Moving Average (MA) & Autoregressive (AR) Models

Ruoqi Yu

Week 5

Section 1

Recap

Big Picture

raw time series $\rightarrow$ stationary process $\rightarrow$ white noise

Pursuing Stationarity

$$f(X_t) = m_t + s_t + X_t$$

- $f()$ Variance stabilizing transform (e.g. $\sqrt{x}$ or $log(x)$)
- m_t deterministic trend (e.g., approximately linear or quadratic)
- s_t deterministic periodic function of know period d, $s_{t+d} = s_t$
- X_t stationary process, e.g. white noise
- Idea: Remove both trend and seasonality so that what remains exhibit stable behavior over time (stability vs stationarity?)
- Instead of deterministic functions, we can also use filters like smoothing and differencing

NEXT!

- We have pursued stationarity (and achieved stability, meaning approximately constant mean and variance)
- Now we can model the autocorrelation structure in this stable series!
- First we will discuss some theory of modeling stationary processes ($\sim$2 weeks)
- Then we'll implement the ideas from theory into applied modeling

NEXT!

One way to think of this next step:

stationary process $\rightarrow$ white noise

means

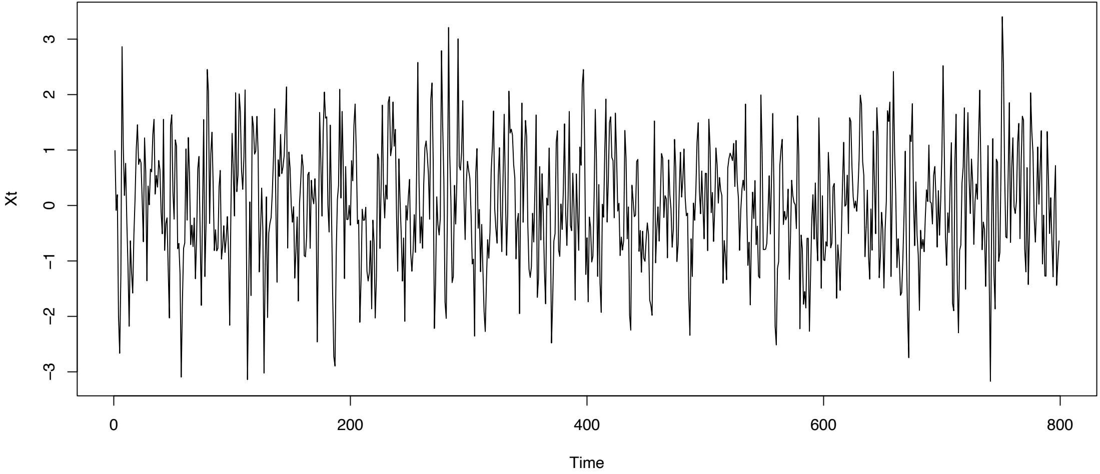
$$X_t = \sum_j a_j W_{t-j}$$

Section 2

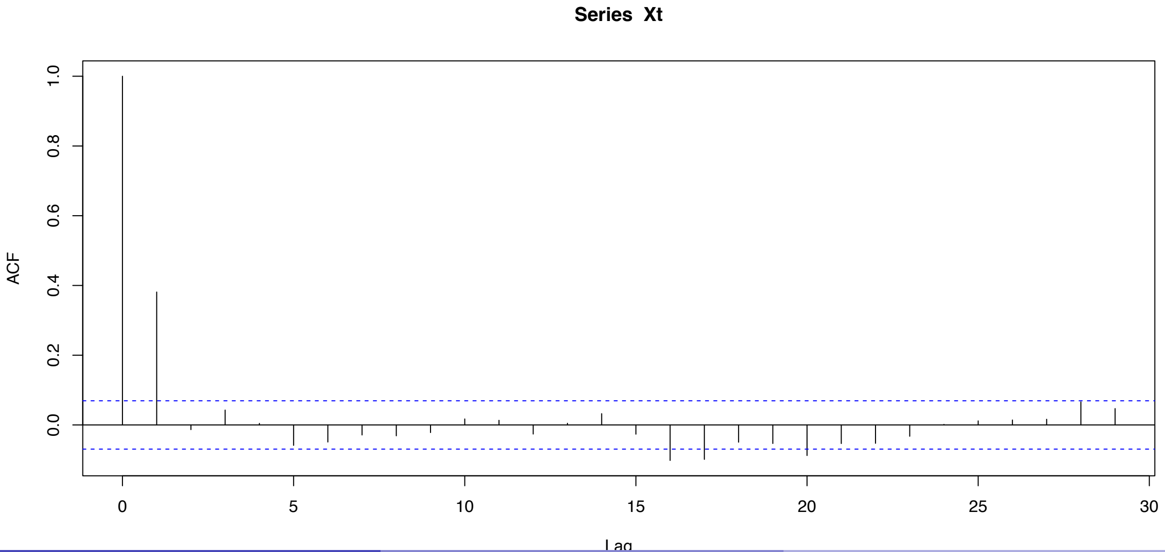
Moving Average models

Motivating Example

- Does this look like White Noise? (does it look stable?)



Motivating Example



So what?

- Given a white noise series $\{W_t\}$ with variance σ^2 and a number $\theta \in R$, set

$$X_t = W_t + \theta W_{t-1}.$$

- This is called a **moving average** of order 1, or MA(1).
- What is the mean? Covariance? Is it stationary?
- Try by yourselves for a few minutes, then we'll derive together

Moving Average Process of Order 1

The series is stationary with mean zero and auto-covariance function (ACVF)

$$\gamma_X(h) = \begin{cases} \sigma^2(1+\theta^2) & h = 0 \\ \theta\sigma^2 & h = 1 \\ 0 & \text{otherwise} \end{cases}$$

As a consequence, X_s and X_t are uncorrelated whenever s and t are two or more time points apart. This time series has *short memory*.

MA(1)

- The autocorrelation function (ACF), for $\{X_t\}$ is given by

$$\rho_X(h) = \frac{\theta}{1+\theta^2}$$

  for $h = 1$ and 0 for $h > 1$.

- What is the maximum value that $\rho_X(1)$ can take?
- The only nonzero value in the theoretical ACF is for lag 1. All other autocorrelations are 0. Thus a sample ACF with a significant autocorrelation only at lag 1 is an indicator of a possible MA(1) model.
- This is our first type of non-white-noise stationary process that we'll explore
- This gives us a tool for modeling noise that has autocorrelation

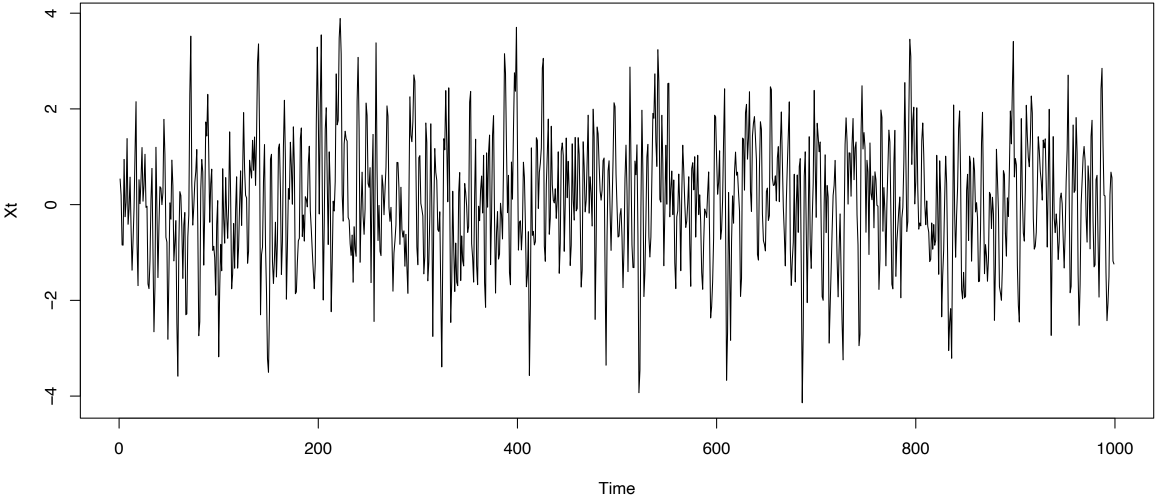
Definition

Let $\ldots, W_{-2}, W_{-1}, W_0, W_1, W_2, \ldots$ be a double infinite white noise sequence. The **moving average model** of order q or **MA(q)** model is defined as

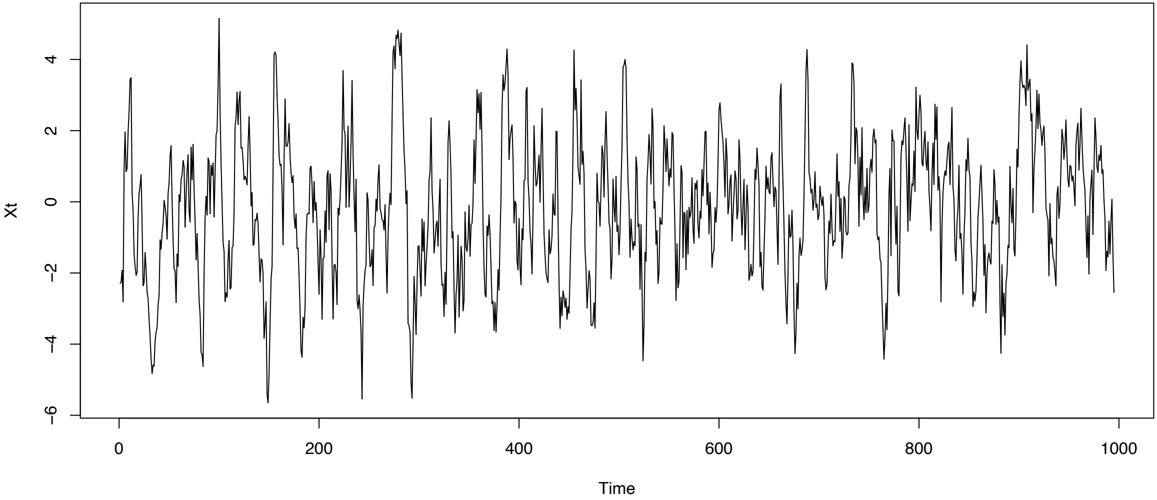
$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}$$

where $\theta_1, \ldots, \theta_q$ are parameters, with $\theta_q \neq 0$.

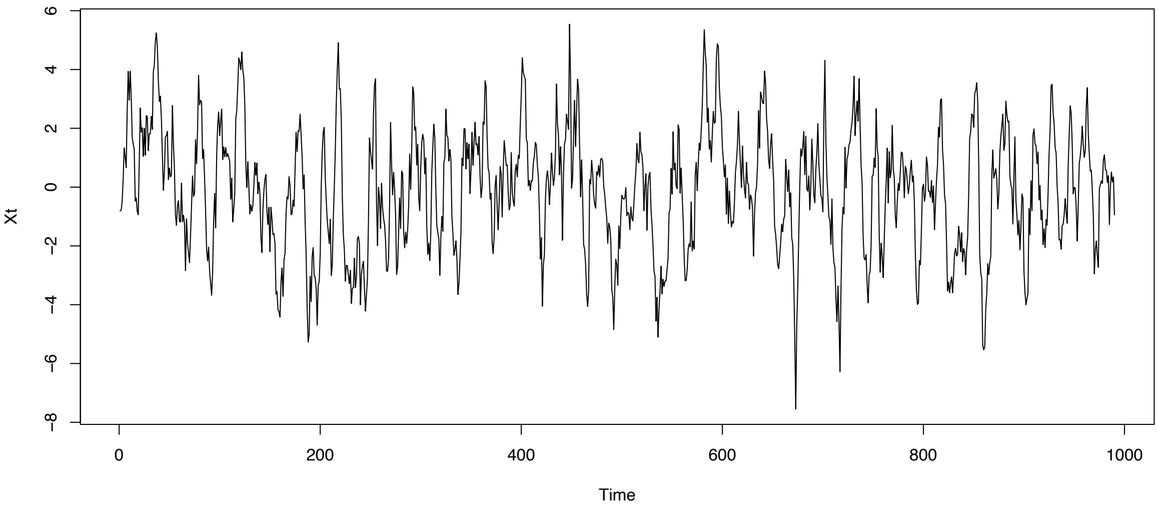
MA(1) $X_t = W_t + 0.9W_{t-1}$



MA(5) $X_t = W_t + 0.9W_{t-1} + \cdots + 0.5W_{t-5}$



MA(9) $X_t = W_t + 0.9W_{t-1} + \cdots + 0.1W_{t-9}$



Autocovariance function of an MA(q) time series:

- The MA(q) model can be concisely written as $X_t = \sum_{j=0}^{q} \theta_j W_{t-j}$ where we take $\theta_0 = 1$.
- The mean of X_t is clearly 0.
- For $h \geq 0$, the covariance between X_t and X_{t+h} is given by

$$\begin{aligned}\text{cov}(X_t, X_{t+h}) &= \text{cov}\left(\sum_{j=0}^{q} \theta_j W_{t-j}, \sum_{k=0}^{q} \theta_k W_{t+h-k}\right) \\ &= \sum_{j=0}^{q} \sum_{k=0}^{q} \theta_j \theta_k \text{cov}(W_{t-j}, W_{t+h-k}).\end{aligned}$$

Autocovariance function of an MA(q) time series:

- Note that because $\{W_t\}$ is white noise, the

$$cov(W_{t-j}, W_{t+h-k}) = \sigma^2 \neq 0$$

  if and only if $t - j = t + h - k$ i.e., if and only if $k = j + h$.

- But because k has to lie between 0 and q, we must have that j has to lie between 0 and $q - h$.
- We thus get:

$$\gamma_X(h) = \begin{cases} \sigma^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h} & h = 0, 1, \ldots, q \\ 0 & if \text{h} > \text{q}. \end{cases}$$

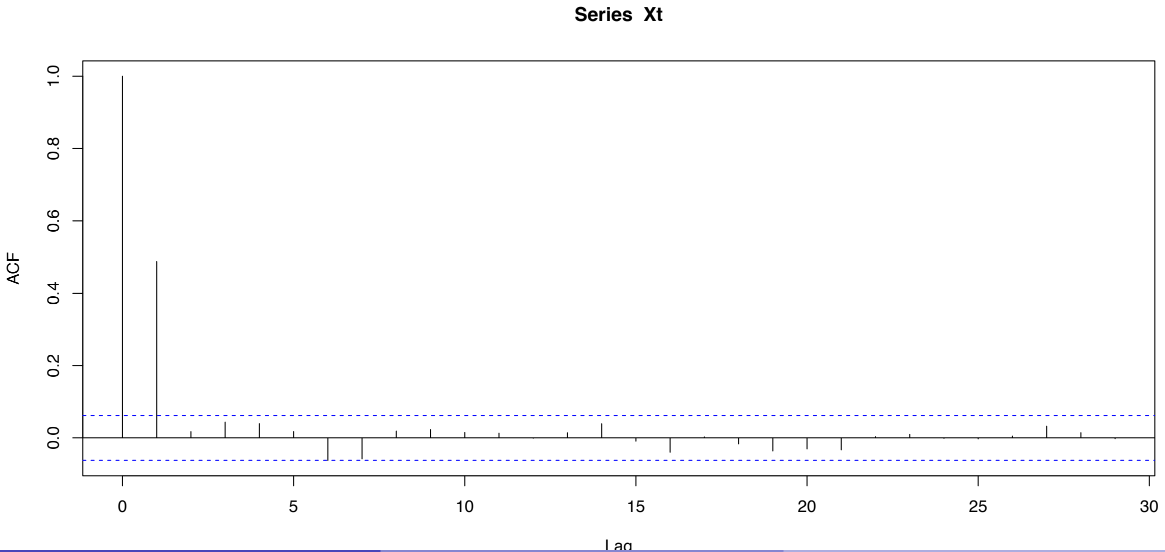
Autocorrelation function of an MA(q) time series:

For the autocorrelation function we thus get

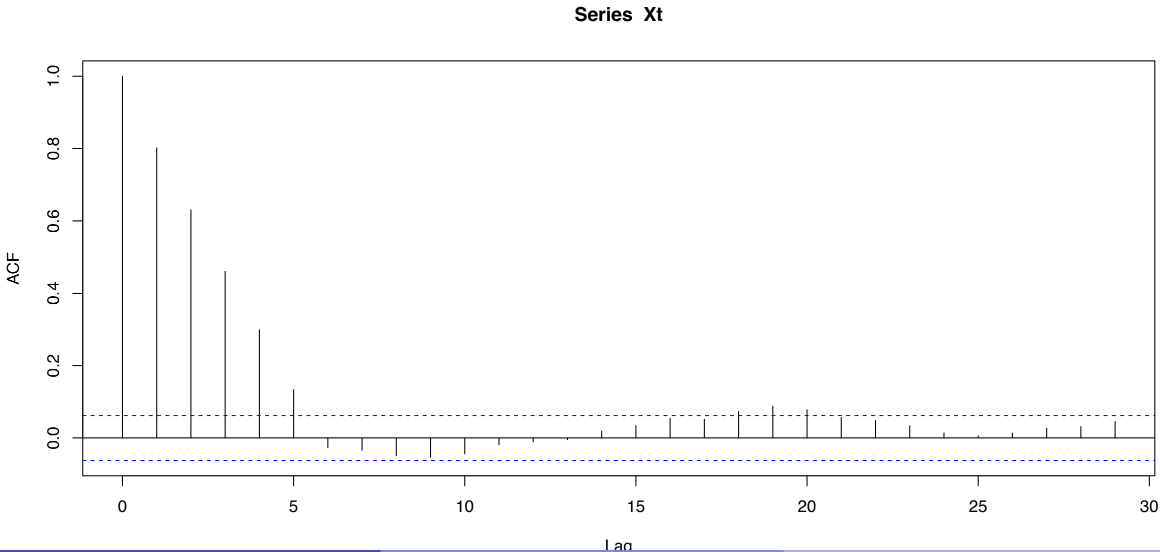
$$\rho_X(h) = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{\sum_{j=0}^{q} \theta_j^2} & h = 0, 1, \ldots, q \\ 0 & h > q \end{cases}$$

Note that the autocovariance and the autocorrelation functions *cut off* after lag q.

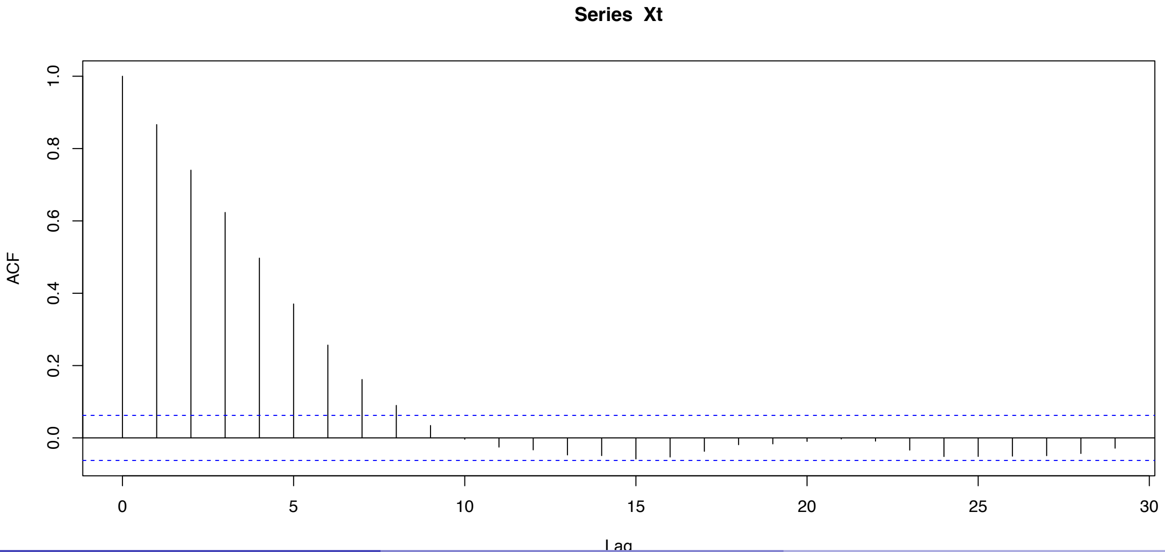
MA(1) $X_t = W_t + 0.9W_{t-1}$



MA(5) $X_t = W_t + 0.9W_{t-1} + \cdots + 0.5W_{t-5}$



MA(9) $X_t = W_t + 0.9W_{t-1} + \cdots + 0.1W_{t-9}$



Theorem: Stationarity of MA(q)

- Theorem: Let $\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots$ be a time series which follows an MA(q) model. Then $\{X_t\}$ is weakly stationary.
- Why?
- Mean is always 0 and $\mathrm{cov}(X_t, X_{t+h})$ does not depend on t, only h.

Moving Average Operator

- The MA(1) process $X_t = W_t + \theta W_{t-1}$ can be written as

$$X_t = \theta(B) W_t$$

  for the polynomial $\theta(z) = 1 + \theta_1 z$.

- Definition: for parameters $\theta_1, \dots, \theta_q$ with $\theta_q \neq 0$ define the **moving average operator** of order q as

$$\theta(B) = 1 + \theta_1 B + \dots \theta_q B^q.$$

- Then we can write the MA(q) model as

$$X_t = \theta(B) W_t,$$

  for a white noise process $\{W_t\}$.

Invertibility: Motivation

- Consider the case of the MA(1) model whose ACVF is given by

$$\begin{aligned}\gamma_X(0) &= \sigma_W^2(1+\theta^2)\\ \gamma_X(1) &= \theta\sigma_W^2\\ \gamma_X(h) &= 0 \text{ for all } h \geq 2.\end{aligned}$$

- Let's say $\theta = 5, \sigma_W^2 = 1$
- But we'd get the same ACVF as for $\theta = 1/5, \sigma_W^2 = 25$.
- In other words, there exist different parameter values that give the same ACVF.
- This implies that one **cannot uniquely** estimate the parameters of an MA(1) model from data.

Invertibility

$$X_t = W_t + \theta W_{t-1}$$

- A natural fix is to consider only those MA(1) for which $|\theta| < 1$:
- This condition is called **invertibility**.

Equivalence of Idea and Definition

- The condition $|\theta| < 1$ for the MA(1) model is equivalent to stating that the moving average polynomial $\theta(z) = 1 + \theta z$ has all roots of magnitude strictly larger than one.
- For $\theta(z) = 1 + \theta z$, force $|\theta| < 1$
- Then for its roots:

$$\text{if } \theta(z) = 0, \text{ then } |z| > 1$$

- The converse carries the same meaning

$$\text{if } |z| \leq 1, \text{ then } \theta(z) \neq 0$$

Definition

An MA(q) model $X_t = \theta(B)W_t$ is said to be **invertible**, if $\theta(z) \neq 0$ for $|z| \leq 1$.

Alternate Definition via Theorem

An MA(q) model $X_t = \theta(B)W_t$ is invertible if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$W_t = \pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j},$$

where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ and $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and $\pi_0 = 1$.

Implication: $AR(\infty)$ representation of invertible $MA(q)$

$$X_t = -\sum_{j=1}^{\infty} \pi_j X_{t-j} + W_t,$$

Example

- Is the following process invertible?

$$X_t = W_t - \frac{11}{8}W_{t-1} + \frac{7}{16}W_{t-2}$$

- What is the autocovariance function $\gamma_Y(h)$ of $Y_t = W_t + 2W_{t-1} - 2W_{t-4}$?

Example Solution

- Yes.
-
$$\gamma_Y(h) = \begin{cases} 9\sigma_W^2 & h = 0 \\ 2\sigma_W^2 & |h| = 1 \\ 0 & |h| = 2 \\ -4\sigma_W^2 & |h| = 3 \\ -2\sigma_W^2 & |h| = 4 \\ 0 & |h| \geq 5. \end{cases}$$

Infinite Order Moving Average: $MA(\infty)$

- This is an $MA(\infty)$ model:

$$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \cdots + \theta_q W_{t-q} + \theta_{q+1} W_{t-q-1} + \dots.$$

  with $\{W_t\}$ as white noise with mean zero and variance σ^2.

- We will write this expression succinctly via

$$X_t = \sum_{j=0}^{\infty} \theta_j W_{t-j}$$

  with θ_0 taken to be 1.

Infinite Order Moving Average

- Infinite sums have convergence issues!
- Note the sum of the infinite geometric series, for $|r| < 1$:

$$a + ar + ar^2 + ar^3 + \ldots = \sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$$

- A sufficient condition which ensures that the infinite sum is finite (almost surely) is $\sum_j |\theta_j| < \infty$.
- In this class, we will always assume this condition when talking about the infinite series $\sum_{j \geq 0} \theta_j W_{t-j}$.

Infinite Order Moving Average

It turns out that $X_t = \sum_{j=0}^{\infty} \theta_j W_{t-j}$ is a stationary process because

$$EX_t = E\left(\sum_{j=0}^{\infty} \theta_j W_{t-j}\right) = \sum_{j=0}^{\infty} \theta_j EW_{t-j} = 0$$

and

$$\begin{aligned} Cov(X_t, X_{t+h}) &= Cov\left(\sum_{j=0}^{\infty} \theta_j W_{t-j}, \sum_{k=0}^{\infty} \theta_k W_{t+h-k}\right) \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \theta_j \theta_k Cov(W_{t-j}, W_{t+h-k}) = \sigma^2 \sum_{j=0}^{\infty} \theta_j \theta_{j+h}. \end{aligned}$$

We could freely interchange the expectation and covariance operators above with the infinite sum because of the condition $\sum_j |\theta_j| < \infty$.

Infinite Order Moving Average

- Note that the expectation EX_t and the covariance $Cov(X_t, X_{t+h})$ do not depend on t and the autocovariance is given by

$$\gamma_X(h) = \sigma^2 \sum_{j=0}^{\infty} \theta_j \theta_{j+h}. \tag{1}$$

  In particular, we get the following

- Thoerem: Let $\ldots, X_{-2}, X_{-1}, X_0, X_1, X_2, \ldots$ be a time series which follows an MA(∞) model. Then $\{X_t\}$ is weakly stationary.

An Interesting $MA(\infty)$

- Fix ϕ with $|\phi| < 1$.
- Choose weights $\theta_j = \phi^j$ in $MA(\infty)$
- $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$
- ACVF:

$$\gamma(h) = \sigma^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+h} = \sigma^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} = \frac{\phi^h \sigma^2}{1 - \phi^2} \text{for } h \geq 0$$

- ACF: $\rho(h) = \phi^h$ for $h \geq 0$.
- Unlike the MA(1), this ACF is strictly non-zero for all lags! But, since $\rho(h)$ drops exponentially as lag increases, this is effectively a stationary time series with short range dependence.

An Interesting $MA(\infty)$

- Here is an important property of this process X_t:

$$\begin{aligned} X_t &= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \dots \\ &= W_t + \phi\left(W_{t-1} + \phi W_{t-2} + \phi^2 W_{t-3} + \dots\right) \\ &= W_t + \phi X_{t-1} \text{ for every } t = \dots, -1, 0, 1, \dots. \end{aligned}$$

- Thus X_t satisfies the following first order *difference equation*:

$$X_t = \phi X_{t-1} + W_t.$$

- For this reason, $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$ is called the **Stationary Autoregressive Process of order one**.

Section 3

Autoregressive Model

Definition of AR(p)

Let $\ldots, W_{-2}, W_{-1}, W_0, W_1, W_2, \ldots$ be a double infinite white noise sequence. The **autoregressive model** of order p or **AR(p)** model is of the form

$$X_t = W_t + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p},$$

where $\phi_1, \ldots, \phi_p$ with $\phi_p \neq 0$ are parameters.

Autoregressive Operator

- We can write the AR(p) model as

$$\phi(B)X_t = W_t,$$

for a white noise process $\{W_t\}$.

- For parameters $\phi_1, \dots, \phi_p$ with $\phi_p \neq 0$ define the **autoregressive operator** of order p as

$$\phi(B) = 1 - \phi_1 B - \dots \phi_p B^p.$$

AR(1) Process

- We will first look at AR(1) processes which satisfy the difference equation

$$X_t - \phi X_{t-1} = W_t.$$

  or equivalently

$$X_t = \phi X_{t-1} + W_t.$$

- Previously seen that when $|\phi| < 1$ the MA(∞) process $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$ solves this difference equation.
- Is it the only solution to the difference equation above?
- No!

What do we mean by solution?

- In practice (empirically/with data), we consider X_t as our residuals.
- Theoretically, however, we're look at an equation that involves white noise (whose properties we understand) and a sequence of unknown random variables,

$$..., X_{t-1}, X_t, X_{t+1}, ...$$

- Thus, we're solving for X, similar to high school algebra class.

Another Solution to $X_t = \phi X_{t-1} + W_t$

- Define X_0 to be an arbitrary random variable that is uncorrelated with the white noise series $\{W_t\}$ and define $X_1, X_2, \dots$ as well as $X_{-1}, X_{-2}, \dots$ using the difference equation $X_t = \phi X_{t-1} + W_t$.
- The resulting sequence surely satisfies $X_t = \phi X_{t-1} + W_t$. Is it stationary?
- NO! Because $X_{-1} = X_0/\phi - W_0/\phi$ and since $|\phi| < 1$ and X_0 and W_0 are uncorrelated, this would give var$(X_{-1}) >$ var(X_0), contradicting stationarity.
- $X_t = \phi X_{t-1} + W_t$ with $|\phi| < 1$ has many solutions but only one stationary solution.

Theorem on AR Stationarity

For some white noise process $\{W_t\}$ and fixed parameter $|\phi| \neq 1$ there exists exactly one time series process $\{X_t\}$ with mean zero which is stationary and solves the difference equation

$$X_t - \phi X_{t-1} = W_t.$$

Sidebar

- Before we prove this theorem, let us analyze what the unique stationary solution of the difference equation is in a rather more heuristic way.
- The difference equation $X_t - \phi X_{t-1} = W_t$ can be rewritten as $\phi(B)X_t = W_t$ where $\phi(B)$ is given by the polynomial $\phi(z) = 1 - \phi z$. Therefore, it is natural that the solution of this equation is

$$X_t = \frac{1}{\phi(B)} W_t.$$

- First consider $|\phi| < 1$. From the formula for the sum of a geometric series, we have

$$\frac{1}{\phi(z)} = (1 - \phi z)^{-1} = 1 + \phi z + \phi^2 z^2 + \phi^3 z^3 + \ldots$$

Sidebar

- As a result, we expect as a stationary solution

$$
\begin{aligned}
X_t &= \frac{1}{\phi(B)} W_t \\
&= \left(I + \phi B + \phi^2 B^2 + \dots\right) W_t \\
&= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \dots \\
&= \sum_{j=0}^{\infty} \phi^j W_{t-j}.
\end{aligned}
$$

Sidebar

- Second consider $|\phi| > 1$. Here, we can write

$$\begin{aligned}
\frac{1}{\phi(z)} &= \frac{1}{1-\phi z} \\
&= \frac{-1}{\phi z}\left(1 - \frac{1}{\phi z}\right)^{-1} \\
&= -\frac{1}{\phi z} - \frac{1}{\phi^2 z^2} - \frac{1}{\phi^3 z^3} - \dots \\
&= -\frac{z^{-1}}{\phi} - \frac{z^{-2}}{\phi^2} - \frac{z^{-3}}{\phi^3} - \dots
\end{aligned}$$

Sidebar

- As a result, we expect as a stationary solution

$$
\begin{aligned}
X_t &= \left(-\frac{B^{-1}}{\phi} - \frac{B^{-2}}{\phi^2} - \frac{B^{-3}}{\phi^3} - \dots\right) W_t \\
&= -\frac{W_{t+1}}{\phi} - \frac{W_{t+2}}{\phi^2} - \frac{W_{t+3}}{\phi^3} - \dots
\end{aligned}
$$

- This is indeed true and we will prove this in the following. The strange part about the equation above is that X_t depends on only future white noise values: $W_{t+1}, W_{t+2}, \dots$.
- As a result, autoregressive processes of order 1 for $|\phi| > 1$ are rarely used in time series modelling.

Proof

- We only present the proof for $|\phi| < 1$. The case for $|\phi| > 1$ is analog.
- We have seen that $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$ is one stationary solution of the difference equation.
- Suppose $\{Y_t\}$ is any other stationary sequence which also satisfies the difference equation, so that we want to show $X_t = Y_t$ is the unique stationary solution. i.e. $Y_t = \phi Y_{t-1} + W_t$.
- In that case, by successively using this equation, we obtain

$$
\begin{aligned}
Y_t &= W_t + \phi Y_{t-1} \\
&= W_t + \phi W_{t-1} + \phi^2 Y_{t-2} \\
&= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \phi^3 Y_{t-3} \\
&= W_t + \phi W_{t-1} + \phi^2 W_{t-2} + \phi^3 W_{t-3} + \phi^4 Y_{t-4} \\
&= \vdots
\end{aligned}
$$

Proof (continued)

- In general, for *every* k, one would have

$$Y_t = \left[\sum_{i=0}^{k} \phi^i W_{t-i}\right] + \phi^{k+1} Y_{t-k-1}$$

- The idea is now to let k approach ∞.
- The first term on the right hand side is

$$\sum_{i=0}^{k} \phi^i W_{t-i}$$

which we have argued converges to $X_t = \sum_{i=0}^{\infty} \phi^i W_{t-i}$ as k goes to infinity.

- If the second term, $\phi^{k+1} Y_{t-k-1}$, goes to 0 as $k \to \infty$, then $Y_t = X_t$ and we're done. We'll do this with mean-square convergence.

Proof (continued) - Mean-Square Convergence

- We want to show
$$\lim_{k\to\infty} E\left[\left(\phi^{k+1}Y_{t-k-1} - 0\right)^2\right] = 0$$
- First note that $E\left[\left(\phi^{k+1}Y_{t-k-1}\right)^2\right] = \phi^{2k+2}EY^2_{t-k-1}$
- We assumed $\{Y_t\}$ is stationary, which means it has time-invariant (constant) mean and variance, implying $E(Y_t^2)$ is time-invariant too as $Var(Y_t) = E(Y_t^2) - [E(Y_t)]^2$. Hence $EY^2_{t-k-1} = EY^2_a$ for any fixed integer a. Let $a = 0:$
$$\phi^{2k+2}EY^2_{t-k-1} = \phi^{2k+2}EY^2_0$$
- As EY_0^2 is a constant and $|\phi| < 1$:
$$\lim_{k\to\infty} E\left[\left(\phi^{k+1}Y_{t-k-1} - 0\right)^2\right] = \lim_{k\to\infty} \phi^{2k+2}EY_0^2 = 0$$
- It follows therefore that Y_t and X_t are the same.

Proof (continued)

- Finally, consider the case $|\phi| = 1$
- Here the difference equation becomes $X_t - X_{t-1} = W_t$ for $\phi = 1$ and $X_t + X_{t-1} = W_t$ for $\phi = -1$.
- These difference equations have **no** stationary solutions.
- Let us see this for $\phi = 1$ (the $\phi = -1$ case is similar).
- Note that $X_t = X_{t-1} + W_t$ means that

$$var(X_t) = var(X_{t-1}) + var(W_t)$$

  as X_{t-1}, W_t are uncorrelated.
- If $var(W_t) > 0$, then $var(X_t) > var(X_{t-1})$. This cannot happen if $\{X_t\}$ were stationary.

AR(1) Summary

1. If $|\phi| < 1$, the difference equation has a unique stationary solution given by $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$. The solution clearly only depends on the present and past values of $\{W_t\}$. It is hence called **causal**.
2. If $|\phi| > 1$, the difference equation has a unique stationary solution given by $X_t = -\sum_{j=1}^{\infty} \phi^{-j} W_{t+j}$. This is **non-causal**.
3. If $|\phi| = 1$, no stationary solution exists.

Reinterpreted Summary

This summary can be reinterpreted in terms of the polynomial $\phi(z) = 1 - \phi z$. The root of this polynomial is $1/\phi$.

1. If the magnitude of the root of $\phi(z)$ is strictly larger than 1, then $\phi(B)X_t = W_t$ has a unique **causal** stationary solution.
2. If the magnitude of the root of $\phi(z)$ is strictly smaller than 1, then $\phi(B)X_t = W_t$ has a unique stationary solution which is **non-causal**.
3. If the magnitude of the root of $\phi(z)$ is exactly equal to one, then $\phi(B)X_t = W_t$ has no stationary solution.

Causality

- Akin to the invertiblity condition for MA(q), we can define the causality condition for general AR(p) processes.
- Definition: An AR(p) model $\phi(B)X_t = W_t$ is said to be **causal**, if $\phi(z) \neq 0$ for $|z| \leq 1$.
- Analog to the invertibility theorem, one gets the following equivalent definition.

Thoerem on Causality: $MA(\infty)$ Representation of Causal AR(p)

An AR(p) model $\phi(B)X_t = W_t$ is causal if and only if the time series $\{X_t\}$ and the white noise $\{W_t\}$ can be written as

$$X_t = \psi(B)W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j},$$

where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and $\psi_0 = 1$.

For an causal $AR(p)$ model there exists a unique stationary solution which only depends on the past values of W_t!

Summary

- $MA(q)$ process:\ invertible ($\theta(z) \neq 0$ for all $|z| \leq 1$) , $AR(\infty)$ representation $\Longrightarrow$ parameters uniquely identified
- AR(p) process:\ causal ($\phi(z) \neq 0$ for all $|z| \leq 1$) , $MA(\infty)$ representation $\Longrightarrow$ unique stationary solution which only depends on the past.
- In the following we will only consider invertible $MA(q)$ models and causal $AR(p)$ models!

Section 4

ARMA

ARMA(p,q)

Definition: A (zero mean) *autoregressive moving average* model of order p and q is of the form

$$\phi(B)X_t = \theta(B)W_t$$

where $\phi(B)$ is the AR operator, $\theta(B)$ is the MA operator, and $\{W_t\}$ is white noise.

Basic ARMA Models

1. White noise ($X_t = W_t$) is ARMA(0,0), with $\phi(z) = \theta(z) = 1$
2. Moving Average is ARMA(0,q), with $\phi(z) = 1$ and $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + ... + \theta_q z^q$
3. Autoregression is ARMA(p,0), with $\theta(z) = 1$ and $\phi(z) = 1 + \phi_1 z + \phi_2 z^2 + ... + \phi_q z^q$

ARMA(p,q)

Rearranged for forecasting:

$$X_t = \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + ...\theta_q W_{t-q}$$

Causal and Invertible

An $ARMA(p, q)$ model (B)X_t $=$ (B)W_t is said to be

1. **invertible** if $\theta(z) \neq 0$ for all $|z| \leq 1$;
2. **causal** if $\phi(z) \neq 0$ for all $|z| \leq 1$.

Example (TSA4e 3.8)

- Is the following process causal and/or invertible?

$$X_t = .4X_{t-1} + .45X_{t-2} + W_t + W_{t-1} + .25W_{t-2}$$

- Move like terms: $X_t - .4X_{t-1} - .45X_{t-2} = W_t + W_{t-1} + .25W_{t-2}$
- Put in operator form: $(1 - .4B - .45B^2)X_t = (1 + B + .25B^2)W_t$

Example (TSA4e 3.8)

- Factor polynomials: $(1 + .5B)(1 - .9B)X_t = (1 + .5B)^2 W_t$
- Cancel common factors: $(1 - .9B)X_t = (1 + .5B)W_t$
- Turns out the original process can be reduced!! To

$$X_t = .9X_{t-1} + W_t + .5W_{t-1}$$

Example (TSA4e 3.8)

- Cancel common factors: $(1-.9B)X_t = (1+.5B)W_t$
- $\theta(z) = 1 + .5B$ has root -2, so it's invertible!
- $\phi(z) = 1 - .9B$ has root $\frac{10}{9}$, so it's causal!

Code

- ARMAacf()
- arima.sim()