Stochastic Processes and Financial Mathematics (37363)

Chapter 8

ARMA processes

Juri Hinz
Modifications by Scott Alexander

School of Mathematical and Physical Sciences, UTS

Autumn 2025

Chapter outline

Topics:

- Definitions
- Motivation
- Filters
- Calculus of shift operator
- Stationarity
- Causality and invertibility
- Tesla example – modelling with ARMA(p, q)
- Non-stationarity
- Tesla example – modelling with ARIMA(p, d, q)

Defintions

We begin by defining the processes we will consider in this chapter.

First is the AR(p) process, i.e. **autoregressive process of order $\boldsymbol{p}$**.

Definition 1 (AR(p) process)

An AR(p) process $X = (X_t)_{t\in\mathbb{Z}}$ is a weakly-stationary process described as the solution of

$$X_t = c + \sum_{j=1}^{p} \phi_j X_{t-j} + Z_t, \quad t \in \mathbb{Z},$$

where $c, \phi_j \in \mathbb{R}$ and $Z = (Z_t)_{t\in\mathbb{Z}}$ is a zero-mean white noise process with variance σ^2.

Note that this process is not weakly-stationary for all values of ϕ_j – we will establish conditions under which it is.

Defintions

Next is the MA(q) process, i.e. **moving-average process of order q**.

Definition 2 (MA(q) process)

A MA(q) process $X = (X_t)_{t\in\mathbb{Z}}$ is weakly-stationary process described by

$$X_t = \mu_X + Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j}, \quad t \in \mathbb{Z},$$

where $\mu_X, \theta_j \in \mathbb{R}$ and $Z = (Z_t)_{t\in\mathbb{Z}}$ is a zero-mean white noise process with variance σ^2.

Unlike the AR(p) process, an MA(q) process is guaranteed to be weakly-stationary for all θ_j as it is finite sum of independent, zero-mean RVs with the same variance.

Defintions

If we combine an AR(p) and a MA(q) processes together we obtain an ARMA(p, q) model, i.e. **autoregressive-moving-average process of order (p, q)**.

Definition 3 (ARMA(p, q) process)

An ARMA(p, q) process is a weakly-stationary process $X = (X_t)_{t \in \mathbb{Z}}$ described as the solution of

$$X_t = c + \sum_{j=1}^{p} \phi_j X_{t-j} + Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j}, \quad t \in \mathbb{Z},$$

where $c, \phi_j, \theta_j \in \mathbb{R}$ and $Z = (Z_t)_{t \in \mathbb{Z}}$ is a zero-mean white noise process with variance σ^2.

Like the AR(p) process, the ARMA(p, q) process is not weakly-stationary for all choices of ϕ_j – we will establish conditions under which it is.

To obtain stationarity conditions we simplify notation by taking $c = 0$.

Motivation

Autoregressive-moving average (ARMA) processes provide a sufficiently large class of stationary processes which are popular in applications due to calculational tractability.

Using ARMA processes for modeling, one benefits from their rich theory and wide range of software applications.

Moreover, the framework of ARMA processes can be extended to describe non-stationary situations (which gives ARIMA, SARIMA, etc. process types).

For many applications, an attempt to describe historical data by way of a stationary time series could start with an appropriate ARMA model.

Terminology. In what follows we will use the term “stationary” in place of “weakly-stationary”.

Motivation

Consider now AR(1) process

$$X_t - \phi X_{t-1} = Z_t, \quad t \in \mathbb{Z}, \tag{1}$$

with given white noise process $Z = (Z_t)_{t \in \mathbb{Z}}$.

For $|\phi| < 1$ there exists a stationary solution to this given by

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}, \quad t \in \mathbb{Z}, \tag{2}$$

which we see can be interpreted as a MA(∞) process.

The form of this can be seen through the back-substitution

$$\begin{aligned} X_t &= \phi X_{t-1} + Z_t \\ &= \phi(\phi X_{t-2} + Z_{t-1}) + Z_t = \phi^2 X_{t-2} + \phi Z_{t-1} + Z_t \\ &= \phi^2(\phi X_{t-3} + Z_{t-2}) + \phi Z_{t-1} + Z_t = \phi^3 X_{t-3} + \phi^2 Z_{t-2} + \phi Z_{t-1} + Z_t \\ &= \cdots . \end{aligned}$$

Motivation

To see that the process is stationary for $|\phi| < 1$ note that

$$E[X_t] = E\left[\sum_{j=0}^{\infty} \phi^j Z_{t-j}\right] = \sum_{j=0}^{\infty} \phi^j E[Z_{t-j}] = 0$$

and

$$\begin{aligned}
\text{cov}(X_t, X_{t+h}) &= \text{cov}\left(\sum_{j=0}^{\infty} \phi^j Z_{t-j}, \sum_{k=0}^{\infty} \phi^k Z_{t+h-k}\right) \\
&= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \phi^j \phi^k \,\text{cov}(Z_{t-j}, Z_{t+h-k}) = \sum_{j=0}^{\infty} \phi^j \phi^{j+h} \,\text{cov}(Z_{t-j}, Z_{t-j}) \\
&= \sigma^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+h} = \sigma^2 \frac{\phi^h}{1-\phi^2}
\end{aligned}$$

which is a function of the time step $h \in \{0, 1, 2, \ldots\}$.

Note that the interchange of summation and expectation/covariance should be justified.

Motivation

Using powers B^j, $j \in \mathbb{Z}$, of the back-shift operator acting on each time series $Y = (Y_t)_{t\in\mathbb{Z}}$ as

$$(B^j Y)_t = Y_{t-j}, \quad t \in \mathbb{Z},$$

we can formally write the equation (1) and its solution $X = (X_t)_{t\in\mathbb{Z}}$ from (2) as

$$(1 - \phi B)X = Z, \quad X = \sum_{j=0}^{\infty} (\phi B)^j Z.$$

This result is appealing, since formally we have

$$(1 - \phi B)^{-1} = \sum_{j=0}^{\infty} (\phi B)^j.$$

This observation shows that there could be a useful *calculus of shift operator*.

Motivation

The ARMA(p, q) process with $c = 0$

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \quad (3)$$

can be re-written using the back-shift operator as

$$\begin{aligned} \underbrace{1X_t - (\phi_1 B^1 X)_t - \cdots - (\phi_p B^p X)_t}_{(\phi(B)X)_t} &= \underbrace{1Z_t + (\theta_1 B^1 Z)_t + \cdots + (\theta_q B^q Z)_t}_{(\theta(B)Z)_t} \\ \phi(B)X &= \theta(B)Z \quad (4) \end{aligned}$$

where ϕ and θ are polynomials

$$\begin{aligned} \phi(z) &= 1 - \phi_1 z - \cdots - \phi_p z^p \\ \theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q \end{aligned}$$

and $z \in \mathbb{C}$.

Motivation

Following intuition, we may search for the solution to

$$\phi(B)X = \theta(B)Z$$

in the form

$$X = \frac{\theta(B)}{\phi(B)}Z,$$

in the sense that

$$X_t = \sum_{j\in\mathbb{Z}} \psi_j (B^j Z)_t = \sum_{j\in\mathbb{Z}} \psi_j Z_{t-j}, \quad t \in \mathbb{Z},$$

where the coefficients $(\psi_j)_{j\in\mathbb{Z}}$ are obtained from the power series expansion

$$\frac{\theta(z)}{\phi(z)} = \sum_{j\in\mathbb{Z}} \psi_j z^j.$$

Motivation

In what follows, we

- introduce calculus of shift operators
- address ARMA processes as stationary solutions to (3)
- using shift calculus, discuss existence and uniqueness of stationary solutions to (4).

For stationarity, the roots of the denominator $\phi(z)$ will be crucial in this context.

Filters

A linear filtering procedure assigns to a stochastic process $X = (X_t)_{t\in\mathbb{Z}}$ a new (filtered) process $Y = (Y_t)_{t\in\mathbb{Z}}$ as

$$Y_t = \sum_{j\in\mathbb{Z}} \psi_j X_{t-j}, \quad t \in \mathbb{Z}\,.$$

As presented, this is an infinite filter and the filtered process at time t depends not only on all past and present values of X, but all future values as well (we will re-visit this later).

To avoid convergence problems, we can suppose that the sum is finite, which is the case if

$$\#\{j : \psi_j \neq 0\} < \infty,$$

i.e. if there are a finite number of terms in the filter.

It turns out that this sort of filtering makes sense also for infinite coefficient sets and exhibits an important property: filtering of a stationary process gives another stationary process.

Filters

Proposition 1 (filtered stationary process)

Let $X = (X_t)_{t\in\mathbb{Z}}$ be a zero-mean stationary process with covariance function γ_X and suppose that filter coefficients are absolutely summable, i.e. $\sum_{t\in\mathbb{Z}} |\psi_j| < \infty$.

Then

$$Y_t = \sum_{j\in\mathbb{Z}} \psi_j X_{t-j}, \quad t \in \mathbb{Z},$$

defines a zero-mean stationary process $Y = (Y_t)_{t\in\mathbb{Z}}$ with covariance function

$$\gamma_Y(h) = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \psi_j \psi_k \gamma_X(h + j - k) \quad h \in \mathbb{Z}\,.$$

If X is white noise of variance σ^2, then the covariance function is

$$\gamma_Y(h) = \sigma^2 \sum_{j\in\mathbb{Z}} \psi_j \psi_{j+h} \qquad h \in \mathbb{Z}\,. \tag{5}$$

Proof.
The mean is time-constant since

$$E[Y_t] = E\left[\sum_{j\in\mathbb{Z}} \psi_j X_{t-j}\right] = \sum_{j\in\mathbb{Z}} \psi_j E[X_{t-j}] = 0, \quad t \in \mathbb{Z},$$

as X is a zero-mean process.

The auto-covariance

$$\begin{aligned}\operatorname{cov}(Y_t, Y_{t+h}) &= \operatorname{cov}\left(\sum_{j\in\mathbb{Z}} \psi_j X_{t-j}, \sum_{k\in\mathbb{Z}} \psi_k X_{t+h-k}\right) \\ &= \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \psi_j \psi_k \operatorname{cov}(X_{t-j}, X_{t+h-k}) \\ &= \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \psi_j \psi_k \gamma_X(h+j-k).\end{aligned}$$

Thus the process Y is stationary as it's covariance function is a weighted sum of the covariance function of X and doesn't depend on t. This result can be strengthened to strictly-stationary if X is Gaussian.

Filters

If X is white noise process with variance σ^2 then

$$
\begin{aligned}
\mathrm{cov}(Y_t, Y_{t+h}) &= \mathrm{cov}\left(\sum_{j\in\mathbb{Z}} \psi_j X_{t-j}, \sum_{k\in\mathbb{Z}} \psi_k X_{t+h-k}\right) \\
&= \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \psi_j \psi_k \,\mathrm{cov}(X_{t-j}, X_{t+h-k}) \\
&= \sum_{j\in\mathbb{Z}} \psi_j \psi_{j+h} \,\mathrm{cov}(X_{t-j}, X_{t-j}) \\
&= \sigma^2 \sum_{j\in\mathbb{Z}} \psi_j \psi_{j+h}
\end{aligned}
$$

as $\mathrm{cov}(X_{t-j}, X_{t+h-k}) = 0$ when

$$t - j \neq t + h - k \Rightarrow k \neq j + h$$

and $\mathrm{cov}(X_{t-j}, X_{t-j}) = \mathrm{var}(X_{t-j}) = \sigma^2$ otherwise.

Calculus of shift operator

Consider a filter with absolutely summable filter coefficients $(\psi_j)_{j\in\mathbb{Z}}$ and corresponding filter function

$$\psi(z) = \sum_{j\in\mathbb{Z}} \psi_j z^j$$

defined as a power series which converges at least on the unit circle in the complex plane

$$z \in \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}.$$

To see the series converges (at least) on the complex unit circle note that

$$\sum_{j\in\mathbb{Z}} |\psi_j z^j| = \sum_{j\in\mathbb{Z}} |\psi_j||z^j| = \sum_{j\in\mathbb{Z}} |\psi_j||1| = \sum_{j\in\mathbb{Z}} |\psi_j| < \infty$$

as the coefficients $(\psi_j)_{j\in\mathbb{Z}}$ are assumed to be absolutely summable.

Calculus of shift operator

The action of the filter on a stationary time series X can be written as

$$Y_t = \sum_{j\in\mathbb{Z}} \psi_j X_{t-j} = \sum_{j\in\mathbb{Z}} \psi_j (B^j X)_t = \Bigg(\underbrace{\Big(\sum_{j\in\mathbb{Z}} \psi_j B^j\Big)}_{\psi(B)} X\Bigg)_t$$

$$= (\psi(B)X)_t$$

where two last terms are understood symbolically.

This suggests to write the filter in terms of the filter function ψ.

This symbolic calculus is advantageous, because composition of filters corresponds to the product of their filter functions, as shown by the next proposition.

Calculus of shift operator

Proposition 2 (shift calculus for composition of filters)

For any function ψ on $\mathbb{T}$ with absolutely summable coefficients in the expansion

$$\psi(z) = \sum_{j\in\mathbb{Z}} \psi_j z^j, \quad z \in \mathbb{T}, \tag{6}$$

agree to write

$$(\psi(B)X)_t = \sum_{j\in\mathbb{Z}} \psi_j (B^j X)_t = \sum_{j\in\mathbb{Z}} \psi_j X_{t-j}, \quad t \in \mathbb{Z}, \tag{7}$$

for each stationary process $X = (X_t)_{t\in\mathbb{Z}}$.

Then for another function ψ' on $\mathbb{T}$ with absolutely summable coefficients

$$\psi'(B)\big(\psi(B)X\big) = \big((\psi'\psi)(B)\big)X \equiv \big(\psi'(B)\psi(B)\big)X.$$

Calculus of shift operator

Proof.

Define the filters $\psi(B)$, $\psi'(B)$ and $(\psi'\psi)(B)$ using the filter functions

$$\psi(z) = \sum_{j \in \mathbb{Z}} \psi_j z^j, \quad \psi'(z) = \sum_{j \in \mathbb{Z}} \psi'_j z^j$$

and

$$(\psi'\psi)(z) \equiv \psi'(z)\psi(z) = \Big(\sum_{j \in \mathbb{Z}} \psi'_j z^j \Big) \Big(\sum_{k \in \mathbb{Z}} \psi_k z^k \Big)$$
$$= \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \psi'_j \psi_k z^j z^k = \sum_{l \in \mathbb{Z}} \underbrace{\Bigg(\sum_{j+k=l} \psi'_j \psi_k \Bigg)}_{\psi''_l} z^l$$

which converges for each $z \in \mathbb{T}$.

Note that with $j + k = l$

$$z^j z^k = (e^{i \arg(z)})^j (e^{i \arg(z)})^k = e^{i \arg(z)(j+k)} = e^{i \arg(z) l} = (e^{i \arg(z)})^l = z^l.$$

Calculus of shift operator

Now, for stationary process X the composition

$$\begin{aligned}
\big(\psi'(B)(\psi(B)X)\big)_t &= \sum_{j\in\mathbb{Z}} \psi'_j \big(\psi(B)X\big)_{t-j} = \sum_{j\in\mathbb{Z}} \psi'_j \Bigg(\sum_{k\in\mathbb{Z}} \psi_k X_{t-j-k} \Bigg) \\
&= \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \psi'_j \psi_k X_{t-j-k} = \sum_{l\in\mathbb{Z}} \underbrace{\Bigg(\sum_{j+k=l} \psi'_j \psi_k \Bigg)}_{\psi''_l} X_{t-l} \\
&= \big((\psi'\psi)(B)X\big)_t
\end{aligned}$$

for all $t \in \mathbb{Z}$.

Exercise.
Justify the changes of summation and ordering the calculations above.

Solution. As class work.

Stationarity

The next theorem establishes that there is a unique stationary solution to the ARMA(p, q) equation (3).

Theorem 1 (Stationarity)

There exists a unique stationary solution to the ARMA(p, q) equation

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

if

$$\phi(z) \neq 0 \quad \text{for all} \quad z \in \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$$

where

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

Proof.

Existence. If ϕ has no root on the unit circle $\mathbb{T}$, then

$$\frac{\theta(z)}{\phi(z)} = \frac{1 + \theta_1 z + \cdots + \theta_q z^q}{1 - \phi_1 z - \cdots - \phi_p z^p}$$

admits a power series representation.

More precisely, there exists an $r > 1$ such that

$$\frac{\theta(z)}{\phi(z)} = \sum_{j \in \mathbb{Z}} \psi_j z^j.$$

for all $z \in \mathbb{C}$ with $r^{-1} < |z| < r$, which is a Laurent series from complex analysis.

Stationarity

In particular,

$$\frac{\theta(z)}{\phi(z)} = \sum_{j\in\mathbb{Z}} \psi_j z^j$$

converges absolutely for each $z \in \mathbb{T}$ and therefore

$$\sum_{j\in\mathbb{Z}} |\psi_j z^j| < \infty,$$

thus the filter coefficients are absolutely summable as

$$\sum_{j\in\mathbb{Z}} |\psi_j| = \sum_{j\in\mathbb{Z}} |\psi_j||1| = \sum_{j\in\mathbb{Z}} |\psi_j||z^j| = \sum_{j\in\mathbb{Z}} |\psi_j z^j| < \infty.$$

Then using Proposition in 1, the process

$$X_t = \sum_{j\in\mathbb{Z}} \psi_j Z_{t-j}, \quad t \in \mathbb{Z},$$

is well-defined and is stationary (the white noise process Z is stationary).

Stationarity

According to (6) and (7) we write

$$X_t = \sum_{j \in \mathbb{Z}} \psi_j Z_{t-j} = \big(\psi(B)Z\big)_t = \left(\frac{\theta(B)}{\phi(B)}Z\right)_t \equiv \left(\frac{\theta}{\phi}(B)Z\right)_t$$

using notation from Proposition 2, or

$$X = \frac{\theta}{\phi}(B)Z.$$

Then using the shift calculus of Proposition 2 we have

$$\phi(B)X = \phi(B)\left(\frac{\theta}{\phi}(B)Z\right) = \left(\phi\frac{\theta}{\phi}\right)(B)Z = \theta(B)Z$$

which is the ARMA(p, q) equation (4) with parameters $(\phi_i)_{i=1}^p$ and $(\theta_j)_{j=1}^q$, σ^2.

Stationarity

Uniqueness. Suppose ϕ has no root on the unit circle $\mathbb{T}$ and $X = (X_t)_{t\in\mathbb{Z}}$ is an arbitrary ARMA(p, q) process with parameters $(\phi_i)_{i=1}^{p}$ and $(\theta_j)_{j=1}^{q}$, i.e. X is stationary and

$$\phi(B)X = \theta(B)Z.$$

Using Laurent series expansion of ϕ^{-1}, we find

$$\frac{1}{\phi}(z) = \sum_{j\in\mathbb{Z}} \psi_j z^j$$

converges absolutely for each $z \in \mathbb{T}$.

Stationarity

Now because

$$\sum_{j \in \mathbb{Z}} |\psi_j| < \infty,$$

on $z \in \mathbb{T}$

$$X = \frac{1}{\phi}(B)\big(\phi(B)X\big) = \frac{1}{\phi}(B)\big(\theta(B)Z\big) = \frac{\theta}{\phi}(B)Z.$$

using the shift calculus of Proposition 2.

That is, each ARMA solution equals to $\frac{\theta}{\phi}(B)Z$ and so uniqueness follows.

Causality and invertibility

Usually, the noise Z appears as an anonymous source of randomness, merely used to define desirable stationary process X by applying a filter to the noise Z giving

$$X = \frac{\theta}{\phi}(B)Z.$$

In many applications, the relation between noise Z and the process X is not of interest, since only μ_X and γ_X are relevant.

However, many arguments and calculations become easier if Z and X are inter-related.

Useful inter-relations are **causality** and **invertibility**.

Causality and invertibility

Consider again the ARMA equation

$$\phi(B)X = \theta(B)Z$$

with given parameters $(\phi_j)_{j=1}^{p}$ and $(\theta_j)_{j=1}^{q}$ and white noise $Z = (Z_t)_{t\in\mathbb{Z}}$.

We know by Theorem 1 that if ϕ has no roots on the unit circle, then the unique solution for X is

$$X = \frac{\theta}{\phi}(B)Z.$$

We can also show using similar arguments that if θ has no roots on the unit circle, then the unique solution for Z is

$$Z = \frac{\phi}{\theta}(B)X.$$

Causality and invertibility

CAUSALITY.

If ϕ has no roots on the unit disc, i.e.

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \quad \text{for all} \quad z \in \mathbb{C} \text{ and } |z| \leq 1$$

then the Laurent series is just the Taylor series, i.e.

$$\frac{\theta}{\phi}(z) = \sum_{j \in \mathbb{Z}} \psi_j z^j \quad \text{becomes} \quad \frac{\theta}{\phi}(z) = \sum_{j=0}^{\infty} \psi_j z^j,$$

and

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j (B^j Z)_t \quad \text{becomes} \quad X_t = \sum_{j=0}^{\infty} \psi_j (B^j Z)_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

In this case, X is called **causal** with respect to Z (only past and present values of the noise enter current process observation).

Causality and invertibility

INVERTIBILITY.

If θ has no roots on the unit disc, i.e.

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0 \quad \text{for all} \quad z \in \mathbb{C} \text{ and } |z| \leq 1$$

then the Laurent series is just the Taylor series, i.e.

$$\frac{\phi}{\theta}(z) = \sum_{j \in \mathbb{Z}} \pi_j z^j \quad \text{becomes} \quad \frac{\phi}{\theta}(z) = \sum_{j=0}^{\infty} \pi_j z^j,$$

so $X = \frac{\theta}{\phi}(B)Z \implies Z = \frac{\phi}{\theta}(B)X$ and

$$Z_t = \sum_{j=-\infty}^{\infty} \pi_j (B^j X)_t \quad \text{becomes} \quad \sum_{j=0}^{\infty} \pi_j (B^j X)_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

In this case, X is called **invertible** with respect to Z (only past and present process values are needed to reconstruct the present noise observation).

Causality and invertibility

WE NEED ONLY CAUSAL PROCESSES.

An interesting insight is that causal ARMA solutions give a sufficiently reach class of ARMA processes.

More precisely, it is possible to prove that if X is an ARMA process then there exists an invertible ARMA process X' (with other parameters) but with the same auto-covariance function $\gamma_X = \gamma_{X'}$.

Thus, if merely auto-covariance is of interest, then it suffices to consider causal ARMA processes only.

For this reason, it makes sense to consider only those ARMA equations where all roots of polynomial ϕ are outside of the unit disc.

Of course, ARMA(p, q) processes are stationary if all roots of polynomial ϕ are outside of the unit disc.

Causality and invertibility

CALCULATION OF AUTOCOVARIANCE γ_X.

Here we consider causal ARMA(p,q) processes, i.e. those with

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \quad \text{for all} \quad z \in \mathbb{C} \text{ and } |z| \leq 1.$$

Under this assumption, the ARMA solution $X = \frac{\theta}{\phi}(B)Z$ is obtained by applying the linear filter $(\psi_j)_{j=0}^{\infty}$ from the Taylor expansion

$$\sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta}{\phi}(z)$$

to the noise Z giving

$$X_t = \sum_{j=0}^{\infty} \psi_j (B^j Z)_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z}.$$

Causality and invertibility

Given filter $(\psi_j)_{j=0}^{\infty}$, we use (5) to calculate the auto-covariance as

$$\gamma_X(h) = \sum_{j=0}^{\infty} \psi_j \psi_{j+h} \sigma^2, \quad h \in \mathbb{Z}.$$

Thus, the main task is to determine the filter coefficients $(\psi_j)_{j\in\mathbb{Z}}$ from

$$\sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta}{\phi}(z).$$

Causality and invertibility

By comparison of powers

$$\underbrace{\left(\psi_0 + \psi_1 z + \psi_2 z^2 + \dots\right)}_{\frac{\theta}{\phi}(z)} \underbrace{\left(1 - \phi_1 z - \dots - \phi_p z^p\right)}_{\phi(z)} = \underbrace{\left(1 + \theta_1 z + \dots + \theta_q z^q\right)}_{\theta(z)}$$

we obtain

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 - \phi_1 \psi_0 &= \theta_1 \\ \psi_2 - \phi_1 \psi_1 - \phi_2 \psi_0 &= \theta_2 \\ &\vdots \end{aligned}$$

etc.

So for $j = 1, 2, \dots,$ the filter coefficients can be obtained recursively using

$$\psi_j = \sum_{k=1}^{j} \phi_k \psi_{j-k} + \theta_j \tag{8}$$

with $\psi_0 = 1$ and $\psi_i = 0$ when $i < 0$.

Causality and invertibility

Example.

Consider an ARMA(1,1) process given by

$$X_t - \underbrace{0.5}_{\phi_1} X_{t-1} = Z_t + \underbrace{0.4}_{\theta_1} Z_{t-1}, \quad t \in \mathbb{Z},$$

with Gaussian white noise, i.e. $Z_t \sim N(0, \sigma^2)$ for all $t \in \mathbb{Z}$.

The polynomial

$$\phi(z) = 1 - 0.5z$$

has a root of 2, which is outside the unit disc $|z| \leq 1$, and so the ARMA(1,1) equation has a unique, stationary solution X that is causal with respect to Z.

On the other hand, the polynomial

$$\theta(z) = 1 + 0.4z$$

has a root of -2.25, which is outside the unit disc $|z| \leq 1$, and so the ARMA(1,1) process is also invertible.

Causality and invertibility

The filter coefficients

$$\sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta}{\phi}(z) = \frac{1 + 0.4z}{1 - 0.5z}$$

can be calculated using (8) as

$$\begin{aligned}
\psi_0 &= 1 \\
\psi_1 &= \phi_1 \psi_0 + \theta_1 = 0.5 * 1 + 0.4 = 0.9 \\
\psi_2 &= \phi_1 \psi_1 + \underbrace{\phi_2}_{0} \psi_0 + \underbrace{\theta_2}_{0} = 0.5 * 0.9 = 0.45 \\
\vdots &\qquad \vdots \\
\psi_j &= \underbrace{\phi_1 \psi_{j-1}}_{0.5^{j-1} * 0.9} + \underbrace{\phi_2}_{0} \psi_{j-2} \underbrace{+ \cdots + \theta_j}_{0} \quad \text{for } j > 2
\end{aligned}$$

from which the auto covariance γ_X can be obtained.

Tesla example – modelling with ARMA(p, q)

Daily close.
In this example we consider the daily close price of Tesla from 2021 onwards.

A plot of this data is shown below.

Tesla example – modelling with ARMA(p, q)

We can use Mathematica to fit an ARMA(p, q) model and Mathematica will determine p, q and estimate the constant c and lag coefficients.

The "best fitting" model as selected by Mathematica is the ARMA(2,1)

$$X_t = 5.09583 + 0.433234 X_{t-1} + 0.545994 X_{t-2} + Z_t + 0.573865 Z_{t-1},$$

with the variance of the noise process $\text{var}(Z_t) \approx 85.4952$.

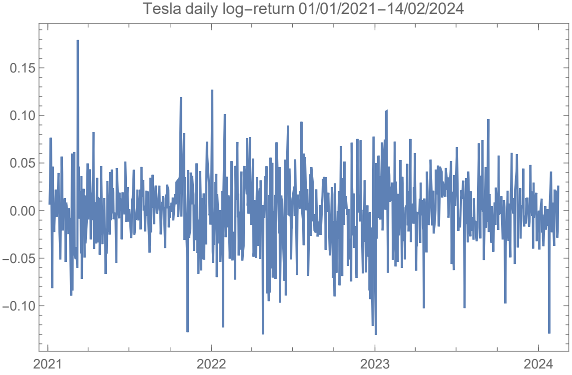
Warning from empirical stock data.
ARMA models should only be fitted to stationary time series, but there is strong evidence that stock price data is not stationary.

Here we should check for potential unit roots (use Dickey-Fuller test or the like).

Tesla example – modelling with ARMA(p, q)

Daily log-return.
Now consider the daily log-returns of Tesla for 2021 onwards.

A plot of this return data is below.



Tesla example – modelling with ARMA(p, q)

The best fitting ARMA(p, q) model is the white noise ARMA(0,0)

$$X_t = -0.000324269 + Z_t$$

with the variance of the noise process $\mathrm{var}(Z_t) \approx 0.00135682$.

Warning from empirical stock return data.
There is strong evidence that stock log-return data contains negligible covariance structure, suggesting our model is somewhat appropriate.

However, there is also strong evidence that (conditional) variance is not constant (perhaps use a GARCH process).

Non-stationarity and differencing

In many situations the data we seek to model will not be stationary (e.g. stock price data).

In this situation that are some techniques that can be applied to transform the data so that the transformed data is stationary (and so can then be modelled with ARMA(p, q) processes as discussed before).

We illustrate these ideas in the context of time series data with a deterministic trend component (which can be extended to cover seasonality and more complex models as well).

Non-stationarity and differencing

DIFFERENCING TO OBTAIN STATIONARITY.

Consider the model for $X = (X_t)_{t\in\mathbb{Z}}$ given by

$$X_t = m_t + Z_t, \quad t \in \mathbb{Z},$$

with $m = (m_t)_{t\in\mathbb{Z}}$ deterministic and $Z = (Z_t)_{t\in\mathbb{Z}}$ white noise, so that X is only stationary if m_t=0.

Non-stationarity and differencing

Linear trend.
If

$$m_t = c_0 + c_1 t$$

then

$$(1 - B)m_t = m_t - m_{t-1} = c_0 + c_1 t - \big(c_0 + c_1(t - 1)\big) = c_1$$

and we see that the linear trend has been eliminated so the process

$$(1 - B)X_t = c_1 + (1 - B)Z_t$$

is stationary (differencing works like differentiation).

The original process X is called $I(1)$ or **integrated of order 1**, i.e. X required first-order differencing to obtain a stationary process.

Non-stationarity and differencing

Quadratic trend.
If

$$m_t = c_0 + c_1 t + c_2 t^2$$

then

$$\begin{aligned}(1-B)^2 m_t &= (1-2B+B^2)m_t = m_t - 2m_{t-1} + m_{t-2} \\ &= c_0 + c_1 t + c_2 t^2 - 2\big(c_0 + c_1(t-1) + c_2(t-1)^2\big) \\ &\quad + \big(c_0 + c_1(t-2) + c_2(t-2)^2\big) \\ &= 2c_2\end{aligned}$$

and we see that the quadratic trend has been eliminated so the process

$$(1-B)^2 X_t = 2c_2 + (1-B)^2 Z_t$$

is stationary (differencing works like differentiation).

The original process X is called $I(2)$ or **integrated of order 2**, i.e. X required second-order differencing to obtain a stationary process.

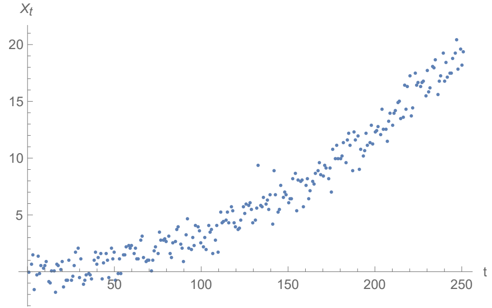
This idea can be extended to higher order polynomial drift.

Non-stationarity and differencing

Example.
Consider the $I(2)$ model with polynomial drift given by

$$X_t = 0.002t + 0.0003t^2 + Z_t$$

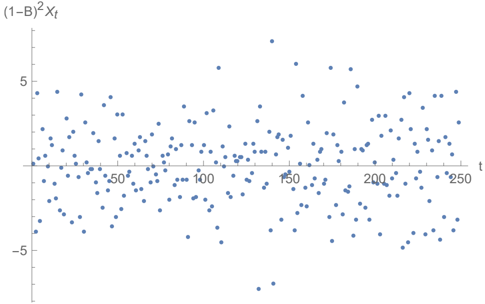
with $Z_t \sim N(0,1)$ for $t \in \{0, 1, \ldots, 250\}$ and $X_0 = 0$.



Non-stationarity and differencing

After differencing twice we have the stationary model

$$(1 - B)^2 X_t = 0.0006 + (1 - B)^2 Z_t.$$



Non-stationarity and differencing

In some situations, the differencing procedure can used to obtain an ARMA(p, q) process.

The original, integrated process is called an ARIMA(p, d, q), where d refers to the order of differencing required to obtain stationarrity.

Definition 4 (ARIMA(p, d, q) process)

An ARIMA(p, d, q) process $X = (X_t)_{t\in\mathbb{Z}}$ is described as

$$\Big(1 - \sum_{j=1}^{p} \phi_j B^j\Big)(1 - B)^d X_t = c + \Big(1 + \sum_{j=1}^{q} \theta_j B^j\Big) Z_t, \quad t \in \mathbb{Z},$$

where $(1 - B)^d X =$ is an ARMA(p, q) process, d refers to the order of differencing, $c, \phi_j, \theta_j \in \mathbb{R}$ and $Z = (Z_t)_{t\in\mathbb{Z}}$ is a zero-mean white noise process with variance σ^2.

Non-stationarity and differencing

Notice that the ARIMA(p, d, q) process has an ARMA$(p + d, q)$ style equation, but is not stationary as it has d unit roots.

The ARIMA(p, d, q) process X is $I(d)$, i.e. integrated order d, while the differenced process

$$Y = (1 - B)^d X$$

is an ARMA(p, q).

That is, the process $Y = (Y_t)_{t\in\mathbb{Z}}$ is the solution of

$$\Big(1 - \sum_{j=1}^{p} \phi_j B^j\Big) Y_t = c + \Big(1 + \sum_{j=1}^{q} \theta_j B^j\Big) Z_t, \quad t \in \mathbb{Z},$$

where the polynomial

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

satisfies

$$\phi(z) \neq 0 \quad \text{for all} \quad z \in \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$$

Tesla example – modelling with ARIMA(p, d, q)

Daily close.
Recall the close price of Tesla for 2021 onwards for which we fitted an ARMA(p, q) model.

Tesla example – modelling with ARIMA(p, d, q)

Now use Mathematica to fit an ARIMA(p, d, q) model.

The ARIMA(0,1,0) model becomes just an ARMA(0,0) model on the differenced data (order 1), i.e. the white noise

$$X_t' = (1 - B)X_t = -0.0696628 + Z_t$$

with the variance of the noise process $\text{var}(Z_t) \approx 85.9329$.

References I