Time Series lecture 3

ARMA revisited

Sofia Olhede

EPFL

March 5, 2025

Lecture outline

1. ARMA polynomial notation

2. Wold decomposition

3. Stationarity and invertibility of ARMA processes

ARMA polynomial notation

ARMA recap

- Recall that an ARMA process has the following form

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} - \sum_{k=0}^{q} \theta_k \varepsilon_{t-k}$$

where $\{\varepsilon_t\}$ is a mean-zero white noise process and $\theta_0 = -1$.

- We could rewrite this as

$$\sum_{j=0}^{p} \phi_j X_{t-j} = \sum_{k=0}^{q} \theta_k \varepsilon_{t-k} \tag{3.1}$$

where $\phi_0 = -1$.

- This is elegant, but we can simplify things a little further with the help of the backshift operator.

The backshift operator

Definition 3.1 (The backshift operator)

Define the backshift operator as a map from one time series $\{X_t\}$ to another time series, say $\{y_t\}$ which simply shifts the first series back one step in time. Formally write

$$\{Y_t\} = \mathrm{B}[\{X_t\}]$$

then for all $t \in \mathbb{Z}$ we have

$$Y_t = X_{t-1}.$$

- Often we will use the informal notation $\mathrm{B}X_t$ to refer to $\mathrm{B}[\{X_t\}]$.
- Note that applying the backshift multiple times will shift the series by multiple lags (i.e. the operator B^k shifts by k lags).

MA and the backshift operator

Take the MA(q) example, we can write this as

$$X_t = \sum_{k=0}^{q} \theta_k \varepsilon_{t-k} \tag{3.2}$$

$$= \sum_{k=0}^{q} \theta_k \mathrm{B}^k \varepsilon_t \tag{3.3}$$

$$= \Theta(\mathrm{B})\varepsilon_t \tag{3.4}$$

where Θ is the polynomial given by

$$\Theta(z) = \sum_{k=0}^{q} \theta_k z^k. \tag{3.5}$$

ARMA and the backshift operator

Now we can apply the same trick for the general ARMA(p, q) process, and write

$$\Phi(\mathrm{B})X_t = \Theta(\mathrm{B})\varepsilon_t \tag{3.6}$$

to specify the ARMA process, where

$$\Theta(z) = \sum_{k=0}^{q} \theta_k z^k, \tag{3.7}$$

$$\Phi(z) = \sum_{j=0}^{p} \phi_j z^j. \tag{3.8}$$

Later in the lecture, we shall explore how properties of the polynomials Θ and Φ relate to properties of the resultant process.

Wold decomposition

Linear combinations of white noise

So far, we have seen finite combinations of white noise processes, but we may want to consider something more general:

$$X_t = \sum_{k=-\infty}^{\infty} g_k \varepsilon_{t-k}, \quad \|g\|_2^2 < \infty. \tag{3.9}$$

where $\{g_k\}$ is a real valued sequence and $\{\varepsilon_t\}$ is mean-zero white noise.

- We have for all $t, \tau \in \mathbb{Z}$

$$\mathbb{E}\left[X_t\right] = 0, \tag{3.10}$$

$$\mathrm{Var}\left(X_t\right) = \|g\|_2^2 \,\mathrm{Var}\left(\epsilon_t\right) < \infty, \tag{3.11}$$

$$\mathrm{Cov}\left(X_t, X_{t+\tau}\right) = \sum_{k=-\infty}^{\infty} g_k g_{k+\tau} \,\mathrm{Var}\left(\varepsilon_t\right). \tag{3.12}$$

Definition 3.2 (General linear process)

If in (3.9) we set $g_{-1},\ g_{-2},\ \cdots = 0$, then we obtain a general linear process:

$$X_t = \sum_{k=0}^{\infty} g_k \varepsilon_{t-k}, \quad \|g\|_2^2 < \infty, \tag{3.13}$$

- Now X_t depends only on the past and the present, making this into a causal process.
- The same more general equations for mean and autocovariance apply, so the process is stationary.

Theorem 3.3 (Wold Decomposition Theorem)

Any stationary process $\{X_t\}$ *can be expressed in the form:*

$$X_t = U_t + V_t,$$

with U_t *and* V_t *uncorrelated such that*

- U_t *has a one–sided linear representation*

  $$U_t = \sum_{k=0}^{\infty} g_k \varepsilon_{t-k},$$

  with $g_0 = 1$, $\|g\|_2^2 < \infty$ *and* ε_t *a mean-zero white noise process uncorrelated with* V_t *so that* $\forall s, t \ \mathbb{E}[\varepsilon_s V_t] = 0$. *The sequences* $\{g_u\}$ *and* $\{\varepsilon_t\}$ *are then uniquely determined.*
- V_t *is singular (can be predicted from its own past with no error).*

- To study such processes, we introduce the function $G(z)$:

$$G(z) = \sum_{k=0}^{\infty} g_k z^k,$$

so that $X_t = G(\mathrm{B})\varepsilon_t$.

- We represent $G(z)$ via its Laurent series (a fancy Taylor series) but will first study it as a ratio:

$$G(z) = \frac{G_1(z)}{G_2(z)}.$$

- Note that the roots of $G_1(z)$ are the roots or zeros of $G(z)$ and the roots of $G_2(z)$ are the poles of $G(z)$.
- Call the zeros of $G_2(z)$ $z_1, z_2, \dots z_p$ ordered so that $z_1, z_2, \dots z_k$ are <u>inside</u> the unit circle $|z| = 1$ and $z_{k+1}, \dots, z_p$ are <u>outside</u> the unit circle $|z| = 1$.

- With this specification the Laurent expansion gives us

$$\begin{aligned}\frac{1}{G_2(z)} &= \sum_{j=1}^{p} \frac{A_j}{z - z_j} \\ &= \sum_{j=1}^{k} \frac{A_j}{z} \sum_{l=0}^{\infty} \left(\frac{z_j}{z}\right)^{l} - \sum_{j=k+1}^{p} \frac{A_j}{z_j} \sum_{l=0}^{\infty} \left(\frac{z_j}{z}\right)^{-l}.\end{aligned}$$

This expansion is convergence on $|z| = 1$. We can therefore replace z by B and arrive at

$$\begin{aligned}\frac{1}{G_2(\mathrm{B})}\varepsilon_t &= \left\{ \sum_{j=1}^{k} A_j \mathrm{B}^{-1} \sum_{l=0}^{\infty} z_j^{l} \mathrm{B}^{-l} - \sum_{j=k+1}^{p} A_j \sum_{l=0}^{\infty} z_j^{-(l+1)} \mathrm{B}^{l} \right\} \varepsilon_t \\ &= \left\{ \sum_{j=1}^{k} A_j \sum_{l=0}^{\infty} z_j^{l} \mathrm{B}^{-l-1} - \sum_{j=k+1}^{p} A_j \sum_{l=0}^{\infty} z_j^{-(l+1)} \mathrm{B}^{l} \right\} \varepsilon_t.\end{aligned}$$

- It therefore follows that

$$\frac{1}{G_2(\mathrm{B})}\varepsilon_t = \sum_{j=1}^{k} A_j \sum_{l=0}^{\infty} z_j^l \underbrace{\varepsilon_{t+1+l}}_{\text{future}} - \sum_{j=k+1}^{p} A_j \sum_{l=0}^{\infty} z_j^{-(l+1)} \underbrace{\varepsilon_{t-l}}_{\text{past+now}} .$$

  Hence if all the roots of $G_2(z)$ are outside the unit circle then only past and present values of X_t are involved. Then the general linear process exists.
- Another way of stating this is that $|G(z)| < \infty$ for $|z| \leq 1$. This means that $G(z)$ is analytic inside and on the unit circle.
- If a particular value of ε_t affects X_t and all subsequent X_t then we say this is an innovations outlier.

Stationarity and invertibility of ARMA processes

- Consider the $MA(q)$ model in this setting. Then

$$X_t = \Theta(\mathrm{B})\varepsilon_t = \varepsilon_t - \theta_{1,q}\varepsilon_{t-1} - \cdots - \theta_{q,q}\varepsilon_{t-q}.$$

Thus we have in the general linear process representation:

$$X_t = \Theta(\mathrm{B})\varepsilon_t \ \Leftrightarrow \Theta^{-1}(\mathrm{B})X_t = \varepsilon_t.$$

- Similarly for the AR model we may write

$$\Phi(\mathrm{B})X_t = \varepsilon_t.$$

Here $\Phi(\mathrm{B})$ has a finite order but $\Phi^{-1}(\mathrm{B})$ has an infinite order.

- Invertibility: Consider inverting the general linear process

$$X_t = G(\mathrm{B})\varepsilon_t \Rightarrow G^{-1}(\mathrm{B})X_t = \varepsilon_t.$$

- The expansion of $G^{-1}(\mathrm{B})$ in powers of B gives its AR form provided that $G^{-1}(\mathrm{B})$ admits a power expansion

$$G^{-1}(z) = \sum_{k=0}^{\infty} h_k z^k,$$

  and that must be analytic on $|z| \leq 1$.
- For a general linear process the model is invertible if $|G^{-1}(z)| < \infty$ for $|z| \leq 1$.
- This means all the poles of $G^{-1}(z)$ are outside the unit circle.
- $X_t = G(\mathrm{B})\varepsilon_t$ is the general linear model. If the poles of $G(z)$ are outside the unit circle, then the zeros of $G^{-1}(z)$ are inside the unit circle.

- For the MA(q) process we have $G(\mathrm{B}) = \Theta(\mathrm{B})$.
- For the AR(p) process we have $\Phi(\mathrm{B})X_t = \varepsilon_t$.
- Thus

$$X_t = \Phi^{-1}(\mathrm{B})\varepsilon_t = G(\mathrm{B})\varepsilon_t$$
$$\Rightarrow G(z) = \Phi^{-1}(z).$$

- Thus in this scenario (AR) the requirement for stationarity is that the roots of $\Phi(z)$ are outside the unit disc.
- For the MA(q) process we have

$$X_t = \Theta(\mathrm{B})\varepsilon_t = G(\mathrm{B})\varepsilon_t.$$

Thus since $\Theta(\mathrm{B}) = G(\mathrm{B})$ is a polynomial of finite order with have $|G(z)| < \infty$ as long as all parameters are finite.

Summary of stationarity and invertibility of ARMA models

We can therefore summarise our understanding as follows:

- An AR(p) must have the roots of $\Phi(z)$ outside $|z| = 1$ to be stationary. It is always invertible.
- An MA(q) is always stationary but must have the roots of $\Theta(z)$ outside $|z| = 1$ to be invertible.
- An ARMA(p, q) must have the roots of $\Phi(z)$ outside $|z| = 1$ to be stationary, and must have the roots of $\Theta(z)$ outside $|z| = 1$ to be invertible.

Characteristic polynomials

Definition 3.4

Recall we can write an ARMA model as

$$\Phi(\mathrm{B})X_t = \Theta(\mathrm{B})\varepsilon_t.$$

We call

- $\Phi(z)$ the characteristic polynomial of the autoregressive part,
- $\Theta(z)$ the characteristic polynomial of the moving average part.

In the specific cases of MA and AR models, this will be shortened to characteristic polynomial, i.e.

- For an AR, $\Phi(\mathrm{B})X_t = \varepsilon_t$, the characteristic polynomial is $\Phi(z)$.
- For an MA, $X_t = \Theta(\mathrm{B})\varepsilon_t$ the characteristic polynomial is $\Theta(z)$.

ARMA Example

Consider the following example, let

$$\{\mathrm{I} - \mathrm{B} + \frac{1}{4}\mathrm{B}^2\}X_t = \{\mathrm{I} + \mathrm{B}\}\varepsilon_t. \tag{3.14}$$

Determine the auto-covariance of X_t assuming $\{\varepsilon_t\}$ is white noise.

- We can cross–multiply equation (3.14) by $X_{t-\tau}$ for $\tau \geq 2$. We then arrive at

  $$X_t X_{t-\tau} - X_{t-1}X_{t-\tau} + \frac{1}{4}X_{t-2}X_{t-\tau} = \varepsilon_t X_{t-\tau} + \varepsilon_{t-1}X_{t-\tau}.$$

  Taking expectations we arrive at

  $$\gamma_\tau - \gamma_{\tau-1} + \frac{1}{4}\gamma_{\tau-2} = 0.$$

- This leaves $\tau = 0, 1$ to figure out. This must be done separately.

- If we set $\tau = 0$ then we get

$$X_t^2 - X_{t-1}X_t + \frac{1}{4}X_{t-2}X_t = \varepsilon_t X_t + \varepsilon_{t-1}X_t.$$

Then taking expectations we get that

$$\gamma_0 - \gamma_1 + \frac{1}{4}\gamma_2 = \sigma^2 + 2\sigma^2.$$

Similarly for $\tau = 1$

$$\gamma_1 - \gamma_0 + \frac{1}{4}\gamma_1 = \sigma^2.$$

A general solution will be of the format

$$\gamma_\tau = \{\beta_{10} + \beta_{11}\tau\}2^{-\tau}, \quad \tau > 0,$$

and by using the initial conditions we may recover the constants.