i\theta}$. The general solution of Yule-Walker equations is

$$\rho(k) = r^k(A_1\cos(\theta k) + A_2\sin(\theta k)), \quad k = 0, \pm 1, \cdots;$$

3. $\alpha_1^2 + 4\alpha_2 = 0$: two common real solutions of the characteristic equations $\lambda_1 = \lambda_2 := \lambda$. The general solution of Yule-Walker equations is

$$\rho(k) = (A_1 + A_2 k)\lambda^k.$$

Two constants A_1, A_2 can be determined by $\rho(0) = 1$ and $\rho(1) = \frac{\alpha_1}{1-\alpha_2}$.

MA(2) process

$$X_t = Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2}, \quad \text{equivalently } X_t = \Theta(B) Z_t.$$

- For invertibility, roots of $1 + \beta_1 y + \beta_2 y^2 = 0$ stay outside the unit circle.
  If $\beta_1^2 - 4\beta_2 \geq 0$, then

$$\left| \frac{-\beta_1 \pm \sqrt{\beta_1^2 - 4\beta_2}}{2\beta_2} \right| > 1.$$

  If $\beta_1^2 - 4\beta_2 < 0$, then

$$\left(\frac{\beta_1}{2\beta_2} \right)^2 + \left(\frac{\sqrt{|\beta_1^2 - 4\beta_2|}}{2\beta_2} \right)^2 > 1.$$

- For the ACF of a MA(2) process , see Example 4.2.

ARMA(1,1) process

$$X_t - \alpha X_{t-1} = Z_t + \beta Z_{t-1} \quad \Leftrightarrow \quad (1-\alpha B)X_t = (1+\beta B)Z_t.$$

- **Causality**: $|\alpha| < 1$. In this case

$$\begin{aligned} X_t &= \frac{1+\beta B}{1-\alpha B} Z_t = (1+\beta B)\sum_{i=0}^{\infty} \alpha^i B^i Z_t \\ &= \left[1 + \sum_{i=1}^{\infty} \alpha^{i-1}(\alpha+\beta)B^i\right] Z_t \\ &= \sum_{i=0}^{\infty} b_i Z_{t-i}, \end{aligned}$$

  where $b_0 = 1,\ b_i = \alpha^{i-1}(\alpha+\beta), \quad \text{for } i = 1, 2, \ldots$

- **Invertibility**: $|\beta| < 1$.

- **ACF of ARMA(1,1).** Under the condition $|\alpha| < 1$,

$$\gamma(0) = \text{Var}(X_t) = \sigma^2 \sum_{i=0}^{\infty} b_i^2 \quad \text{and} \quad \gamma(k) = \cdots = \sum_{i=0}^{\infty} b_i b_{i+k} \sigma^2.$$

Thus, ACF at lag k follows

$$\rho(k) = \sum_{i=0}^{\infty} b_i b_{i+k} \Big/ \sum_{i=0}^{\infty} b_i^2,$$

with

$$\sum_{i=0}^{\infty} b_i^2 = 1 + (\alpha+\beta)^2 \sum_{i=1}^{\infty} \alpha^{2i-2} = 1 + \frac{(\alpha+\beta)^2}{1-\alpha^2}$$

and

$$\sum_{i=0}^{\infty} b_i b_{i+k} = \alpha^{k-1}(\alpha+\beta) + \frac{(\alpha+\beta)^2 \alpha^k}{1-\alpha^2}.$$

Therefore,

$$\rho(k) = \frac{\alpha^{k-1}(\alpha+\beta)(1+\alpha\beta)}{1+2\alpha\beta+\beta^2}, \quad \text{for } k = 1, 2, \ldots,$$

i.e. exponential decay, $\rho(k) = \alpha\rho(k-1)$, for $k = 2, 3, \ldots$

Example 4.6

For an ARMA(2,1) process

$$Y_t = 0.3Y_{t-1} + 0.2Y_{t-2} + \epsilon_t + \theta\epsilon_{t-1}, \quad \epsilon_t \sim WN(0, \sigma^2)$$

given $\rho_1 = 0.8, \rho_2 = 0.44$. Find ρ_4.

4.5 ARIMA Processes

Definition 4.6 (The ARIMA(p, d, q) process)

If d is a nonnegative integer, then $\{X_t\}$ is an **ARIMA(p,d,q) process**, if $Y_t := (1 - B)^d X_t$ is an ARMA(p, q) process.

Remark.

1) The term "integrated" refers to the process of differencing a time series to achieve stationarity. A time series is said to be integrated of order d (denoted as $I(d)$) if it becomes stationary after being differenced d times.

2). $d = 1 \Rightarrow Y_t = X_t - X_{t-1} \Rightarrow Y \sim \text{ARIMA}(p, 1, q)$.

3). When $d = 0$, X reduces to an ARMA(p, q) process.

Definition 4.7 (The seasonal ARIMA process)

Let Φ, ϕ, Θ and θ be polynomials of order P, p, Q and q, respectively. If X satisfies $\Phi(B^S)\phi(B)\nabla_S^D \nabla^d X_t = \Theta(B^S)\theta(B)\epsilon_t$ for some S, it is called a seasonal ARIMA process with order $(p, d, q) \times (P, D, Q)$ and seasonal period S, written as ARIMA$(p, d, q) \times (P, D, Q)_S$, where $\nabla_S^D = (1 - B^S)^D$ is the seasonal differencing.