Lecture 4: Lag Operator, MA Models, and AR Models

Raul Riva

FGV EPGE

October, 2025

Intro

Intro

- We will start covering a very useful tool for modelling: ARMA models;
- They combine autoregressive (AR) and moving average (MA) components;
- In a certain sense, they are the most fundamental models for **stationary** time series;
- There will be a neat theorem about this: **Wold's decomposition theorem**;

Intro

- We will start covering a very useful tool for modelling: ARMA models;
- They combine autoregressive (AR) and moving average (MA) components;
- In a certain sense, they are the most fundamental models for **stationary** time series;
- There will be a neat theorem about this: **Wold's decomposition theorem**;
- But first we need to get acquainted with the **lag operator** L;
- Never underestimate this guy!

The Lag Operator

The Lag Operator

- An important vector space for us is the space of sequences;
- Remember that time series are sequences of random variables;

The Lag Operator

- An important vector space for us is the space of sequences;
- Remember that time series are sequences of random variables;
- An *operator* for us is a mapping from sequences to sequences;
- The **lag operator** L generates a new sequence by lagging the original sequence by one period;

$$L(y_t) = y_{t-1}$$

- This is an informal way of saying

$$\{\cdots, y_{t-1}, y_t, y_{t+1}, \cdots\} \mapsto \{\cdots, y_{t-2}, y_{t-1}, y_t, \cdots\}$$

The Lag Operator Algebra

Some useful tricks:

- $L(\alpha y_t) = \alpha y_{t-1} = \alpha L(y_t), \forall \alpha \in \mathbb{R}$
- $L(y_t + z_t) = y_{t-1} + z_{t-1} = L(y_t) + L(z_t)$
- It is a *linear operator* $\implies$ all the machinery from functional analysis applies!

The Lag Operator Algebra

Some useful tricks:

- $L(\alpha y_t) = \alpha y_{t-1} = \alpha L(y_t), \forall \alpha \in \mathbb{R}$
- $L(y_t + z_t) = y_{t-1} + z_{t-1} = L(y_t) + L(z_t)$
- It is a *linear operator* $\implies$ all the machinery from functional analysis applies!

The most important property:

- $L^2(y_t) = L(L(y_t)) = L(y_{t-1}) = y_{t-2}$;
- $L^k(y_t) = y_{t-k}$ for any integer $k > 0$;
- $L^0(y_t) = y_t = I(y_t)$, the identity operator;

The Lag Operator Algebra

Some useful tricks:

- $L(\alpha y_t) = \alpha y_{t-1} = \alpha L(y_t), \forall \alpha \in \mathbb{R}$
- $L(y_t + z_t) = y_{t-1} + z_{t-1} = L(y_t) + L(z_t)$
- It is a *linear operator* $\implies$ all the machinery from functional analysis applies!

The most important property:

- $L^2(y_t) = L(L(y_t)) = L(y_{t-1}) = y_{t-2}$;
- $L^k(y_t) = y_{t-k}$ for any integer $k > 0$;
- $L^0(y_t) = y_t = I(y_t)$, the identity operator;

The polynomial operator $\phi(L) = a_0 I + a_1 L + a_2 L^2 + a_3 L^3 + \ldots + a_p L^p$ is such that:

$$\phi(L) y_t = a_0 y_t + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p}$$

Can we invert an operator?

- Question: for a given operator T, can we find an operator S such that $ST = I$?
- In other words, if I see $T(y_t)$, can I recover the original sequence y_t?

Can we invert an operator?

- Question: for a given operator T, can we find an operator S such that $ST = I$?
- In other words, if I see $T(y_t)$, can I recover the original sequence y_t?
- As you imagine, this depends on the operator...
- Yes, you can invert the lag operator: $L^{-1}(y_t) = y_{t+1}$;
- But not this one (dumb example): $T(y_t) = 0 \implies$ obviously you cannot invert it;

Can we invert an operator?

- Question: for a given operator T, can we find an operator S such that $ST = I$?
- In other words, if I see $T(y_t)$, can I recover the original sequence y_t?
- As you imagine, this depends on the operator...
- Yes, you can invert the lag operator: $L^{-1}(y_t) = y_{t+1}$;
- But not this one (dumb example): $T(y_t) = 0 \implies$ obviously you cannot invert it;

It will be very useful to study the invertibility the polynomials $\phi(L)$:

- Let's start with baby steps: assume $\phi(L) = I - aL$ for some $a \neq 0$;
- Challenge: can you invert this operator?

Can we ever invert $\phi(L)$?

- Inverting this operator means finding another operator $\psi(L)$ such that $\psi(L)\phi(L) = I$;
- Focus on scalar sequences, abuse notation, and write "1" instead of I;
- If L were a number (*it is not a number!!!*), we would like to write $\psi(L) = \frac{1}{1-aL}$;
- If $|aL| < 1$, we would be able to write: $\psi(L) = 1 + aL + a^2L^2 + a^3L^3 + ...$
- The problem with this intuition is that L is not a number!
- What does "$|aL| < 1$" even mean???

Can we ever invert $\phi(L)$?

- Let's explore this intuition a bit more;
- Consider $\psi_n(L) = 1 + aL + a^2L^2 + ... + a_nL^n$. Then:

$$\psi_n(L)\phi(L) = (1 + aL + a^2L^2 + ... + a_nL^n)(1 - aL) = 1 - a^{n+1}L^{n+1}$$

Hence, $\psi_n(L)\phi(L)(y_t) = y_t - a^{n+1}L^{n+1}(y_t) = y_t - a^n y_{t-n}$ for any sequence $\{y_t\}_{t\in\mathbb{Z}}$.

Can we ever invert $\phi(L)$?

- Let's explore this intuition a bit more;
- Consider $\psi_n(L) = 1 + aL + a^2L^2 + ... + a_nL^n$. Then:

$$\psi_n(L)\phi(L) = (1 + aL + a^2L^2 + ... + a_nL^n)(1 - aL) = 1 - a^{n+1}L^{n+1}$$

Hence, $\psi_n(L)\phi(L)(y_t) = y_t - a^{n+1}L^{n+1}(y_t) = y_t - a^n y_{t-n}$ for any sequence $\{y_t\}_{t\in\mathbb{Z}}$.

- If $|a| < 1$, then $\lim_{n\to\infty} a^n = 0$ and $a^n y_{t-n}$ is probably very small;
- Even if y_t is stochastic, but stationary, $a^n y_{t-n}$ converges to zero almost surely;

Can we ever invert $\phi(L)$?

- Let's explore this intuition a bit more;
- Consider $\psi_n(L) = 1 + aL + a^2L^2 + ... + a_nL^n$. Then:

$$\psi_n(L)\phi(L) = (1 + aL + a^2L^2 + ... + a_nL^n)(1 - aL) = 1 - a^{n+1}L^{n+1}$$

Hence, $\psi_n(L)\phi(L)(y_t) = y_t - a^{n+1}L^{n+1}(y_t) = y_t - a^n y_{t-n}$ for any sequence $\{y_t\}_{t\in\mathbb{Z}}$.

- If $|a| < 1$, then $\lim_{n\to\infty} a^n = 0$ and $a^n y_{t-n}$ is probably very small;
- Even if y_t is stochastic, but stationary, $a^n y_{t-n}$ converges to zero almost surely;
- But if $|a| \geq 1$, then $a^n y_{t-n}$ never dies out!
- So, it seems that the invertibility of $\phi(L)$ depends on the value of a...

Theorem 4.40 from Rynne and Youngson (2000)

Theorem

Let B be a Banach space equipped with a certain norm $\|.\|$. For any operator $A : B \to B$, let $\|A\| \equiv \sup_{||x|| \leq 1} \|Ax\|$. If $T : B \to B$ is a bounded linear operator and $\|I - T\| < 1$, then T is invertible with inverse:

$$T^{-1} = \sum_{k=0}^{\infty} (I - T)^k.$$

- The space of paths generated by second-order stationary processes is a Banach space when equipped with the supremum norm $\|.\|_\infty$;
- $\|y\|_\infty = \sup_t \mathbb{E}|y_t|$. Recall that finite variance implies that this is finite;
- The lag operator is linear and bounded:

$$\|L\| = \sup_{||y||_\infty \leq 1} ||Ly||_\infty = \sup_{\{y\}_{t\in\mathbb{Z}} : \sup_t \mathbb{E}|y_t| \leq 1} ||Ly||_\infty = \sup_{\{y\}_{t\in\mathbb{Z}} : \sup_t \mathbb{E}|y_t| \leq 1} \left(\sup_t \mathbb{E}|y_{t-1}| \right) = 1$$

In our example

- $\phi(L) = I - aL$ is linear and bounded (why?);
- $\|I - \phi(L)\| = \|I - I + aL\| = |a|$;
- $\|I - \phi(L)\| < 1 \iff |a| < 1$;

In our example

- $\phi(L) = I - aL$ is linear and bounded (why?);
- $\|I - \phi(L)\| = \|I - I + aL\| = |a|$;
- $\|I - \phi(L)\| < 1 \iff |a| < 1$;

Now we use the theorem to get exactly what intuition that hinted at before:

$$\psi(L) \equiv \phi^{-1}(L) = \sum_{k=0}^{\infty} (aL)^k = I + aL + a^2L^2 + a^3L^3 + \dots$$

In our example

- $\phi(L) = I - aL$ is linear and bounded (why?);
- $\|I - \phi(L)\| = \|I - I + aL\| = |a|$;
- $\|I - \phi(L)\| < 1 \iff |a| < 1$;

Now we use the theorem to get exactly what intuition that hinted at before:

$$\psi(L) \equiv \phi^{-1}(L) = \sum_{k=0}^{\infty} (aL)^k = I + aL + a^2L^2 + a^3L^3 + \dots$$

In terms of notation, everything below is equivalent:

- $\phi(L)^{-1} \equiv \frac{1}{1-aL} \equiv \frac{1}{\phi(L)}$
- Given two polynomials $\phi_1(L)$ and $\phi_2(L)$, the expression $\frac{\phi_1(L)}{\phi_2(L)}$ means *"first, apply the inverse of $\phi_2(L)$, then apply $\phi_1(L)$"*;

Questions?

A More Complicated Example

Let's study a more complicated polynomial of order 2. Assume that:

$$\phi(L) = I - a_1 L - a_2 L^2$$

What conditions on parameters would imply invertibility?

Hint: what does the Fundamental Theorem of Algebra say about a polynomial of order p, in general?

A More Complicated Example

Let's study a more complicated polynomial of order 2. Assume that:

$$\phi(L) = I - a_1 L - a_2 L^2$$

What conditions on parameters would imply invertibility?

Hint: what does the Fundamental Theorem of Algebra say about a polynomial of order p, in general?

Every polynomial of order p has exactly p (maybe complex!) roots. So you can write:

$$\phi(L) = \alpha (L - \lambda_1 I)(L - \lambda_2 I) = (\alpha\lambda_1\lambda_2) \underbrace{\left(I - \frac{1}{\lambda_1} L\right)}_{\equiv \phi_1(L)} \underbrace{\left(I - \frac{1}{\lambda_2} L\right)}_{\equiv \phi_2(L)}$$

for some (maybe complex) numbers α, λ_1, and λ_2.

A More Complicated Example

- $\phi_1(L)$ and $\phi_2(L)$ are both invertible if and only if $|\lambda_1| > 1$ and $|\lambda_2| > 1$. Why?

A More Complicated Example

- $\phi_1(L)$ and $\phi_2(L)$ are both invertible if and only if $|\lambda_1| > 1$ and $|\lambda_2| > 1$. Why?
- Now, let's define $\psi(L) \equiv \frac{1}{\alpha\lambda_1\lambda_2}\phi_1(L)^{-1}\phi_2(L)^{-1}$
- It's clear that $\psi(L)\phi(L) = I$. So, by definition, $\psi(L) = \phi(L)^{-1}$;

A More Complicated Example

- $\phi_1(L)$ and $\phi_2(L)$ are both invertible if and only if $|\lambda_1| > 1$ and $|\lambda_2| > 1$. Why?
- Now, let's define $\psi(L) \equiv \frac{1}{\alpha\lambda_1\lambda_2}\phi_1(L)^{-1}\phi_2(L)^{-1}$
- It's clear that $\psi(L)\phi(L) = I$. So, by definition, $\psi(L) = \phi(L)^{-1}$;
- Conclusion: $\phi(L)$ is invertible if and only if all its roots are *outside the unit circle*;
- Just a fancy way of saying that the modulus of all roots must be strictly greater than 1;
- This generalizes to polynomials of any order p;

What to do in practice?

Given a polynomial $\phi(L) = I - a_1 L - a_2 L^2 - ... - a_p L^p$:

- Focus on its "sister polynomial" $f(x) = 1 - a_1 x - a_2 x^2 - ... - a_p x^p$;
- Find the roots of f (numerically, of course);
- Check the modulus of all roots: if they are all greater than 1, then $\phi(L)$ is invertible.
- Coding tools like Matlab, Python, R, etc all have commands to check for invertibility.

Questions?

MA Models (Finally!!!)

MA Models (Finally!!!)

- A very important building block is the **Moving Average (MA) model**;
- Let ϵ_t be a white noise process with zero mean and variance σ^2;
- The MA model of order q is defined as:

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} = \mu + \sum_{i=0}^{q} \theta_i \epsilon_{t-i}$$

  where μ is the mean of the process and θ_i are the parameters of the model ($\theta_0 \equiv 1$).
- This is a tool to model dependence that will *completely die* after q periods!

MA Models (Finally!!!)

- A very important building block is the **Moving Average (MA) model**;
- Let ϵ_t be a white noise process with zero mean and variance σ^2;
- The MA model of order q is defined as:

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} = \mu + \sum_{i=0}^{q} \theta_i \epsilon_{t-i}$$

  where μ is the mean of the process and θ_i are the parameters of the model ($\theta_0 \equiv 1$).

- This is a tool to model dependence that will *completely die* after q periods!

Notice we can write $y_t = \mu + \theta(L)\epsilon_t$ where $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q$.

Is this process stationary?

- The mean: $\mathbb{E}[y_t] = \mu + \mathbb{E}[\epsilon_t] + \sum_{i=1}^{q} \theta_i \mathbb{E}[\epsilon_{t-i}] = \mu$
- The variance: $\mathsf{Var}(y_t) = \mathsf{Var}(\epsilon_t) + \sum_{i=1}^{q} \theta_i^2 \mathsf{Var}(\epsilon_{t-i}) = \sigma^2(1 + \sum_{i=1}^{q} \theta_i^2)$
- What about autocovariances? What should $\gamma_h \equiv \mathsf{Cov}(y_t, y_{t-h})$ be?

Is this process stationary?

- The mean: $\mathbb{E}[y_t] = \mu + \mathbb{E}[\epsilon_t] + \sum_{i=1}^{q} \theta_i \mathbb{E}[\epsilon_{t-i}] = \mu$
- The variance: $\text{Var}(y_t) = \text{Var}(\epsilon_t) + \sum_{i=1}^{q} \theta_i^2 \text{Var}(\epsilon_{t-i}) = \sigma^2(1 + \sum_{i=1}^{q} \theta_i^2)$
- What about autocovariances? What should $\gamma_h \equiv \text{Cov}(y_t, y_{t-h})$ be?

There will be two cases. First, let $h > q$. Then $\gamma_h = 0$. Why? What's the intuition for this?

Is this process stationary?

- The mean: $\mathbb{E}[y_t] = \mu + \mathbb{E}[\epsilon_t] + \sum_{i=1}^{q} \theta_i \mathbb{E}[\epsilon_{t-i}] = \mu$
- The variance: $\text{Var}(y_t) = \text{Var}(\epsilon_t) + \sum_{i=1}^{q} \theta_i^2 \text{Var}(\epsilon_{t-i}) = \sigma^2(1 + \sum_{i=1}^{q} \theta_i^2)$
- What about autocovariances? What should $\gamma_h \equiv \text{Cov}(y_t, y_{t-h})$ be?

There will be two cases. First, let $h > q$. Then $\gamma_h = 0$. Why? What's the intuition for this?

Now, what about $h \leq q$?

Let's do simple cases

If $q = 1$, then $y_t - \mu = \epsilon_t + \theta_1 \epsilon_{t-1}$. Hence:

$$\gamma_1 = \mathbb{E}[(y_t - \mu)(y_{t-1} - \mu)] = \mathbb{E}[(\epsilon_t + \theta_1 \epsilon_{t-1})(\epsilon_{t-1} + \theta_1 \epsilon_{t-2})] = \sigma^2 \theta_1$$

Let's do simple cases

If $q = 1$, then $y_t - \mu = \epsilon_t + \theta_1 \epsilon_{t-1}$. Hence:

$$\gamma_1 = \mathbb{E}[(y_t - \mu)(y_{t-1} - \mu)] = \mathbb{E}[(\epsilon_t + \theta_1 \epsilon_{t-1})(\epsilon_{t-1} + \theta_1 \epsilon_{t-2})] = \sigma^2 \theta_1$$

If $q = 2$, then $y_t - \mu = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$. Hence:

$$\gamma_1 = \mathbb{E}[(y_t - \mu)(y_{t-1} - \mu)] = \mathbb{E}[(\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})(\epsilon_{t-1} + \theta_1 \epsilon_{t-2} + \theta_2 \epsilon_{t-3})] = \sigma^2(\theta_1 + \theta_2 \theta_1)$$

and very importantly $\gamma_2 \neq 0$ here as well

$$\gamma_2 = \mathbb{E}[(y_t - \mu)(y_{t-2} - \mu)] = \mathbb{E}[(\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})(\epsilon_{t-2} + \theta_1 \epsilon_{t-3} + \theta_2 \epsilon_{t-4})] = \sigma^2 \theta_2$$

General case for γ_h

- If $h \leq q$, then $\gamma_h = \sigma^2 \left(\theta_h + \theta_{h+1}\theta_1 + \theta_{h+2}\theta_2 + \cdots + \theta_q\theta_{q-h}\right)$
- If $h > q$, then $\gamma_h = 0$;

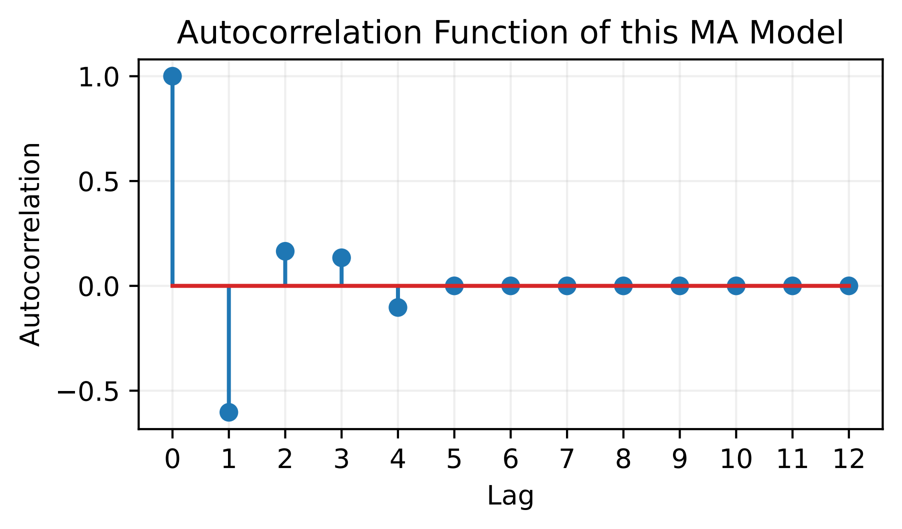
General case for γ_h

- If $h \leq q$, then $\gamma_h = \sigma^2 \left(\theta_h + \theta_{h+1}\theta_1 + \theta_{h+2}\theta_2 + \cdots + \theta_q\theta_{q-h}\right)$
- If $h > q$, then $\gamma_h = 0$;
- True or false? "Any MA(q) is stationary".

General case for γ_h

- If $h \leq q$, then $\gamma_h = \sigma^2 \left(\theta_h + \theta_{h+1}\theta_1 + \theta_{h+2}\theta_2 + \cdots + \theta_q\theta_{q-h}\right)$
- If $h > q$, then $\gamma_h = 0$;
- True or false? "Any MA(q) is stationary".
- The case $q = \infty$ is actually well-defined. In that case we write $y_t = \mu + \sum_{i=0}^{\infty} \theta_i \epsilon_{t-i}$.
- This is a stationary process as long as the variance of y_t is finite
- This is ensured by the following condition: $\sum_{i=0}^{\infty} \theta_i^2 < \infty$.
- Under this condition, any MA(q) process, even with $q = \infty$, is stationary!
- It's also common to focus on autocorrelations, which are defined as $\rho_h \equiv \frac{\gamma_h}{\gamma_0}$.

Quick Visualization ($\theta_1 = -0.8, \theta_2 = 0.5, \theta_3 = 0.1, \theta_4 = -0.2$)



AR Models

AR Models

- The MA-type dependence might not be suitable for all time series data;
- Example: effects of shocks might be long-lasting and never approach zero;
- In such cases, we can use **Autoregressive (AR) models**;

AR Models

- The MA-type dependence might not be suitable for all time series data;
- Example: effects of shocks might be long-lasting and never approach zero;
- In such cases, we can use **Autoregressive (AR) models**;

An AR(p) process satisfies the following dynamics:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

AR Models

- The MA-type dependence might not be suitable for all time series data;
- Example: effects of shocks might be long-lasting and never approach zero;
- In such cases, we can use **Autoregressive (AR) models**;

An AR(p) process satisfies the following dynamics:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

- In the lag operator notation:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i L^i y_t + \epsilon_t \implies (1 - \sum_{i=1}^{p} \phi_i L^i) y_t = \mu + \epsilon_t \implies \Phi(L) y_t = \mu + \epsilon_t$$

where $\Phi(L) \equiv I - \sum_{i=1}^{p} \phi_i L^i$;

Is this process stationary?

- At this point, we are already good friends with $\Phi(L)$;
- This is the same type of polynomial we talked about!
- What are the conditions for the invertibility of $\Phi(L)$?

Is this process stationary?

- At this point, we are already good friends with $\Phi(L)$;
- This is the same type of polynomial we talked about!
- What are the conditions for the invertibility of $\Phi(L)$?
- **The roots of** $f(x) = 1 - a_1 x - a_2 x^2 - \cdots - a_p x^p$ **should lie outside the unit circle**;
- If we can invert $\Phi(L)$, we already know that $\Phi(L)$ will have infinite terms;
- If $\Phi^{-1}(L)$ exists, we can write:

$$y_t = \Phi(L)^{-1}\mu + \Phi(L)^{-1}\epsilon_t$$

Is this process stationary?

- At this point, we are already good friends with $\Phi(L)$;
- This is the same type of polynomial we talked about!
- What are the conditions for the invertibility of $\Phi(L)$?
- **The roots of** $f(x) = 1 - a_1 x - a_2 x^2 - \cdots - a_p x^p$ **should lie outside the unit circle**;
- If we can invert $\Phi(L)$, we already know that $\Phi(L)$ will have infinite terms;
- If $\Phi^{-1}(L)$ exists, we can write:

$$y_t = \Phi(L)^{-1}\mu + \Phi(L)^{-1}\epsilon_t$$

But $\Phi(L)^{-1}\mu$ is just a number. And $\Phi(L)^{-1}\epsilon_t$ is an MA(∞) process!

Is this process stationary?

- At this point, we are already good friends with $\Phi(L)$;
- This is the same type of polynomial we talked about!
- What are the conditions for the invertibility of $\Phi(L)$?
- **The roots of** $f(x) = 1 - a_1 x - a_2 x^2 - \cdots - a_p x^p$ **should lie outside the unit circle**;
- If we can invert $\Phi(L)$, we already know that $\Phi(L)$ will have infinite terms;
- If $\Phi^{-1}(L)$ exists, we can write:

$$y_t = \Phi(L)^{-1}\mu + \Phi(L)^{-1}\epsilon_t$$

But $\Phi(L)^{-1}\mu$ is just a number. And $\Phi(L)^{-1}\epsilon_t$ is an MA(∞) process!

(You will prove in the problem set that this specific MA process is stationary!)

What are the moments?

Consider the AR(1) process $y_t = \mu + \phi_1 y_{t-1} + \epsilon_t$ and assume $|\phi_1| < 1$.

- First: is this guy stationary?

What are the moments?

Consider the AR(1) process $y_t = \mu + \phi_1 y_{t-1} + \epsilon_t$ and assume $|\phi_1| < 1$.

- First: is this guy stationary?

Let's compute its mean:

$$\mathbb{E}(y_t) = \mu + \phi_1 \mathbb{E}(y_{t-1}) + \mathbb{E}(\epsilon_t) = \mu + \phi_1 \mathbb{E}(y_{t-1})$$

Since $\mathbb{E}(y_t) = \mathbb{E}(y_{t-1})$ (why?), we have that $\mathbb{E}(y_t) = \frac{\mu}{1-\phi_1}$.

What are the moments?

Consider the AR(1) process $y_t = \mu + \phi_1 y_{t-1} + \epsilon_t$ and assume $|\phi_1| < 1$.

- First: is this guy stationary?

Let's compute its mean:

$$\mathbb{E}(y_t) = \mu + \phi_1 \mathbb{E}(y_{t-1}) + \mathbb{E}(\epsilon_t) = \mu + \phi_1 \mathbb{E}(y_{t-1})$$

Since $\mathbb{E}(y_t) = \mathbb{E}(y_{t-1})$ (why?), we have that $\mathbb{E}(y_t) = \frac{\mu}{1-\phi_1}$.

Now, we analyze the variance using a similar trick:

$$\text{Var}(y_t) = \text{Var}(\phi_1 y_{t-1} + \epsilon_t) = \phi_1^2 \text{Var}(y_{t-1}) + \text{Var}(\epsilon_t) = \phi_1^2 \text{Var}(y_t) + \sigma^2 \implies \text{Var}(y_t) = \frac{\sigma^2}{1 - \phi_1^2}$$

What are the moments:

- The first autocovariance is:

$$\gamma_1 = \mathsf{Cov}(y_t, y_{t-1}) = \mathsf{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-1}) = \phi_1 \mathsf{Var}(y_{t-1}) = \phi_1 \mathsf{Var}(y_t)$$

What are the moments:

- The first autocovariance is:

$$\gamma_1 = \text{Cov}(y_t, y_{t-1}) = \text{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-1}) = \phi_1 \text{Var}(y_{t-1}) = \phi_1 \text{Var}(y_t)$$

- The second autocovariance is:

$$\gamma_2 = \text{Cov}(y_t, y_{t-2}) = \text{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-2}) = \phi_1 \text{Cov}(y_{t-1}, y_{t-2}) = \phi_1 \gamma_1 = \phi_1^2 \text{Var}(y_t)$$

What are the moments:

- The first autocovariance is:

$$\gamma_1 = \text{Cov}(y_t, y_{t-1}) = \text{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-1}) = \phi_1 \text{Var}(y_{t-1}) = \phi_1 \text{Var}(y_t)$$

- The second autocovariance is:

$$\gamma_2 = \text{Cov}(y_t, y_{t-2}) = \text{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-2}) = \phi_1 \text{Cov}(y_{t-1}, y_{t-2}) = \phi_1 \gamma_1 = \phi_1^2 \text{Var}(y_t)$$

- A similar argument will establish that, in general, $\gamma_h = \phi \gamma_{h-1} \implies \gamma_h = \phi^h \text{Var}(y_t)$.

What are the moments:

- The first autocovariance is:

$$\gamma_1 = \text{Cov}(y_t, y_{t-1}) = \text{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-1}) = \phi_1 \text{Var}(y_{t-1}) = \phi_1 \text{Var}(y_t)$$

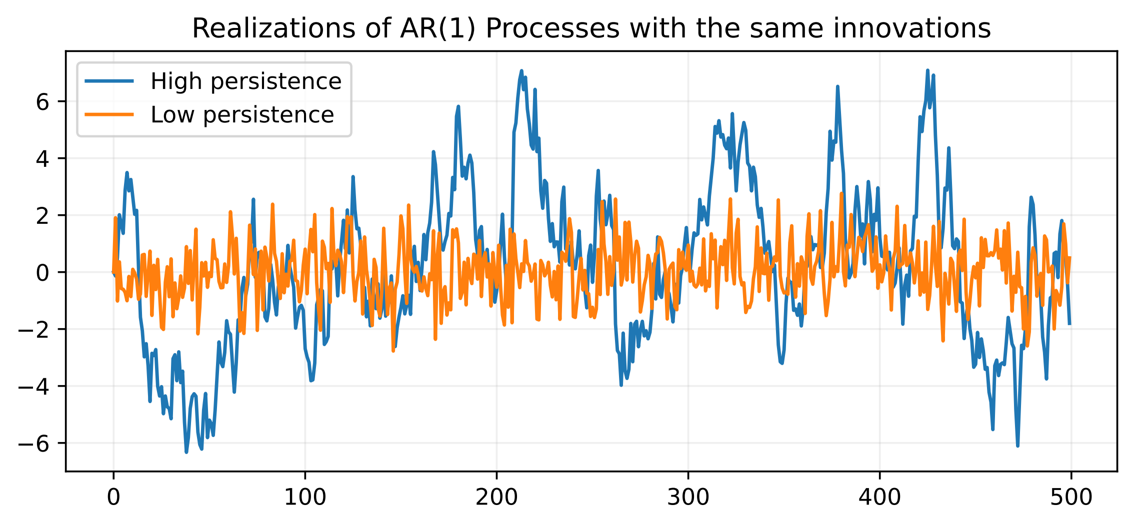
- The second autocovariance is:

$$\gamma_2 = \text{Cov}(y_t, y_{t-2}) = \text{Cov}(\mu + \phi_1 y_{t-1} + \epsilon_t, y_{t-2}) = \phi_1 \text{Cov}(y_{t-1}, y_{t-2}) = \phi_1 \gamma_1 = \phi_1^2 \text{Var}(y_t)$$

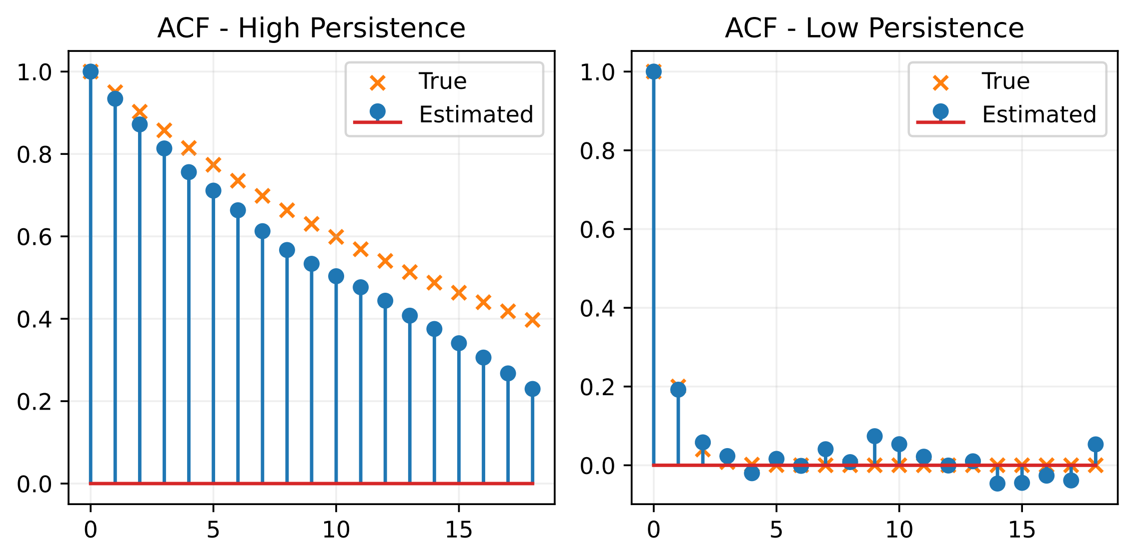
- A similar argument will establish that, in general, $\gamma_h = \phi \gamma_{h-1} \implies \gamma_h = \phi^h \text{Var}(y_t)$.

How is this different from the MA(1) case?

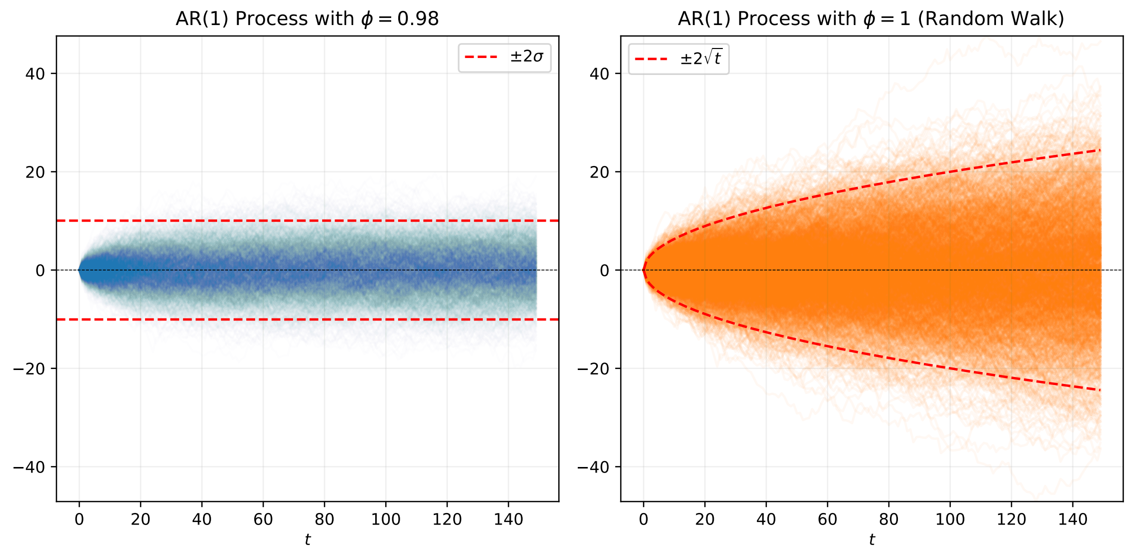
Quick Visualization



Quick Visualization of Autocorrelation Functions $\rho_h \equiv \gamma_h / \gamma_0$



What happens when $\phi_1 = 1$?



The General AR(p) Case

- Consider the general AR(p) process $y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$;
- Let's assume it is stationary;
- The mean is easy, it follows the same tricks as before: $\mathbb{E}(y_t) = \frac{\mu}{1-\phi_1-\phi_2 \ldots - \phi_p}$

The General AR(p) Case

- Consider the general AR(p) process $y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$;
- Let's assume it is stationary;
- The mean is easy, it follows the same tricks as before: $\mathbb{E}(y_t) = \frac{\mu}{1-\phi_1-\phi_2 ...-\phi_p}$

The variance is more involved:

$$\begin{aligned}
\gamma_0 &= Cov(y_t, y_t) \\
&= Cov(\mu + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \epsilon_t, y_t) \\
&= \phi_1 \gamma_1 + \phi_2 \gamma_2 + ... + \phi_p \gamma_p + \sigma^2
\end{aligned}$$

The General AR(p) Case

- Consider the general AR(p) process $y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$;
- Let's assume it is stationary;
- The mean is easy, it follows the same tricks as before: $\mathbb{E}(y_t) = \frac{\mu}{1-\phi_1-\phi_2 ... -\phi_p}$

The variance is more involved:

$$\begin{aligned}
\gamma_0 &= Cov(y_t, y_t) \\
&= Cov(\mu + \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + \epsilon_t, y_t) \\
&= \phi_1 \gamma_1 + \phi_2 \gamma_2 + ... + \phi_p \gamma_p + \sigma^2
\end{aligned}$$

In general, recalling that $\gamma_h = \gamma_{-h}$, we will have (please verify this at home):

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + ... + \phi_p \gamma_{h-p}, \qquad h = 1, 2, 3, ...$$

This generates a system of equations that can be solved recursively (a computer will do it).

Questions?

The End

References

- Chapter 2 from Hamilton's book for the Lag Operator;
- Chapter 3 from Hamilton's book for the definition of AR and MA models;