Autoregressive Moving Average Models (ARMA)[1],[2]

The ARMA models were proposed by Whittle (1951) to capture the correlation that may be generated through lagged linear relations in time series.

ARMA models can accurately approximate stationary processes:

For any stationary process with autocovariance γ such that

$$\lim_{h\to\infty} \gamma(h) = 0,$$

and any integer k>0, there is an ARMA process $\{X_t\}$ for which

$$\gamma_X(h) = \gamma(h),\ h = 0,1,\ldots,k.$$

ARMA processes are generated by using white noise as the 'forcing terms' in a set of linear difference equations with constant coefficients.

[1] Robert Shumway, David Stoffer. Time Series Analysis and its Applications with R Examples, 4th edition, Springer, 2017 (chapter 3)

[2] Peter Brockwell, Richard Davis. Time Series: Theory and Methods, Springer-Verlag, 1987 (chapter 3)

ARMA

<u>Def</u>. The backshift operator B is defined by

$$B X_t = X_{t-1}$$

and it is extended to powers $B^2 X_t = B(B X_t) = X_{t-2}$ and so on. Thus, $B^k X_t = X_{t-k}$.

The inverse operator B^{-1} is the forward-shift operator

$$X_t = B^{-1} B X_t = B^{-1} X_{t-1}$$

Autoregressive Models (AR)

They are based on the assumption that X_t can be explained as a function of p past values $X_{t-1}, \ldots, X_{t-p}$.

<u>Def.</u> An AR(p) model is of the form

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

X_t is stationary, $w_t \sim wn(0, \sigma_w^2)$ and $\phi_1, \ldots, \phi_p$ are parameters ($\phi_p \neq 0$).

Using the backshift operator, an AR(p) has the form

$$\underbrace{\left(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p\right)}_{\overset{\text{not}}{=}\ \phi(B) \text{ the autoregressive operator}} \cdot X_t = w_t$$

AR(1) and causality

Example The AR(1) model

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

or equivalent, iterating backwards k times

$$X_t = \phi^k X_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j}$$

I Provided that $|\phi| < 1$ and $\mathrm{Var}(X_t) < \infty$, we have that

$$\lim_{k\to\infty} E\left[\left(X_t - \sum_{j=0}^{k-1} \phi^j \cdot w_{t-j}\right)^2\right] = \lim_{k\to\infty} E\left[\left(\phi^k X_{t-k}\right)^2\right] =$$

$$= \lim_{k\to\infty} \phi^{2k} E\left[(X_{t-k})^2\right] = 0$$

$$\left[X_n \xrightarrow[ms]{\text{mean square}} X \quad \text{iff} \quad E\left[(X_n - X)^2\right] \xrightarrow[n\to\infty]{} 0\right]$$

AR(1) and causality

So, $X_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}$ — AR(1) can be represented as a linear process

$$E(X_t) = \sum_{j=0}^{\infty} \phi^j E(w_{t-j}) = 0$$

$$\gamma(h) = \operatorname{cov}(X_{t+h}, X_t) = E\left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j}\right)\left(\sum_{k=0}^{\infty} \phi^k w_{t-k}\right)\right]$$

$$= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \cdot \phi^j = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j}$$

$$= \frac{\sigma_w^2 \cdot \phi^h}{1-\phi^2}, \quad h \geq 0$$

AR(1) and causality

So, $X_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}$ — AR(1) can be represented as a linear process

$$E(X_t) = \sum_{j=0}^{\infty} \phi^j E(w_{t-j}) = 0$$

$$\gamma(h) = cov(X_{t+h}, X_t) = E\left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j}\right)\left(\sum_{k=0}^{\infty} \phi^k w_{t-k}\right)\right]$$

$$= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \cdot \phi^j = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j}$$

$$= \frac{\sigma_w^2 \cdot \phi^h}{1-\phi^2}, \quad h \geq 0$$

For a stationary process, $\gamma(h) = \gamma(-h)$ so we only exhibit it for $h \geq 0$.

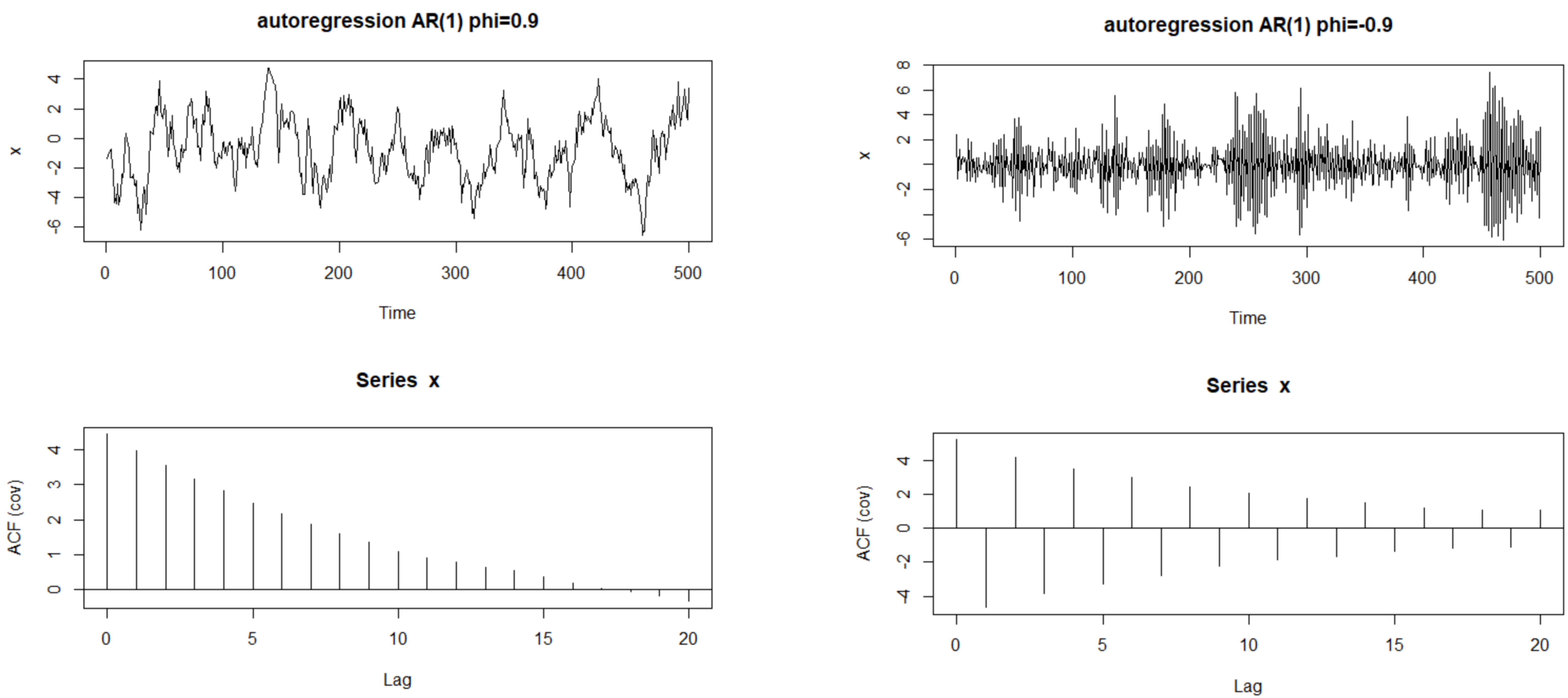
The ACF $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h, \quad h \geq 0$

$$\rho(h) = \phi \rho(h-1), \quad h = 1, 2, \ldots$$

AR(1) and causality

Example Two AR(1) processes, one with $\phi = 0.9$ and the other with $\phi = -0.9$; in both cases $\sigma_w^2 = 1$.

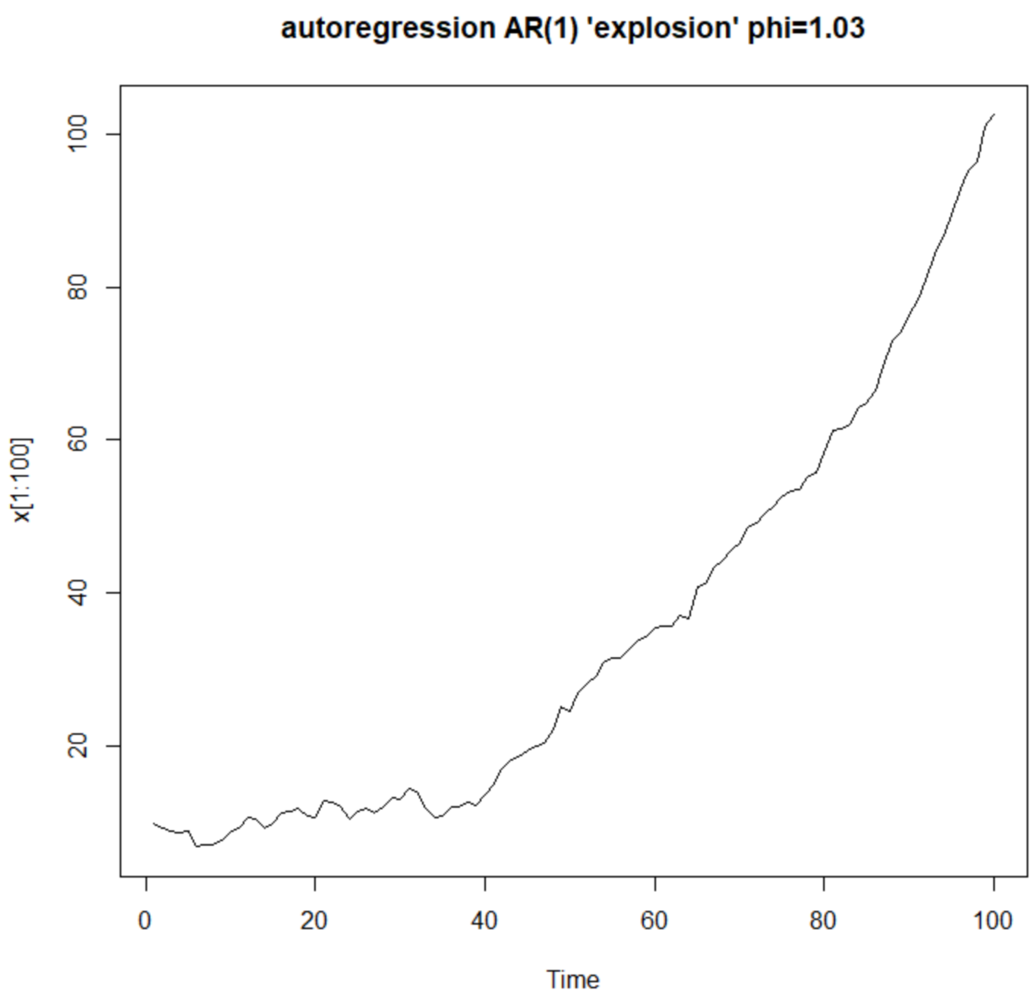
$$x_t = \phi x_{t-1} + w_t$$



AR(1) and causality

II if $|\phi|>1$, the process is called "explosive" because the values of the time series grow exponentially.

$|\phi|^j \xrightarrow[j\to\infty]{} \infty \implies \sum_{j=0}^{k-1} \phi^j w_{t-j}$ does not converge in mean square as $k\to\infty$



AR(1) and causality

Still, we can obtain a stationary model as follows:

$$X_{t+1} = \phi X_t + w_{t+1}$$

$$X_t = \frac{1}{\phi} X_{t+1} - \frac{1}{\phi} w_{t+1} = \dots$$

$$= \phi^{-k} X_{t+k} - \sum_{j=1}^{k-1} \phi^{-j} w_{t+j}$$

AR(1) and causality

Because $|\phi|^{-1} < 1$, then

$$X_t = -\sum_{j=1}^{\infty} \phi^{-j} w_{t+j}$$

$$E(X_t) = 0$$

$$\gamma(h) = \text{cov}(X_{t+h}, X_t) = \text{cov}\left(-\sum_{j=1}^{\infty} \phi^{-j} w_{t+h+j}, -\sum_{k=1}^{\infty} \phi^{-k} w_{t+k}\right)$$

$$= \frac{\sigma_w^2 \phi^{-h} \cdot \phi^{-2}}{1-\phi^{-2}}$$

So, X_t is stationary.

We notice that X_t is correlated with $\{w_s\}_{s>t}$, which is unnatural.

AR(1) and causality

<u>Def</u>. When a process does not depend on the future, such as AR(1) when $|\phi|<1$, we call it <u>causal</u>.

<u>Obs</u>. Every "explosion" has a "cause"

$$X_t = \phi X_{t-1} + w_t \quad , \quad |\phi|>1 \quad \text{and} \quad w_t \sim \text{iid } N(0, \sigma_w^2)$$

The causal process defined by

$$Y_t = \phi^{-1} Y_{t-1} + V_t \quad , \quad V_t \sim \text{iid } N(0, \sigma_w^2 \phi^{-2})$$

is stochastically equivalent to X_t.

AR(1) and causality

Example. $X_t = 2X_{t-1} + w_t$ with $\sigma_w^2 = 1$ and
$Y_t = \frac{1}{2} Y_{t-1} + v_t$ with $\sigma_v^2 = \frac{1}{4}$ are equivalent (Y_t is causal!)

Excluding explosive models from consideration is not a problem because they have causal counterparts.

III if $|\phi| = 1$, the AR(1) has no stationary solution.

Moving Average Models (MA)

They assume that on the right-hand side of the defining equation, there is a linear combination of white noise.

<u>Def</u>. An MA(q) model is of the form

$$X_t = w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}, \text{ where}$$

$w_t \sim wn(0, \sigma_w^2)$ and $\theta_1, \dots, \theta_q$ are parameters ($\theta_q \neq 0$).

in an equivalent form,

$$X_t = \underbrace{(1 + \theta_1 B + \dots + \theta_q B^q)}_{\overset{\text{not}}{=}\ \theta(B) \text{ the moving average operator}} w_t$$

Moving Average Models (MA)

<u>Obs.</u> MA is stationary for any values of $\theta_1, \dots, \theta_q$.

<u>Example</u> The MA(1) model

$$X_t = w_t + \theta w_{t-1}$$

$$E(X_t) = 0$$

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma_w^2 & , h=0 \\ \theta \sigma_w^2 & , h=1 \\ 0 & , h>1 \end{cases}$$

$$\rho(h) = \begin{cases} 1 & , h=0 \\ \frac{\theta}{1+\theta^2} & , h=1 \\ 0 & , h>1 \end{cases}$$

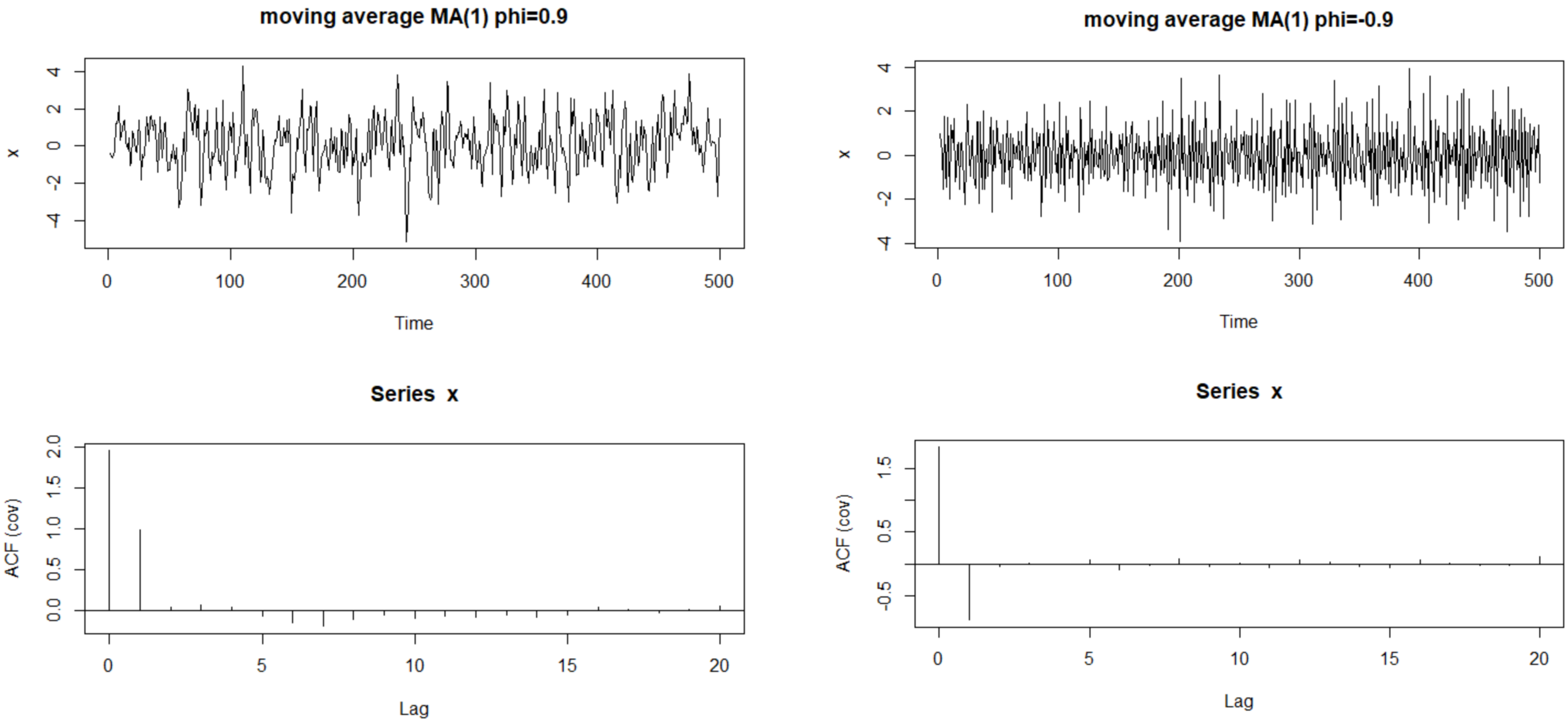
Obs. $|\rho(1)| \le \frac{1}{2} \quad \forall \theta$

X_t is correlated with X_{t-1}, but not with $X_{t-2}, X_{t-3}, \dots$

(compare it to AR(1))

Moving Average Models (MA)

Example Two MA(1) processes, one with $\theta = 0.9$ and the other with $\theta = -0.9$; in both cases $\sigma_w^2 = 1$.



Moving Average Models (MA)

<u>Non-uniqueness of MA models</u>

<u>Obs</u>. For MA(1), $\rho(h)$ is the same for θ and $\frac{1}{\theta}$

$\gamma(h)$ is the same for $\left(\sigma_w^2 = 1 ; \theta = 5\right)$ and $\left(\sigma_w^2 = 25 ; \theta = \frac{1}{5}\right)$.

Thus, the MA(1) processes

$$X_t = W_t + \frac{1}{5} W_{t-1}, \quad W_t \sim \text{iid } N(0, 25)$$

$$Y_t = V_t + 5 V_{t-1}, \quad V_t \sim \text{iid } N(0, 1)$$

are stochastically equivalent because of normality (all finite distributions are the same).

Moving Average Models (MA)

By mimicking the criterion of causality for AR models, we choose the MA model with an infinite AR representation. Such a process is called <u>invertible</u>.

To find out which model is the invertible one, we reverse the roles of x_t and w_t

$$w_t = -\theta w_{t-1} + x_t$$

if $|\theta| < 1$, then $w_t = \sum_{j=0}^{\infty} (-\theta)^j x_{t-j}$ - that is the infinite AR representation of the model.

So, we will choose the MA(1) model with $\theta = \frac{1}{5}$, $\sigma_w^2 = 25$ because it is invertible.

ARMA

They are models for stationary time series

<u>Def</u>. A time series $\{X_t\}$, $t = 0, \pm 1, \pm 2 \dots$ is ARMA (p, q) if it is stationary and

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q},$$

with $\phi_p \neq 0$, $\theta_q \neq 0$ and $\sigma_w^2 > 0$, $w_t \sim wn(0, \sigma_w^2)$.

It can be written more concisely as $\phi(B) X_t = \theta(B) w_t$.

ARMA

Def. The AR and MA polynomials are defined as:

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \quad , \ \phi_p \neq 0$$

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \quad , \ \theta_q \neq 0$$

where z is a complex number.

Def. An ARMA(p,q) model is called causal if X_t can be written as

$$X_t = \sum_{j=0}^{\infty} \psi_j \cdot w_{t-j} = \psi(B) w_t ,$$

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

ARMA

<u>Theorem 1</u> An ARMA(p,q) model is causal iff $\phi(z) \neq 0$ for $|z| \leq 1$. The coefficients ψ can be determined by solving

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j \cdot z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1.$$

A stationary solution X_t to the ARMA equation $\Phi(B)X_t=\Theta(B)W_t$ exists iff Φ has no roots on the unit circle.

Causality implies stationarity.

ARMA

<u>Def</u>. An ARMA(p,q) model is called invertible if X_t can be written as

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

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

<u>Theorem 2</u> An ARMA(p,q) model is invertible iff $\theta(z) \neq 0$ for $|z| \leq 1$. The coefficients π can be determined by solving

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1$$

<u>Obs</u>. The condition "$\phi(z) \neq 0$ for $|z| \leq 1$" can be rephrased as "the roots of $\phi(z)$ lie outside the unit circle".

ARMA

Example ARMA(1, 1) model

$$X_t = 0.9\, X_{t-1} + 0.5\, W_{t-1} + W_t$$

$$(1 - 0.9B)\, X_t = (1 + 0.5B)\, W_t$$

$$\phi(z) = (1 - 0.9z)$$
$$\theta(z) = (1 + 0.5z)$$

The model is causal because $\phi(z) = 1 - 0.9z = 0$ when $z = 10/9$ (outside the unit circle)

ARMA

ψ - weights are obtained from $\phi(z) \cdot \psi(z) = \theta(z)$

$$(1-0.9z)(1+\psi_1 z + \psi_2 z^2 + \dots) = 1 + 0.5z$$

$$1 + (\psi_1 - 0.9)z + (\psi_2 - 0.9\psi_1)z^2 + \dots = 1 + 0.5z$$

$$\psi_1 - 0.9 = 0.5$$

$$\psi_j - 0.9\psi_{j-1} = 0 \, , \text{ for } j > 1$$

$$\psi_j = 1.4 \cdot (0.9)^{j-1} \text{ for } j \geq 1$$

$$x_t = w_t + 1.4 \sum_{j=1}^{\infty} 0.9^{j-1} w_{t-j}$$

The model is invertible because $\theta(z) = (1 + 0.5z) = 0$ when $z = -2$ (outside the unit circle)

ARMA

π - weights are obtained from $\theta(z)\,\pi(z) = \phi(z)$

$$(1+0.5z)(1+\pi_1 z + \pi_2 z^2 + \dots) = 1 - 0.9z$$

$$1 + (\pi_1 + 0.5)z + (\pi_2 + 0.5\pi_1)z^2 + \dots = 1 - 0.9z$$

$$\pi_1 + 0.5 = -0.9$$

$$\pi_j + 0.5\pi_{j-1} = 0 \quad \text{for } j > 1$$

$$\pi_j = (-1.4)\cdot(-0.5)^{j-1}, \quad j \geq 1$$

$$W_t = X_t - 1.4 \sum_{j=1}^{\infty} (-0.5)^{j-1} \cdot X_{t-j}$$

Particularizations of Theorem 1

1) The MA(q) process is causal $\forall\, \theta_1, \dots, \theta_q$.

2) The process AR(1) is causal iff $\phi(z) = 1 - \phi_1 z$ has its root outside the unit circle, i.e. $|\phi_1| < 1$.

Particularizations of Theorem 1

3) The AR(2) process $X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = w_t$ is causal iff $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$ has its roots outside the unit circle.

if we denote the roots by z_1 and z_2, then $|z_1| > 1$ and $|z_2| > 1$.
z_1 and z_2 may be real or a complex conjugate pair.

Viète formulas

$$\left.\begin{aligned} z_1 + z_2 &= -\frac{\phi_1}{\phi_2} \\ z_1 \cdot z_2 &= -\frac{1}{\phi_2} \end{aligned}\right\} \Rightarrow \quad \begin{aligned} \phi_1 &= z_1^{-1} + z_2^{-1} \\ \phi_2 &= -(z_1 z_2)^{-1} \end{aligned}$$

Particularizations of Theorem 1

in the complex case

$$z_1 = x + iy \qquad |z_1| > 1, \ |z_2| > 1 \ (\text{that is } x^2+y^2>1)$$
$$z_2 = x - iy$$

$$\phi_1 + \phi_2 = \frac{1}{z_1} + \frac{1}{z_2} - \frac{1}{z_1 z_2} = \frac{z_1+z_2-1}{z_1 z_2} = \frac{2x-1}{x^2+y^2} < 1 \quad \text{because}$$

$$\text{from } x^2+y^2>1 \Rightarrow (x-1)^2+y^2>0$$

$$\phi_2 - \phi_1 = -\frac{1}{z_1 z_2} - \frac{1}{z_1} - \frac{1}{z_2} = \frac{-2x-1}{x^2+y^2} < 1 \quad \text{because}$$

$$\text{from } x^2+y^2>1 \Rightarrow (x+1)^2+y^2>0$$

$$|\phi_2| = \frac{1}{|z_1 z_2|} = \frac{1}{x^2+y^2} < 1$$

Similar reasoning for the real case.