Bertinoro 2012

Time Series

Marco Lippi and Umberto Triacca

Program

1. Stationarity
    - 1.2 The autocovariance function
    - 1.3 White noise processes
    - 1.4 Finite moving averages and autoregressive processes
    - 1.5 Infinite moving averages
    - 1.6 ARMA processes
    - 1.7 Long-memory processes: an example
2. Prediction
    - 2.1 The best predictor of a stationary process
    - 2.2 The best linear predictor
    - 2.3 The Wold representation theorem
    - 2.4 The Wold representation for ARMA processes

Program (continued)

3. Estimation of stationary processes
    3.1 Estimating the mean and the covariance function
    3.2 Maximum likelihood estimation of the parameters of ARMA processes
4. Multivariate stationary processes
    4.1 Generalizing stationarity, white noise, ARMA, to higher dimension
    4.2 Wold representation
    4.3 Inverting a matrix $A(L)$
    4.4 Cointegration
5. Non-stationary processes. Cycle and trend
    5.1 Trend stationary and difference stationary processes
    5.2 Explaining the fluctuations of the GDP: the debate on the relative importance of trend and cycle

Definition of a stochastic process

A stochastic process is (1) a correspondence associating a stochastic variable with each integer t:

$$x:\ t \mapsto x_t$$

(2) for any integer s and $t_1 \leq t_2 \leq \cdots \leq t_s$, the probability measure

$$\mathrm{P}(a_1 \leq x_{t_1} \leq b_1,\ a_2 \leq x_{t_2} \leq b_2, \ldots,\ a_s \leq x_{t_s} \leq b_s)$$

Saying that a stochastic process is given means that not only the stochastic variables x_t are given, but also the joint probability measure, for all s-tuples

$$x_{t_1},\ x_{t_2},\ \ldots,\ x_{t_s}$$

Notation. x_t for the stochastic variable associated with t. $\{x_t,\ t \in \mathbb{Z}\}$ for the process, but we also use , for short, x_t, meaning the process, instead of $\{x_t,\ t \in \mathbb{Z}\}$.

Definition of a stationary stochastic process

So we have a probability space $\mathcal{S}_t = (\Omega_t, \mathcal{F}_t, \mathrm{P})$ for each t, and the products

$$\Omega_{t_1} \times \Omega_{t_2} \times \cdots \times \Omega_{t_s}$$

with their probability measures

$$\mathrm{P}(a_1 \leq x_{t_1} \leq b_1,\ a_2 \leq x_{t_2} \leq b_2, \ldots,\ a_s \leq x_{t_s} \leq b_s)$$

Equivalently, with each s and $t_1 \leq t_2 \leq \cdots \leq t_s$ we can associate the probability distribution function

$$F^x_{t_1,t_2,\ldots,t_s}(b_1, b_2, \ldots, b_s) = \mathrm{P}(x_{t_1} \leq b_1,\ x_{t_2} \leq b_2, \ldots,\ x_{t_s} \leq b_s)$$

Definition of a stationary stochastic process

Definition. A stochastic process is strongly stationary if

$$\mathrm{P}(a_1 \leq x_{t_1} \leq b_1,\ a_2 \leq x_{t_2} \leq b_2, \ldots,\ a_s \leq x_{t_s} \leq b_s)$$
$$= \mathrm{P}\left(a_1 \leq x_{t_1+k} \leq b_1,\ a_2 \leq x_{t_2+k} \leq b_2, \ldots,\ a_s \leq x\right.$$

for all s, $t_1 \leq t_2 \leq \cdots \leq t_s$, and all integer k. We also say strictly stationary.

Equivalently, using the probability distributions,

$$F^x_{t_1,t_2,\ldots,t_s} = F^x_{\tau_1,\tau_2,\ldots,\tau_s}$$

where $\tau_j = t_j + k$.

For example,
the probability that x_1 lies between 1 and 2, AND x_2 lies between -3 and 3, and
the probability that x_{11} lies between 1 and 2, AND x_{12} lies between -3 and 3,
are the same.

Stochastic processes: examples

1. Let $\mathcal{S}_t = \mathcal{S} = (\Omega, \mathcal{F}, \mathrm{P})$. Let

$$x_t = A \text{ for all } t.$$

This is an extreme example in which the function $x : \ t \mapsto x_t$ is constant. We can say that x_t is a constant stochastic process. This process is of course stationary

2. A slightly modified example is

$$x_t = (-1)^t A.$$

3. At the other extreme, let the variables x_t be IID. Stationary. In this case $\mathrm{P}(x_{t_1} \in I_1,\ x_{t_2} \in I_2) = \mathrm{P}(x_{t_1} \in I_1)\mathrm{P}(x_{t_2} \in I_2)$.

4. Non stationary processes are

$$y_t = a + bt + x_t, \qquad z_t = tx_t,$$

where x_t is IID.

Stochastic processes: examples

5. Another important example of non-stationarity

$$w_t = \begin{cases} x_t & \text{if } t < 0 \\ 1 + x_t & \text{if } t \geq 0 \end{cases}$$

where x_t is ID. The mean changes but there is no trend.

Stochastic processes: realizations

Consider the space

$$\mathbb{R}^{\mathbb{Z}} = \cdots \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots$$

containing all two-sided infinite sequences of real numbers

$$y = (\cdots \; y_{-1} \, y_0 \, y_1 \; \cdots)$$

The probability

$$\mathrm{P}(a_1 \leq y_{t_1} \leq b_1,\; a_2 \leq y_{t_2} \leq b_2, \ldots,\; a_s \leq y_{t_s} \leq b_s)$$

can be interpreted as the probability of the subset B of $\mathbb{R}^{\mathbb{Z}}$ containing the sequences y that pass through $[a_1 \; b_1]$ at t_1, through $[a_2 \; b_2]$ at t_2, etc. In other words, the sequences of B can go wherever they want, provided that they behave in the prescripted way at $t_1,\; t_2,\; \ldots,\; t_s$.

Stochastic processes: realizations

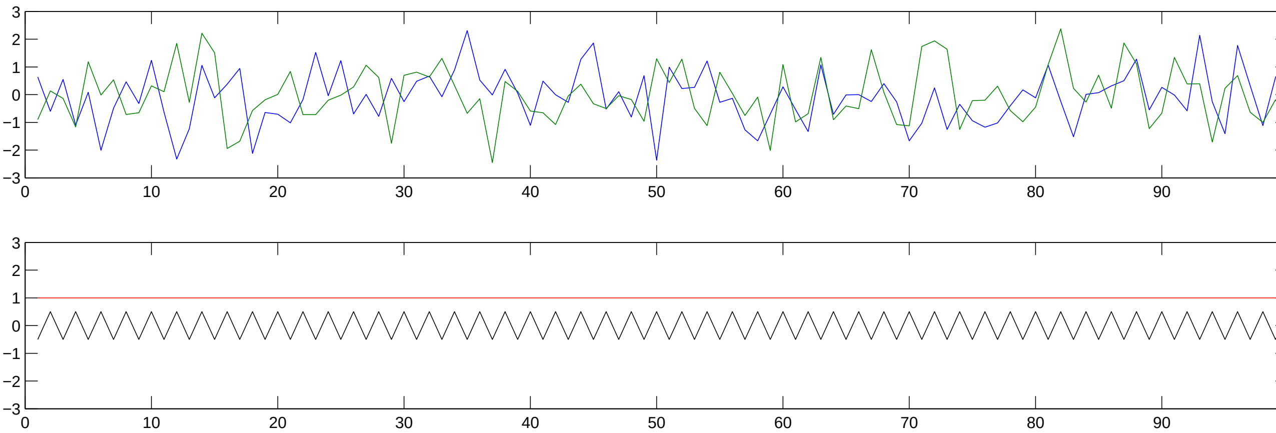
The space $\mathbb{R}^{\mathbb{Z}}$, endowed with the probability measure just defined, is called the space of realizations, or space of trajectories, of the stochastic process x_t.

For example,

1. If $x_t = A$, the constant process, then the realization is a constant sequence with probability one.
2. If $x_t = (-1)^t A$, then with probability one the realization is

$$(\cdots \quad -a \quad a \quad -a \quad a \quad \cdots)$$

Stochastic processes: realizations



The red line in the lower panel is a realization of $x_t = A$, the black line of $x_t = (-1)^t A$. In the upper panel we have two realizations of x_t, which is IID, normal, with zero mean and unit variance.

Redefining stochastic processes

Starting with the stochastic process $\{x_t,\ t \in \mathbb{Z}\}$ we have defined a probability space on $\mathbb{R}^{\mathbb{Z}}$. There is a correspondence between events in the spaces Ω_t and events in $\mathbb{R}^{\mathbb{Z}}$. For example, to the event $\{\omega \in \Omega_4,$ such that $x_4(\omega) = 5\}$ there corresponds in $\mathbb{R}^{\mathbb{Z}}$ the event

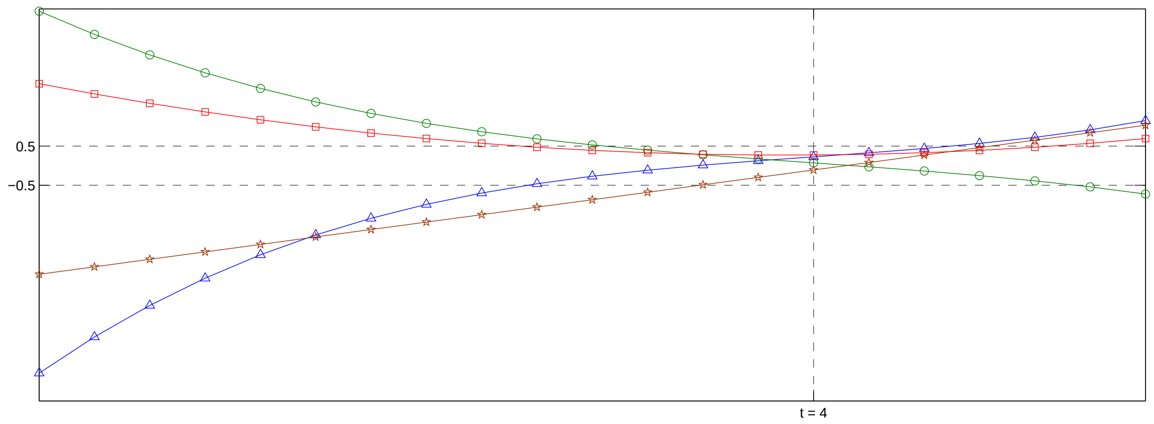
All trajectories $\{y_t,\ t \in \mathbb{Z}\}$ such that $y_4 = 5$

Moreover, to the stochastic variable $x_\tau : \ \Omega_\tau \longrightarrow \mathbb{R}$, there corresponds $X_\tau : \mathbb{R}^{\mathbb{Z}} \longrightarrow \mathbb{R}$ which is defined as follows:

$$X_\tau(\{y_t,\ t \in \mathbb{Z}\}) = y_\tau = x_\tau(\text{any } \omega \in \Omega_\tau \text{ such that } x_\tau(\omega) = y_\tau).$$

That is: If the trajectory $\{y_t,\ t \in \mathbb{Z}\}$ passes through z at $t = \tau$, then the stochastic variable X_τ takes the value z, which is precisely the value that x_τ would take for any ω of the subset of Ω_τ corresponding to $\{y_t,\ t \in \mathbb{Z}\}$.

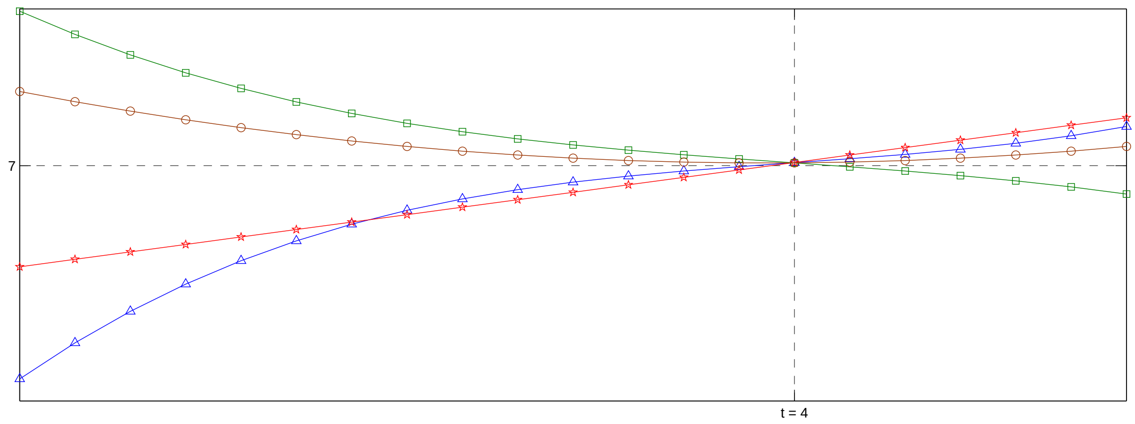
Redefining stochastic processes



The four trajectories belong to the event:

All trajectories that at $t = 4$ take values between -0.5 and 0.5.

Redefining stochastic processes



The four trajectories belong to the event:

All trajectories that take value 7 at $t = 4$.

The variable X_4 takes value 7 for all these elementary events of $\mathbb{R}^{\mathbb{Z}}$.

Redefining stochastic processes

It is quite obvious that the process $\{X_t,\ t \in \mathbb{Z}\}$ has the same probability distributions as the process $\{x_t,\ t \in \mathbb{Z}\}$, i.e.

$$F^X_{t_1,t_2,\ldots,t_s} = F^x_{t_1,t_2,\ldots,t_s}$$

The advantage of $\{X_t,\ t \in \mathbb{Z}\}$ is that all the variables are defined on the same probability space $\mathbb{R}^{\mathbb{Z}}$.

This is the basis for the following definition of a stochastic process.
Let $\mathcal{S} = (\Omega, \mathcal{F}, \mathrm{P})$ be a probability space and $\mathcal{L}$ be the set of all real stochastic variables defined on $\mathcal{S}$. A function

$$x :\ \mathbb{Z} \longrightarrow \mathcal{L}$$

is called a stochastic process.

$$t \longmapsto x_t, \qquad x_t : \Omega \longrightarrow \mathbb{R}$$

The autocovariance function

Suppose that x_t is strongly stationary and has finite second moment. Let $\mu = \mathrm{E}(x_t)$. Then consider

$$\mathrm{E}[(x_t - \mu)(x_{t-k} - \mu)]. \qquad (*)$$

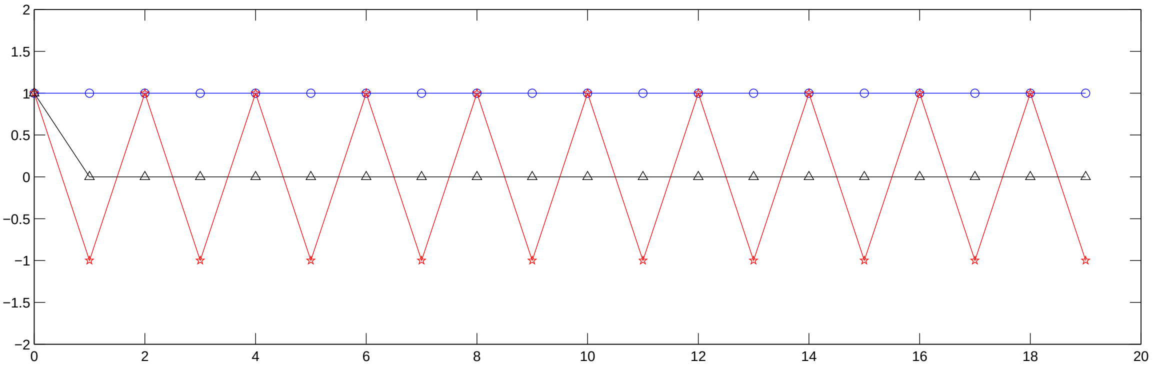
Since the joint distribution of x_t and x_{t-k} depends on k but not on t, then $(*)$ depends on k but not on t. We denote it by γ_k, or γ_k^x if necessary. The function associating γ_k with the integer k is called autocovariance function.

It easily seen that $\gamma_k = \gamma_{-k}$:

$$\gamma_k = \mathrm{E}[(x_t - \mu)(x_{t-k} - \mu)] = \mathrm{E}[(x_{t+k} - \mu)(x_t - \mu)] = \mathrm{E}[(x_t - \mu)(x_{t-(-k)} - \mu)] = \gamma_{-k}$$

Moreover $|\gamma_k| \leq \gamma_0 = \mathrm{var}(x_t)$.

The autocovariance function



In the figure we have the autocovariance function (first 20 values) of the constant process, blue circles, the process $x_t = (-1)^t A$, red asterisks, the IID process, black triangles. All processes have unit variance.

Weakly stationary processes

Suppose that x_t is a stationary process with finite second moments, that (i) $\mathrm{E}(x_t)$ is independent of t, call it μ, and

$$\mathrm{E}[(x_t - \mu)(x_{t-k} - \mu)] \qquad (*)$$

is independent of t. Then we say that x_t is weakly stationary. We denote $(*)$ by γ_k.

If x_t is strongly stationary and has finite second moments, then x_t is weakly stationary.

Of course a weakly stationary process is not necessarily strongly stationary. As an example, let $\gamma_k = 0$ for $k \neq 0$, $\gamma_0 = \mathrm{var}(x_t) = 1$. Moreover, assume that x_t is normal for $t \neq 0$ and uniform for $t = 0$.

If x_t is weakly stationary and normal then it is strongly stationary.

Unless explicitly stated, stationary processes will be weakly stationary.

White noise

We say that the stationary process x_t is a white noise if $\mathrm{E}(x_t) = 0$ and

$$\gamma_k = 0 \quad \text{for } k \neq 0$$

The IID process is of course a white noise.

The autocovariance function of a white noise has only one non-zero value.

Moving averages of a white noise

Start with a white noise process u_t. The process

$$x_t = \sum_{j=-m}^{m} a_j u_{t-j}$$

is called a moving average of u_t.

The process x_t is stationary. For, the mean of x_t is zero, obvious. Moreover, to compute the autocovariances,

$$\begin{aligned} x_t &= \cdots + a_1 u_{t-1} + \ a_0 u_t \ + a_{-1} u_{t+1} + \cdots \\ x_{t-1} &= \cdots + a_0 u_{t-1} + a_{-1} u_t + a_{-2} u_{t+1} + \cdots \end{aligned}$$

We have:

$$\gamma_k^x = \sigma_u^2 \textstyle\sum_{j=-m+k}^{m} a_j a_{j-k}, \qquad \text{in particular} \quad \gamma_0^x = \sigma_u^2 \sum_{j=-m}^{m} a_j^2$$

Moving averages of a white noise

For the moving average

$$x_t = \sum_{j=-m}^{m} a_j u_{t-j},$$

we have

$$\gamma_s^x = 0 \quad \text{for } s > 2m+1$$

Example: $x_t = u_t + 2u_{t-1} - 6u_{t-2}$:

$$\gamma_0^x = 41\sigma_u^2, \quad \gamma_1^x = -10\sigma_u^2, \quad \gamma_2^x = -6\sigma_u^2, \quad \gamma_s^x = 0 \text{ for } s > 2$$

Note that the autovariance function of the process $\{x_t,\ t \in \mathbb{Z}\}$ and that of $\{x_{t-h},\ t \in \mathbb{Z}\}$ are identical. For example $y_t = u_{t+1} + 2u_t - 6u_{t-1}$ and $x_t = y_{t-1}$.

Moving average of a general process

We can construct a moving average with any process x_t:

$$y_t = \sum_{j=-m}^{m} a_j x_{t-j}.$$

If x_t is stationary then y_t is stationary. For, $\mathrm{E}(y_t) = \mu \sum a_j$. Moreover:

$$\begin{aligned}
\mathrm{E}(y_t y'_{t-k}) &= \mathrm{E}\left[(a_m \; a_{m-1} \; \cdots \; a_{-m+1} \; a_{-m}) \begin{pmatrix} x_{t-m} \\ x_{t-m+1} \\ \vdots \\ x_{t+m-1} \\ x_{t+m} \end{pmatrix} (x_{t-m-k} \; x_{t-m+1-k} \; \cdots \; x \right. \\
&= a \begin{pmatrix} \gamma^x_k & \gamma^x_{k-1} & \cdots & \gamma^x_{k-2m} \\ \gamma^x_{k+1} & \gamma^x_k & \cdots & \gamma^x_{k-2m+1} \\ & & \ddots & \\ \gamma^x_{k+2m} & \gamma^x_{k+2m-1} & \cdots & \gamma^x_k \end{pmatrix} a'
\end{aligned}$$

The lag operator

If x_t is a stochastic process, define the lag operator, denoted by L, by

$$Lx_t = x_{t-1}.$$

Moreover, for $k > 1$,

$L^k x_t = L(L^{k-1} x_t),\ \ L^0 = 1$ where 1 here means the identity operator: $1x_t = x_t$.

Lastly $Fx_t = L^{-1}x_t = x_{t+1}$.

We also define functions of L,

$$\begin{aligned} a(L) &= a_{-m}L^{-m} + \cdots + a_{-1}L^{-1} + a_0 + a_1 L + \cdots + a_m L^m \\ &= a_{-m}F^m + \cdots + a_{-1}F + a_0 + a_1 L + \cdots + a_m L^m \end{aligned}$$

Moving averages can be rewritten as:

$$x_t = a(L)u_t = a_{-m}u_{t+m} + \cdots + a_{-1}u_{t+1} + a_0 u_t + a_1 u_{t-1} + \cdots + a_m u_{t-m}.$$

The lag operator

Examples:

$$x_t = u_t + 2u_{t-1} - 6u_{t-2} = (1 + 2L - 6L^2)u_t, \quad y_t = u_{t+1} + 2u_t - 6u_{t-1} = (F + 2 - 6L)$$

We have seen that the autocovariance function of x_t and $L^k x_t$ are equal for all k.

We have

$$a(L)[b(L)u_t] = [a(L)b(L)]u_t,$$

where $a(L)b(L)$ is obtained by the ordinary algebraic rules. Example:

$$(F + 2 - 6L)[(1 + 5L^2)u_t] = [F + 2 - L + 10L^2 - 30L^3]u_t.$$

Check this result.

Infinite moving averages of a white noise

We can in principle consider a moving average of a white noise with an infinite number of terms:

$$x_t = \sum_{j=-\infty}^{\infty} a_j u_{t-j} = \cdots + a_m u_{t-m} + \cdots + a_0 u_t + \cdots + a_{-m} u_{t+m} + \cdots$$

provided that

$$\sum_{j=-\infty}^{\infty} a_j^2 < \infty. \qquad (*)$$

Condition $(*)$ ensures that x_t has finite variance and covariances:

$$\gamma_0^x = \sigma_u^2 \sum_{j=-\infty}^{\infty} a_j^2, \qquad \gamma_k^x = \sigma_0^x \sum_{j=-\infty}^{\infty} a_j a_{j-k}$$

(Do you remember the Cauchy-Schwartz inequality? $\left|\sum c_j d_j\right| \leq \sqrt{\sum c_j^2}\sqrt{\sum d_j^2}$.)

Infinite moving averages of a white noise

To see why infinite moving averages are important consider the difference equation:

$$z_t = \alpha z_{t-1} + u_t,$$

where u_t is a white noise, $\alpha < 1$, and z_t is an unknown stochastic process.

Suppose that x_t is a solution. Then:

$$x_t = \alpha x_{t-1} + u_t = \alpha(\alpha x_{t-2} + u_{t-1}) + u_t = [u_t + \alpha u_{t-1}] + \alpha^2 x_{t-2} = [u_t + \alpha u_{t-1} + \alpha^2 u_t$$

Thus an obvious candidate solution is

$$x_t = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \cdots = \sum_{j=0}^{\infty} \alpha^j u_{t-j}$$

And it works

$$u_t + \alpha x_{t-1} = u_t + \alpha(u_{t-1} + \alpha u_{t-2} + \cdots) = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \cdots = x_t$$

Infinite moving averages of a white noise

Note that other solutions exist. If A is any stochastic variable, and $y_t = x_t + A\alpha^t$, then

$$u_t + \alpha y_{t-1} = u_t + \alpha(x_{t-1} + A\alpha^{t-1}) = [u_t + \alpha x_{t-1}] + A\alpha^t = x_t + A\alpha^t = y_t$$

and viceversa, if y_t is a solution then $y_t = x_t + B\alpha^t$ for some stochastic variable B.

Thus x_t is the only stationary solution.

In conclusion, infinite moving averages arise as solutions of elementary and sensible problems.

Infinite moving averages of a white noise

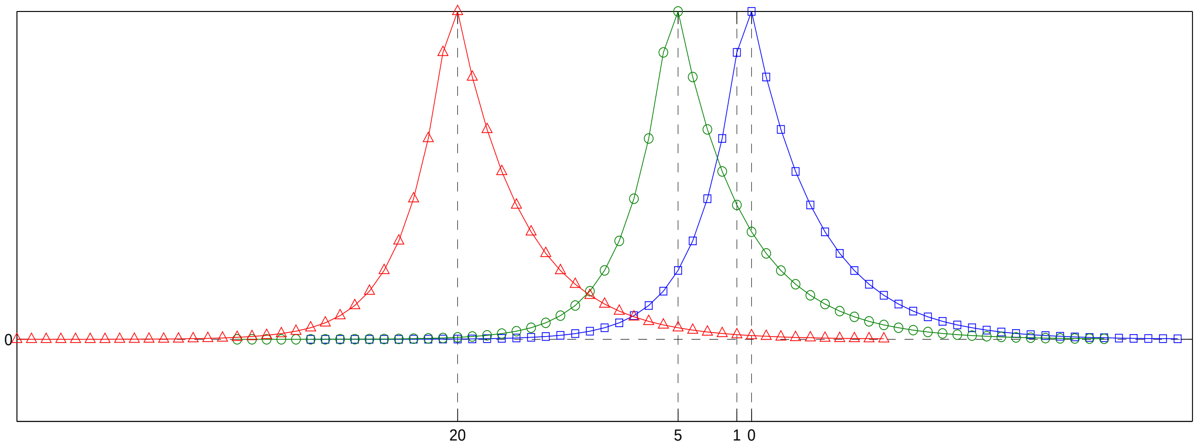
Infinite moving averages of a white noise are a large class of stochastic processes. We will see that every stationary stochastic process can be decomposed into an infinite moving average of a white noise plus a "deterministic" component.

A very important property of infinite moving averages of a white noise is that the autocovariance function vanishes when the lag tends to infinity. Precisely, if $x_t = \sum a_h u_{t-h}$,

$$\lim_{k\to\infty} \gamma_k^x = 0$$

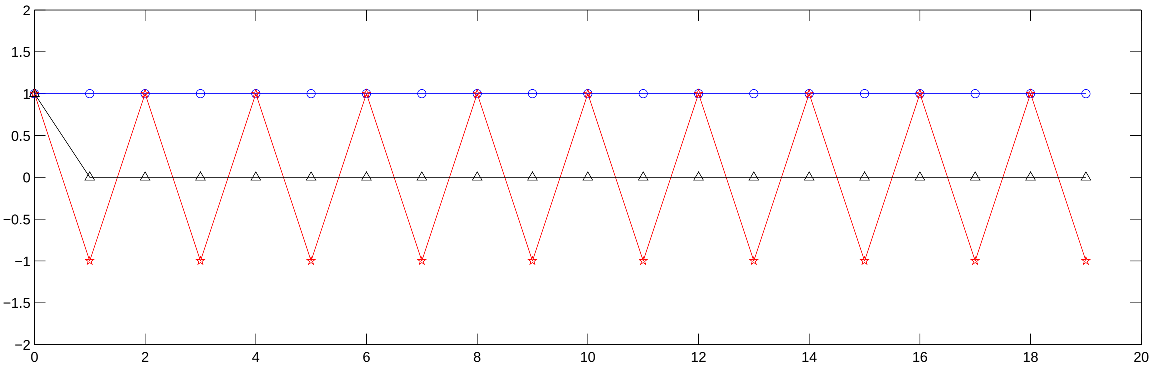
This can be easily grasped: $\gamma_k^x = \sum_{h=-\infty}^{\infty} a_h a_{h-k} \sigma_u^2$. The formula can be "seen" as obtained by displaying the array a_j on top of the same array shifted to the left by k integer steps, then taking the products "vertically" and summing. For example, for $k = 1$

$$\begin{matrix} \cdots & a_1 & a_0 & a_{-1} & \cdots \\ \cdots & a_0 & a_{-1} & a_{-2} & \cdots \end{matrix}$$



Infinite moving averages. The squares on the blue line are the coefficients a_j. The graph has a belly and two tails, a stylization of the fact that the coefficients must decline to zero as j tends to infinity (they must be square summable). The green and the red graphs are obtained by shifting the blue line back of 5 and 20 positions respectively. It is apparent that increasing the lag the belly of the blue graph corresponds more and more to the smaller and smaller values of the tail of the shifted graph, while the belly of the shifted graph corresponds to the tail of the blue graph.

Infinite moving averages of a white noise



Now we know that there exist processes that are not moving averages of a white noise. The autocovariance function of the processes $x_t = A$ and $x_t = (-1)^t A$ does not decline: see the blue and red lines above.

ARMA processes

We have discussed the stationary solution of

$$z_t = \alpha z_{t-1} + u_t.$$

Rewrite it as $z_t - \alpha z_{t-1} = u_t$ and then as

$$(1 - \alpha L) z_t = u_t$$

We call this equation an autoregressive equation of order one, and the stationary solution an autoregressive process of order one, or an AR(1). An obvious generalization is

$$z_t = \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} + u_t, \text{ that is } \quad (1 - \alpha_1 L - \cdots - \alpha_p L^p) z_t = u_t$$

an autoregressive equation of order p.

ARMA processes

We want to solve the equation

$$(1 - \alpha_1 L - \cdots - \alpha_p L^p) z_t = u_t.$$

This means looking for a moving average of u_t that solves the equation.

For $p = 1$ we found that if $|\alpha| < 1$ the solution is

$$x_t = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \cdots$$

We can define the inverse of $1 - \alpha L$:

$$\frac{1}{1 - \alpha L} = (1 - \alpha L)^{-1} = 1 + \alpha L + \alpha^2 L^2 + \cdots$$

and write

$$x_t = (1 - \alpha L)^{-1} u_t = (1 + \alpha L + \alpha^2 L^2 + \cdots) u_t$$

The expansion

$$\frac{1}{1-\alpha L} = 1 + \alpha L + \alpha L^2 + \cdots \qquad (*)$$

closely resembles the Taylor expansion

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots$$

Both must be handled with extreme care. If $|x| \geq 1$ the Taylor series does not converge. In the same way, if $|\alpha| \geq 1$ application of the operator $(*)$ to u_t does not produce a finite-variance stochastic variable.

However, for $|\alpha| > 1$ the operator $1 - \alpha L$ has an inverse with an interesting expansion:

$$\frac{1}{1-\alpha L} = \frac{F}{F-\alpha} = -\alpha^{-1} F \frac{1}{1-\alpha^{-1}F} = -\alpha^{-1}F(1 + \alpha^{-1}F + \alpha^{-2}F^2 + \cdots)$$

where $F = L^{-1}$.

However, for $|\alpha| > 1$ the operator $1 - \alpha L$ has an inverse with an interesting expansion:

$$\frac{1}{1-\alpha L} = \frac{F}{F-\alpha} = -\alpha^{-1}F\frac{1}{1-\alpha^{-1}F} = -\alpha^{-1}F(1+\alpha^{-1}F+\alpha^{-2}F^2+\cdots)$$

where $F = L^{-1}$.

This can be used to solve the equation

$$z_t = 2z_{t-1} + u_t, \quad \text{that is} \quad (1-2L)z_t = u_t$$

The solution is

$$x_t = -\frac{1}{2}(1 + \frac{1}{2}F + \frac{1}{2^2}L^2 + \cdots)u_t = -\frac{1}{2}u_{t+1} - \frac{1}{2^2}u_{t+2} + \cdots$$

thus a moving average of future values of u_t.

This can be used to solve the equation

$$z_t = 2z_{t-1} + u_t, \quad \text{that is} \quad (1 - 2L)z_t = u_t$$

The solution is

$$x_t = -\frac{1}{2}(1 + \frac{1}{2}F + \frac{1}{2^2}L^2 + \cdots)u_t = -\frac{1}{2}u_{t+1} - \frac{1}{2^2}u_{t+2} + \cdots$$

thus a moving average of future values of u_t.

The same result can be seen immediately if

$$z_t = 2z_{t-1} + u_t \text{ is rewritten as } z_{t-1} = \frac{1}{2}z_t - \tfrac{1}{2}u_t,$$

that is

$$z_t = \frac{1}{2}z_{t+1} + v_t,$$

where $v_t = -\frac{1}{2}u_{t+1}$.

ARMA processes

Thus the equation $(1 - \alpha L)z_t = u_t$ has a stationary solution which is a moving average of u_t, in the past if $|\alpha| < 1$,

$$x_t = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \cdots$$

in the future if $|\alpha| > 1$,

$$x_t = -\alpha^{-1} u_{t+1} - \alpha^{-2}_{t+2} - \cdots$$

Only when $|\alpha| = 1$ there is no stationary solution. More precisely, the equations

$$z_t = z_{t-1} + u_t, \qquad z_t = -z_{t-1} + u_t$$

have no stationary solutions unless $u_t = 0$. In that case the solutions are $x_t = A$ and $x_t = (-1)^t A$, respectively.

ARMA processes

Let us go back to the autoregressive equation of order p:

$$(1-\alpha_1 L-\cdots-\alpha_p L^p)z_t=u_t$$

We solve it by factoring the polynomial into first order factors. Precisely:

$$\begin{aligned}1-\alpha_1 L- \ & \alpha_2 L^2-\cdots-\alpha_p L^p=-\alpha_p(L-\gamma_1)(L-\gamma_2)\cdots(L-\gamma_p)\\ &=[-\alpha_p(-1)^p\gamma_1\gamma_2\cdots\gamma_p](1-\delta_1 L)(1-\delta_2 L)\cdots(1-\delta_p L)=(1-\delta_1\end{aligned}$$

where $\delta_j=1/\gamma_j$.

Thus the autoregressive equation can be rewritten as

$$(1-\delta_1 L)(1-\delta_2 L)\cdots(1-\delta_p L)z_t=u_t$$

and the solution is obtained by inverting the polynomials $1-\delta_j L$ one after another.

ARMA processes

Thus the autoregressive equation can be rewritten as

$$(1-\delta_1 L)(1-\delta_2 L)\cdots(1-\delta_p L)z_t = u_t$$

and the solution is obtained by inverting the polynomials $1-\delta_j L$ one after another.

More precisely, if the roots of the polynomial $1-\alpha_1 L-\cdots-\alpha_p L^p$, that is the γ's, are greater than 1 in modulus, that is if $|\delta_j|<1$, then for the stationary solution we have

$$\begin{aligned} x_t &= (1-\delta_1 L)^{-1}(1-\delta_2 L)^{-1}\cdots(1-\delta_p L)^{-1}u_t \\ &= (1+\delta_1 L+\delta_1^2 L^2+\cdots)\cdots(1+\delta_p L+\delta_p^2 L^2+\cdots)u_t \\ &= \left[1+A_1 L+A_2 L^2+\cdots\right]u_t \\ &= u_t+A_1 u_{t-1}+A_2 u_{t-2}+\cdots \end{aligned}$$

where

$$A_k = \sum_{k_1+k_2+\cdots+k_p=k} \delta_1^{k_1}\delta_2^{k_2}\cdots\delta_p^{k_p}$$

In particular, $A_1=\delta_1+\delta_2+\cdots+\delta_p$.

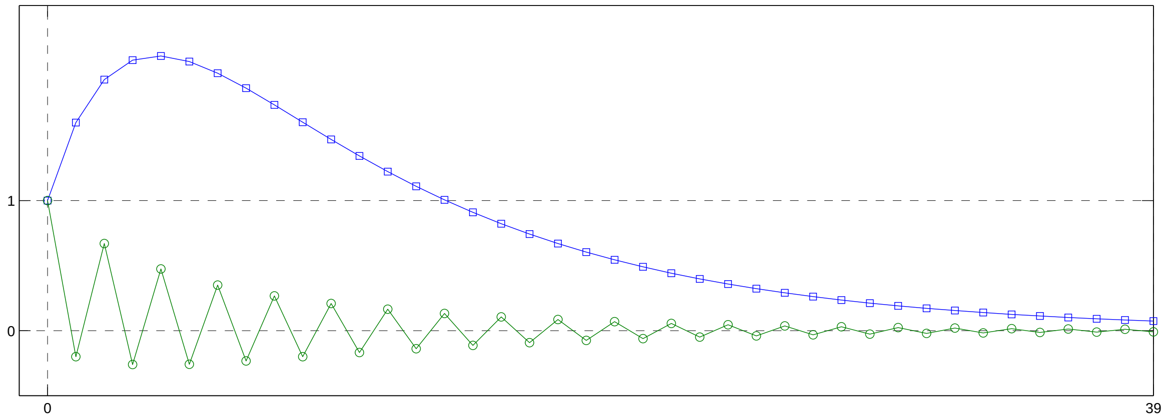
ARMA processes

About the roots of the autoregressive polynomial, let us insist that that the root of $1 - \alpha L$ is $1/\alpha$. Thus, by requiring that the root is greater than 1 in modulus we require that α is smaller than 1 in modulus.

ARMA processes

Example: $(1 - \alpha_1 L - \alpha_2 L^2) z_t = u_t$. The solution is

$$x_t = u_t + [\delta_1 + \delta_2] u_{t-1} + [\delta_1^2 + \delta_1 \delta_2 + \delta_2^2] u_{t-2} + [\delta_1^3 + \delta_1^2 \delta_2 + \delta_1 \delta_2^2 + \delta_2^3] u_{t-3} + \cdots$$

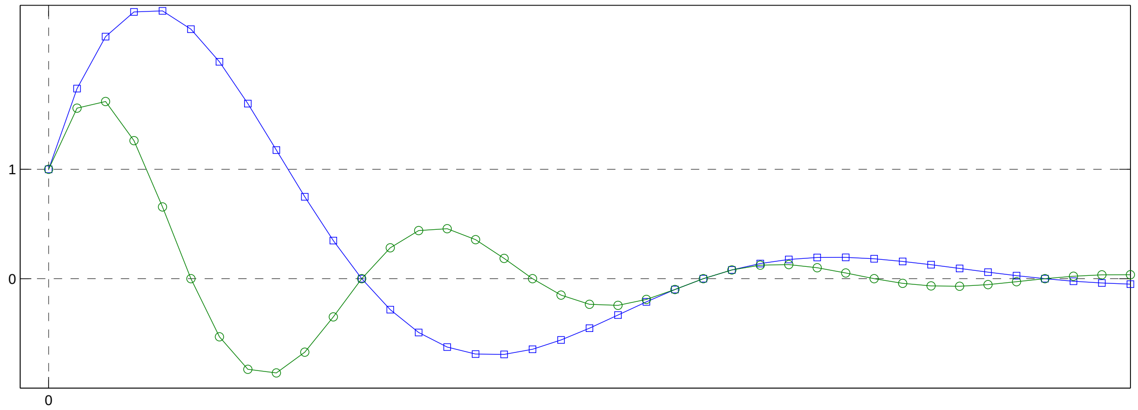


The figure shows the coefficients corresponding to $\delta_1 = 0.9$, $\delta_2 = 0.7$, blue line, $\delta_1 = -0.9$, $\delta_2 = 0.7$, green line.

ARMA processes

Example: $(1-\alpha_1 L - \alpha_2 L^2)z_t = u_t$. The solution is

$$x_t = u_t + [\delta_1 + \delta_2]u_{t-1} + [\delta_1^2 + \delta_1\delta_2 + \delta_2^2]u_{t-2} + [\delta_1^3 + \delta_1^2\delta_2 + \delta_1\delta_2^2 + \delta_2^3]u_{t-3} + \cdots$$



The blue line and green show the coefficients corresponding to the complex conjugate solutions $\delta_1 = 0.9(\cos\frac{2\pi}{\tau} + i\sin\frac{2\pi}{\tau})$, $\delta_2 = \bar{\delta}_1$, for $\tau = 24$ and 12 respectively. The coefficients are real, as expected (see also the formula above).

ARMA processes

Another technique to obtain the moving average corresponding to the autoregressive equation

$$(1 - \alpha_1 L - \alpha_2 L^2) z_t = u_t$$

is the following. Suppose that $x_t = (1 + a_1 L + a_2 L^2 + \cdots) u_t$ is the solution. Then

$$(1 - \alpha_1 L - \alpha_2 L^2)(1 + a_1 L + a_2 L^2 + \cdots) u_t = u_t$$

i.e.

$$\Big[1 + (a_1 - \alpha_1) L + (a_2 - \alpha_1 a_1 - \alpha_2) L^2 + \cdots + (a_k - \alpha_1 a_{k-1} - \alpha_2 a_{k-2}) L^k + \cdots$$

Since u_t is a white noise all the coefficients, apart the first, must be zero. Thus we find the difference equation

$a_k - \alpha_1 a_{k-1} - \alpha_2 a_{k-2} = 0,$ for $k \geq 2$, with initial conditions $a_1 = \alpha_1$ and $a_2 = \alpha_1 a_1$

Lastly, examining again the formula

$$\begin{aligned} x_t &= u_t + [\delta_1 + \delta_2]u_{t-1} + [\delta_1^2 + \delta_1\delta_2 + \delta_2^2]u_{t-2} + [\delta_1^3 + \delta_1^2\delta_2 + \delta_1\delta_2^2 + \delta_2^3]u_{t-3} + \cdots \\ &= u_t + a_1u_{t-1} + a_2u_{t-2} + \cdots \end{aligned}$$

we see that if, say, $|\delta_1| \geq |\delta_2|$, then the k-th coefficient of the moving average, in modulus, is smaller or equal to $k|\delta_1|^k$. Thus the coefficients a_k tend to zero faster that any geometric sequence ρ^k with $|\delta_1| < \rho < 1$. We say that the coefficients a_k decline geometrically with rate $|\delta_1|$, which is the modulus of the "worst" δ, that is the one that is closer to 1.

ARMA processes

In general, the coefficients of the moving average

$$x_t = u_t + A_1 u_{t-1} + A_2 u_{t-2} + \cdots$$

solution of the autoregressive equation

$$(1 - \alpha_1 L - \cdots - \alpha_p L^p) z_t = u_t$$

tend to zero geometrically at rate $|\delta_1|$, where, with no loss of generality, $|\delta_1| \geq |\delta_s|$, this meaning that δ_1 is the worst of the δ's.

ARMA processes

The stationary solution of

$$(1 - \alpha_1 L - \cdots - \alpha_p L^p) z_t = u_t,$$

under the assumption that all the roots of $1 - \alpha_1 L - \cdots - \alpha_p L^p$ are greater than 1 in modulus, is called an AR(p) process. An AR(p) process has a moving average representation in the contemporaneous and past values of u_t:

$$x_t = u_t + A_1 u_{t-1} + A_2 u_{t-2} + \cdots$$

If some of the roots γ_j of $1 - \alpha_1 L - \cdots - \alpha_p L^p$ are smaller than 1 then

$$1 - \alpha_1 L - \cdots - \alpha_p L^p = (1 - \gamma_1 L)(1 - \gamma_2 L) \cdots (1 - \gamma_p L)$$

can still be inverted but not in L alone, so that x_t in this case will be a two-sided moving average of u_t.

For example, suppose that $1-\alpha_1 L-\alpha_2 L^2 = (1-\delta_1 L)(1-\delta_2 L)$, with $|\delta_1| < 1$ but $|\delta_2| > 1$. Then

$$(1-\delta_1 L)^{-1}(1-\delta_2 L)^{-1} = -(1+\delta_1 L+\delta_1^2 L^2+\cdots)(\delta_2^{-1}F+\delta_2^{-2}F^{-2}+\cdots)$$

In this case

$$x_t = \cdots + A_1 u_{t-1} + A_0 u_t + A_{-1} u_{t+1} + \cdots$$

with

$$\begin{aligned} A_0 &= -\delta_1\delta_2^{-1} - \delta_1^2\delta_2^{-2} - \cdots \\ A_1 &= -\delta_1^2\delta_2^{-1} - \delta_1^3\delta_2^{-2} - \cdots \\ A_{-1} &= -\delta_2^{-1} - \delta_1\delta_2^{-2} - \cdots \\ &\vdots \end{aligned}$$

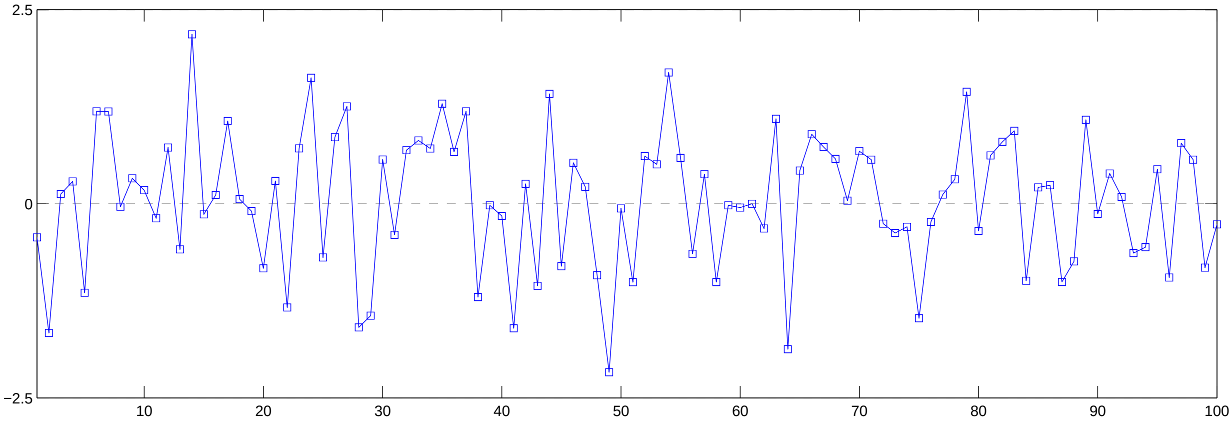
(Check that the coefficients A_k and A_{-k} tend to zero geometrically for $k \to +\infty$.)

But if some of the roots γ has unit modulus, in that case we say that the autoregressive polynomial has unit roots, then the autoregressive equation

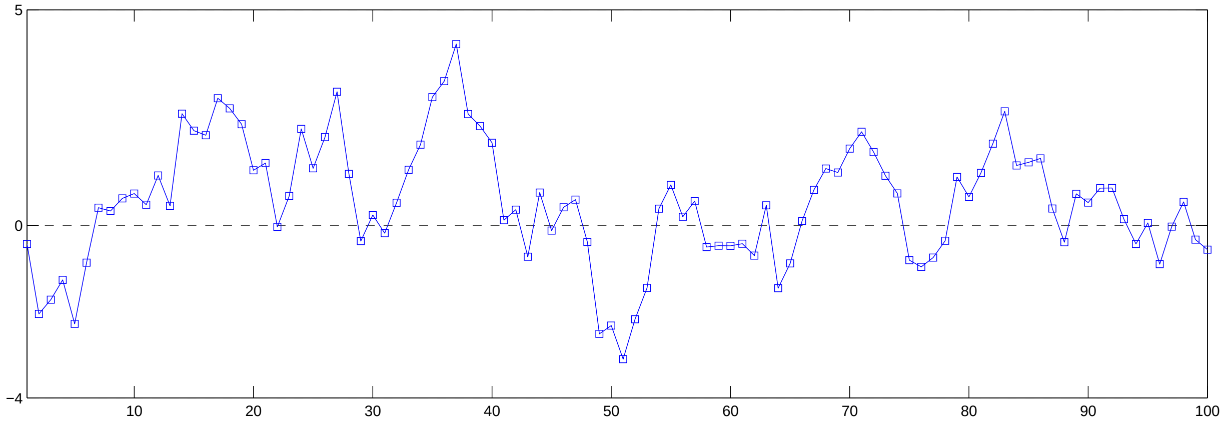
$$(1 - \alpha_1 L - \cdots - \alpha_p L^p) z_t = u_t,$$

has no stationary solution. (Unit roots will be discussed later on.)

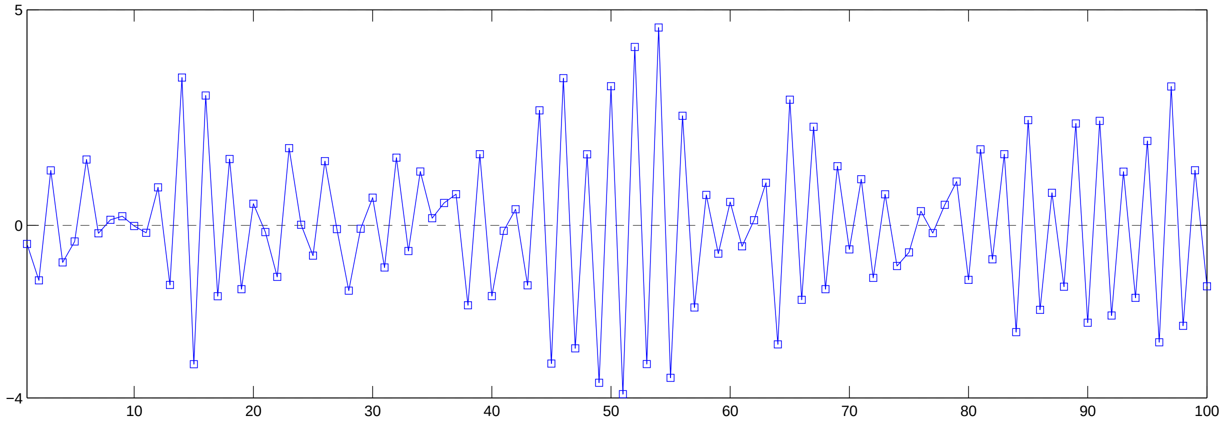
We assume that the autoregressive polynomial has no root inside or on the unit circle, i.e. that $|\gamma_j| > 1$ for all j or that $|\delta_j| < 1$ for all j. In this case we say that the stability, or stationarity condition is fulfilled. (But we know that roots inside the unit circle do not imply that a stationary solution does not exist, they only produce moving average containing future values of u_t. It is only with roots of unit modulus that stationary solutions do not exist.)



ARMA processes. Plot of u_t, normally distributed white noise.



ARMA processes. Plot of $x_t = (1 - 0.9L)^{-1} u_t$.



ARMA processes. Plot of $x_t = (1 + 0.9L)^{-1} u_t$. The equation is

$$z_t = -0.9 z_{t-1} + u_t.$$

ARMA processes

The autocovariance function of an autoregressive process can be obtained very easily. Assume that $p = 2$:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + u_t \qquad (*)$$

We know that

$$x_t = u_t + A_1 u_{t-1} + A_2 u_{t-2} + \cdots \qquad (**)$$

Multiplying both sides of $(*)$ by $x_{t-k} = u_{t-k} + A_1 u_{t-k-1} + \cdots$, for $k \geq 1$, we obtain

$$\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2}$$

(Yule-Walker equations) same as $(*)$ apart from u_t.

The first two equations, for $k = 1$ and $k = 2$, are

$$\begin{aligned}\gamma_1 &= \alpha_1\gamma_0 + \alpha_2\gamma_{-1} = \alpha_1\gamma_0 + \alpha_2\gamma_1 \\ \gamma_2 &= \alpha_1\gamma_1 + \alpha_2\gamma_0\end{aligned}$$

(recall that $\gamma_{-1} = \gamma_1$). Multiplying by x_t, and using $(**)$, we get $\gamma_0 = \alpha_1\gamma_1 + \alpha_2\gamma_2 + \sigma_u^2$. Thus

$$\begin{aligned}\gamma_0 &= \alpha_1\gamma_1 + \alpha_2\gamma_2 + \sigma_u^2 \\ \gamma_1 &= \alpha_1\gamma_0 + \alpha_2\gamma_1 \\ \gamma_2 &= \alpha_1\gamma_1 + \alpha_2\gamma_0\end{aligned}$$

This system determines γ_0, γ_1 and γ_2, then $\gamma_3 = \alpha_1\gamma_2 + \alpha_2\gamma_1$, etc.

ARMA processes

For $p = 1$, $z_t = \alpha z_{t-1} + u_t$ the equation is $\gamma_k = \alpha\gamma_{k-1}$ for $k \geq 0$. Using the same trick as before we get

$$\begin{aligned} \gamma_0 &= \alpha\gamma_1 + \sigma_u^2 \\ \gamma_1 &= \alpha\gamma_0 \\ \gamma_k &= \alpha^k\gamma_0 \end{aligned}$$

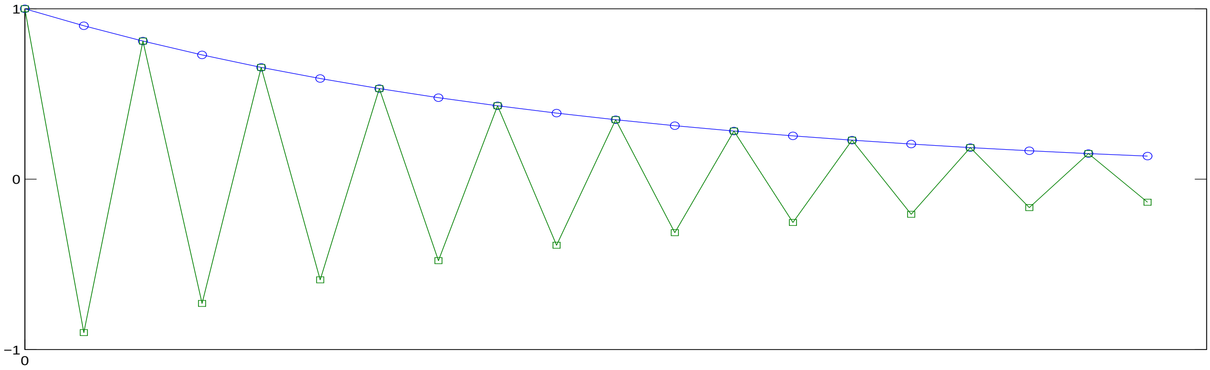
This gives

$$\gamma_0 = \frac{\sigma_u^2}{1-\alpha^2}, \quad \gamma_k = \alpha^k \frac{\sigma_u^2}{1-\alpha^2}$$

ARMA processes

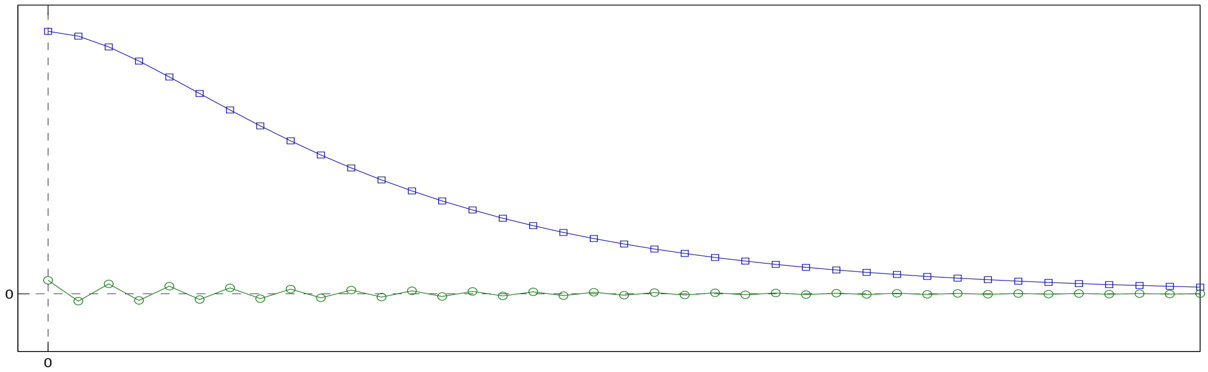
For $p = 1$:

$$\gamma_0 = \frac{\sigma_u^2}{1-\alpha^2}, \qquad \gamma_k = \alpha^k \frac{\sigma_u^2}{1-\alpha^2}$$



Autocovariance function for AR(1) processes: $\alpha = 0.9$, blue line and $\alpha = -0.9$, green line.

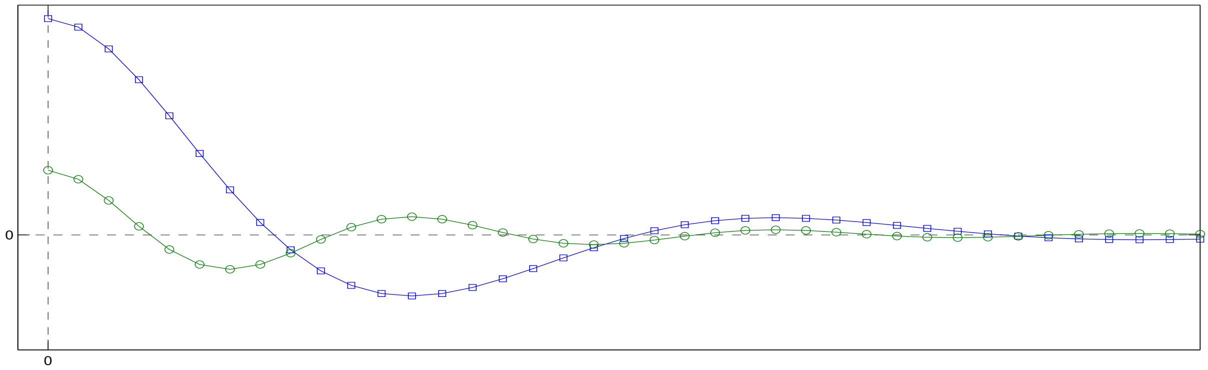
ARMA processes



Autocovariance function for AR(2) processes, with roots 0.9 and 0.7, blue line, -0.9 and 0.7, green line:

$$x_t = 1.6x_{t-1} - .63x_{t-2} + u_t, \qquad x_t = -0.2x_{t-1} + 0.63x_{t-2} + u_t,$$

respectively. (Compare with the graph of the moving-average coefficients.)



Autocovariance function for AR(2) processes, with roots $0.9\left(\cos\frac{2\pi}{\tau}+i\sin\frac{2\pi}{\tau}\right)$ and its conjugate, for $\tau=24$, blue line, $\tau=12$, green line:

$$x_t = 1.74x_{t-1} - .81x_{t-2} + u_t, \qquad x_t = 1.55x_{t-1} - .81x_{t-2} + u_t,$$

respectively. (Compare with the graph of the moving-average coefficients.)

ARMA processes

The stationary solution of

$$(1 - \alpha_1 L - \cdots - \alpha_p L^p) z_t = (1 + \beta_1 L + \cdots + \beta_q L^q) u_t, \qquad (*)$$

which is an autoregressive-moving-average equation, is called an ARMA(p,q).

Let $a(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p$ and $b(L) = 1 + \beta_1 L + \cdots + \beta_q L^q$. The solution of $(*)$ is

$$x_t = a(L)^{-1} b(L) u_t$$

1. Assuming that the roots of $a(L)$ lie outside of the unit circle,

$$a(L)^{-1} b(L) = a(L)^{-1} + \beta_1 a(L)^{-1} L + \cdots + \beta_q a(L)^{-1} L^q, \qquad (**)$$

which clearly shows that $a(L)^{-1} b(L) u_t$ is a one-sided moving average of u_t.

2. $(**)$ also shows that the coefficients of $a(L)^{-1} b(L)$ decline geometrically at rate $|\delta_1|$.

ARMA processes

Note that the roots of the polynomial $1 + \beta_1 L + \cdots + \beta_q L^q$ play no role for the moment. If the roots of $a(L)$ lie outside of the unit circle, then the equation

$$a(L)z_t = b(L)u_t$$

has the solution $a(L)^{-1}b(L)$, which is a one-sided moving average, irrespective of whether the roots of $b(L)$ are outside, on or inside the unit circle.

For example,

$$x_t = (1 + 0.5L)u_t, \quad \text{and} \quad y_t = (1 + 2L)u_t$$

are “equally” stationary. However, for reasons that will be discussed in the sequel we usually assume that ARMA(p,q) processes fulfill the invertibility condition, i.e. that the roots of $b(L)$ lie outside of the unit circle.

ARMA processes

In conclusion, the stationary solution of

$$a(L)z_t = b(L)u_t,$$

where $a(L)$ and $b(L)$ fulfill the stability and the invertibility condition respectively, meaning that their roots lie outside of the unit circle, is called an ARMA(p,q).

Long-memory processes

Now consider the moving average

$$x_t = u_t + \frac{1}{2}u_{t-1} + \frac{1}{3}u_{t-2} + \cdots$$

Is this moving average admissible? The answer is yes, because the sum of the squared coefficients

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$$

converges, although the sum of the coefficients

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots$$

does not.

Long-memory processes

This moving average cannot be produced as the stationary solution of an ARMA equation. For, we know that solution of ARMA equations have coefficients that converge to zero geometrically, as fast as $|\delta_1|^k$. In our case, the coefficients of the moving average converge to zero arithmetically (this is the term used), and are not even summable.

Remember that we already have examples of stationary processes that are not moving averages of a white noise. Now we have an example of a process that is a moving average of a white noise but is not an ARMA.

Given the moving average

$$x_t = \sum_{j=1}^{\infty} a_j u_{t-j},$$

if the coefficients a_j tend to zero geometrically, like in the ARMA case, we say that the process has short memory. If the series $\sum a_j^2$ is summable but $\sum |a_j|$ is not summable, like in the example above, we say that the process has long memory.

References

For stochastic processes, stationarity, ARMA processes, see:

1. Hamilton, J.D. (1994) Time series Analysis, Princeton University Press, Chapters 1, 2, 3.

2. Hansen, B.E. (2008) Econometrics, University of Wisconsin, available on www.ssc.wisc.edu/~bhansen, Chapter 10.

3. Brockwell, P.J. and Davis, R.A. (1996) Time series: Theory and Methods, Springer, Chapters 1 and 3.