Intro to ARMA models

FISH 507 – Applied Time Series Analysis

Mark Scheuerell

15 Jan 2019

Topics for today

Review

- White noise
- Random walks

Autoregressive (AR) models

Moving average (MA) models

Autoregressive moving average (ARMA) models

Using ACF & PACF for model ID

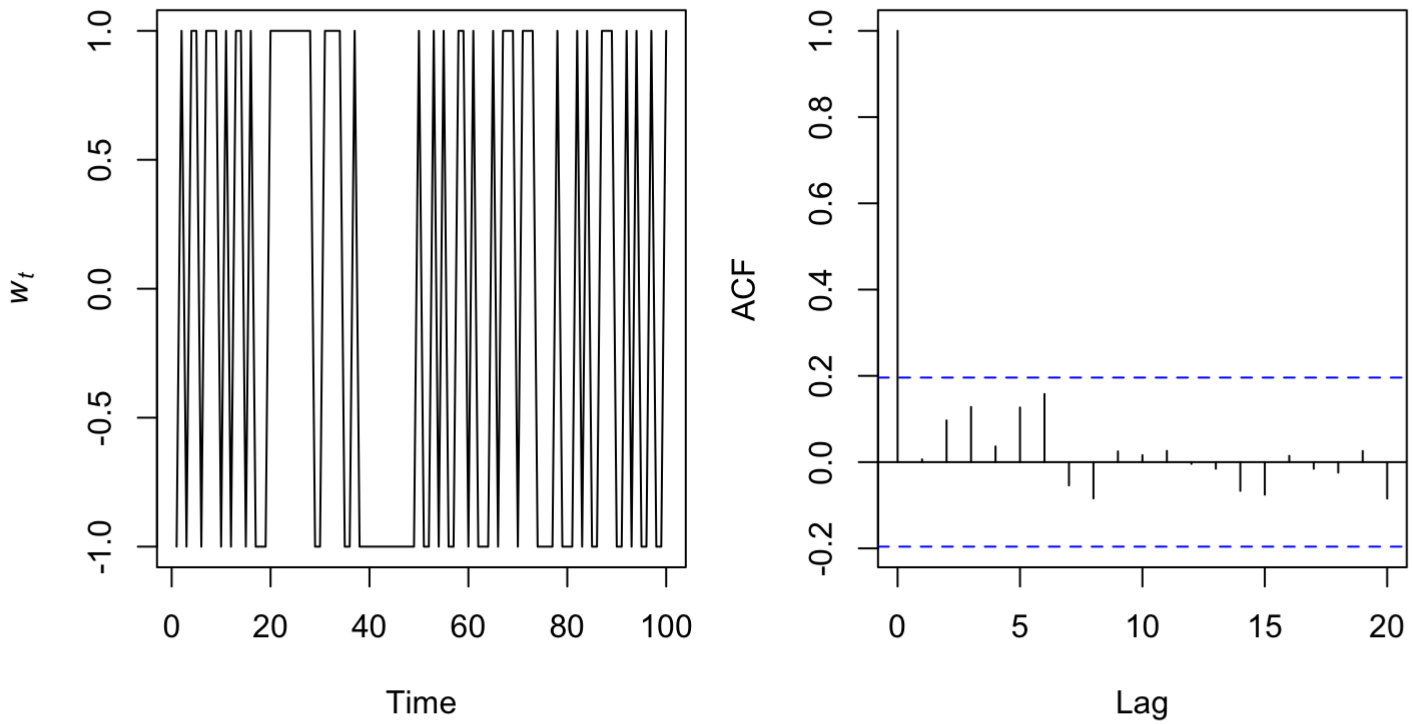
White noise (WN)

A time series $\{w_t\}$ is discrete white noise if its values are

1. independent
2. identically distributed with a mean of zero

The distributional form for $\{w_t\}$ is flexible

White noise (WN)



$w_t = 2e_t - 1; e_t \sim \text{Bernoulli}(0.5)$

Gaussian white noise

We often assume so-called *Gaussian white noise*, whereby

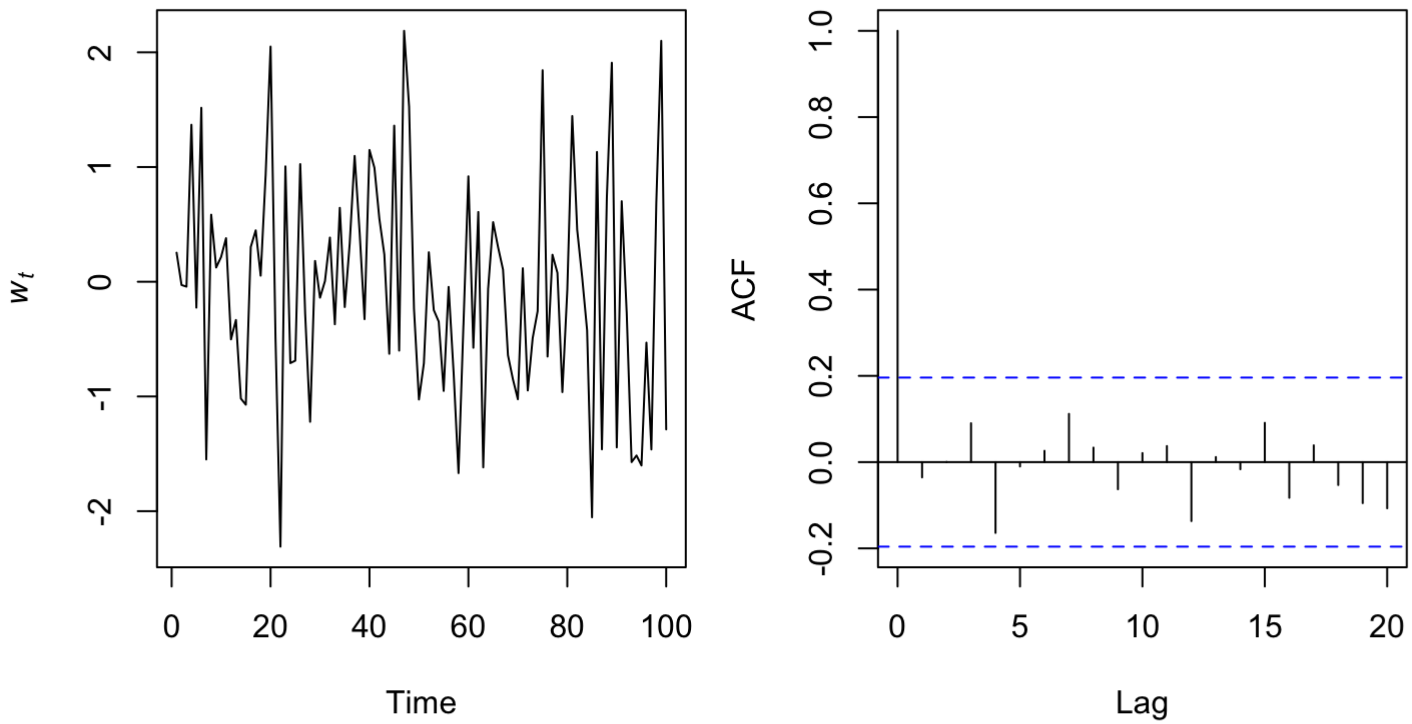
$$w_t \sim \mathrm{N}(0, \sigma^2)$$

and the following apply as well

autocovariance: $\gamma_k = \begin{cases} \sigma^2 & \text{if } k = 0 \\ 0 & \text{if } k \geq 1 \end{cases}$

autocorrelation: $\rho_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \geq 1 \end{cases}$

Gaussian white noise



$w_t \sim N(0, 1)$

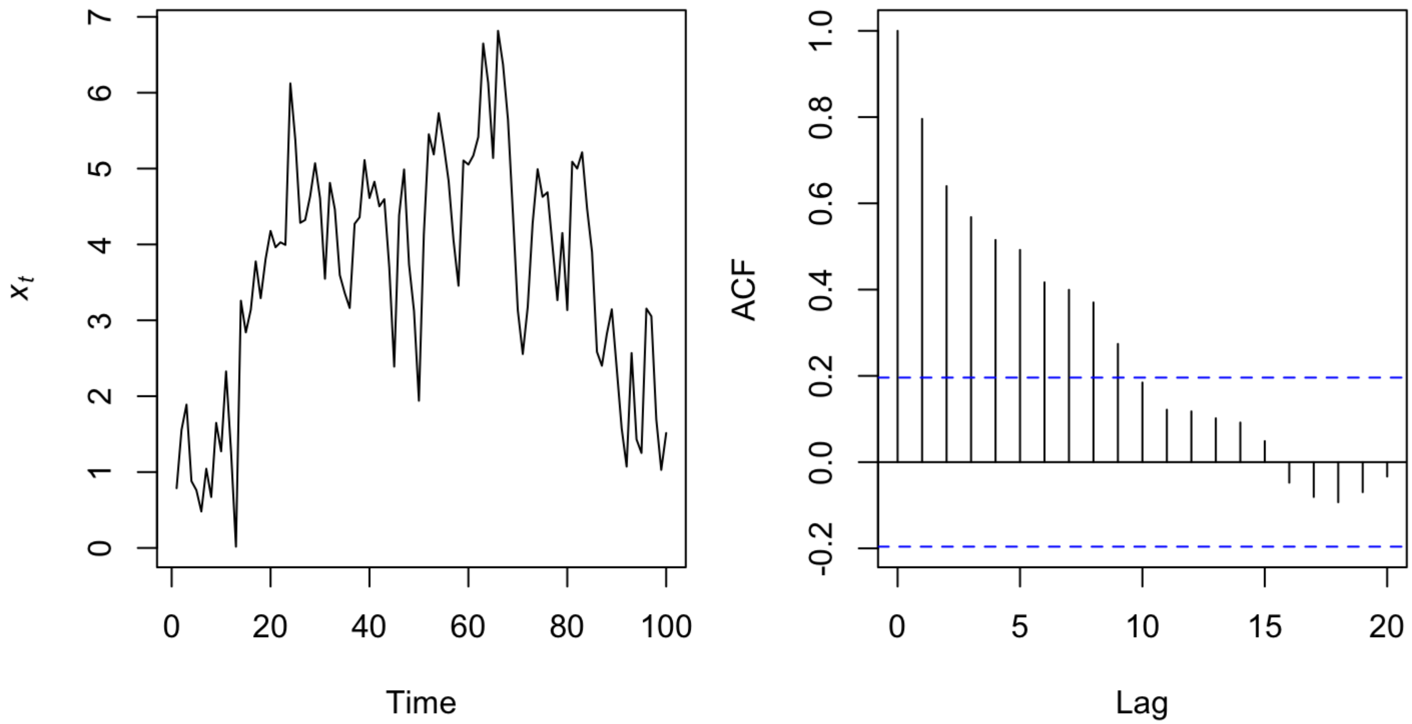
Random walk (RW)

A time series $\{x_t\}$ is a random walk if

1. $x_t = x_{t-1} + w_t$

2. w_t is white noise

Random walk (RW)



$$x_t = x_{t-1} + w_t; w_t \sim N(0, 1)$$

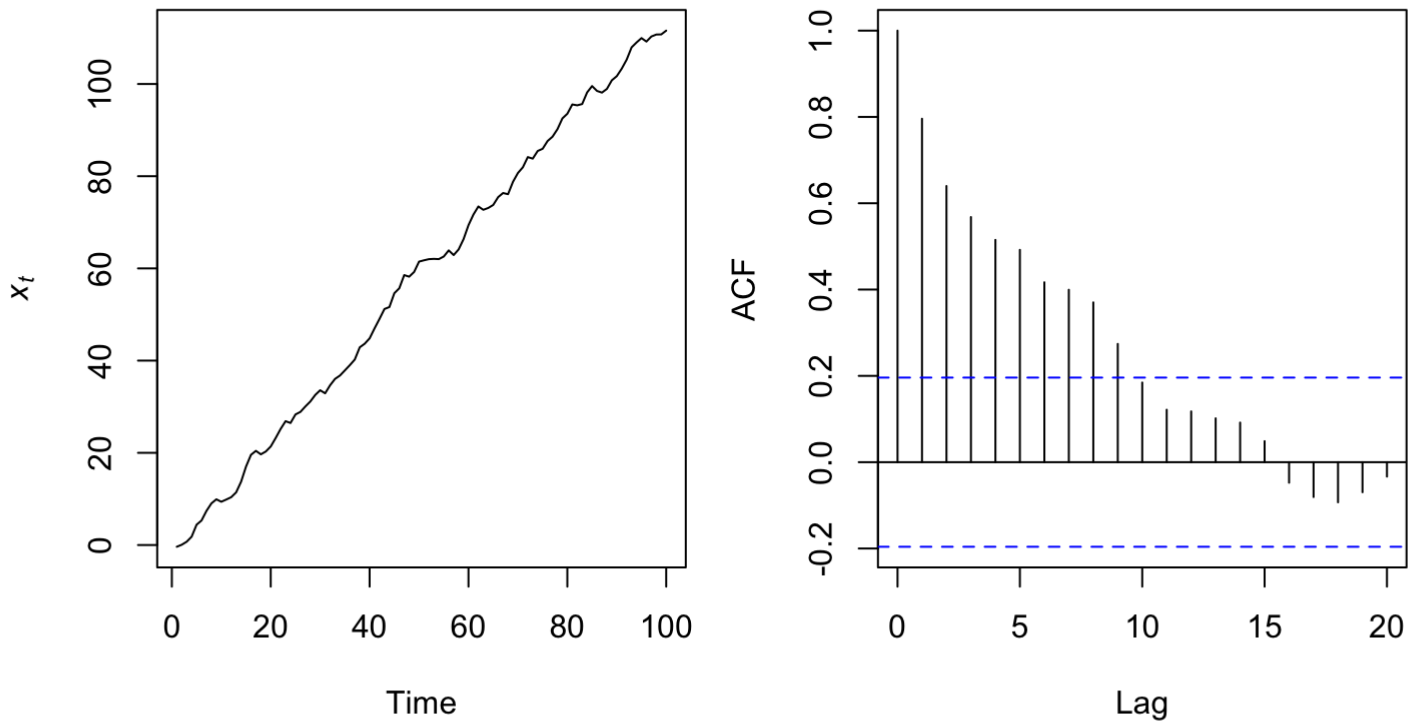
Biased random walk

A *biased random walk* (or *random walk with drift*) is written as

$$x_t = x_{t-1} + u + w_t$$

where u is the bias (drift) per time step and w_t is white noise

Biased random walk



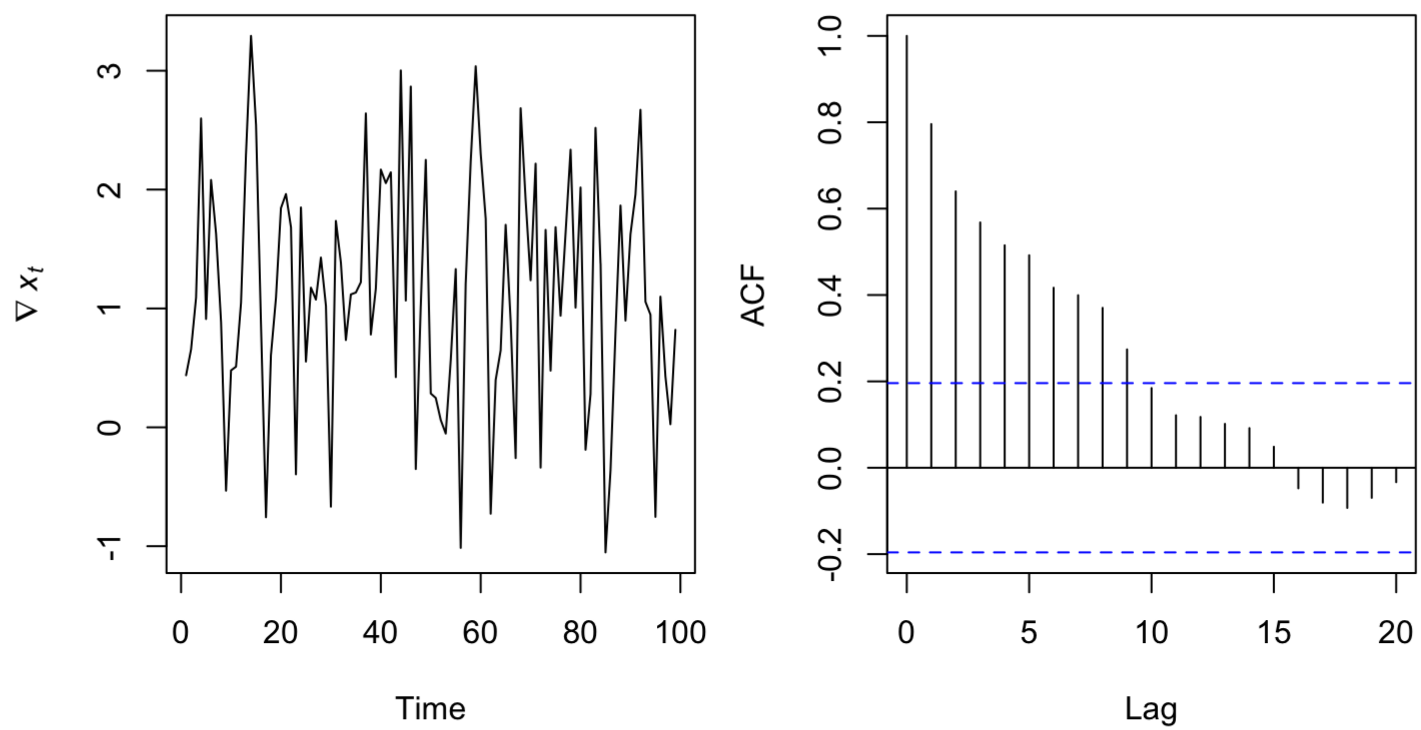
$$x_t = x_{t-1} + 1 + w_t ; w_t \sim N(0, 1)$$

Differencing a biased random walk

First-differencing a biased random walk yields a constant mean (level) u plus white noise

$$
\begin{aligned}
\nabla x_t &= x_{t-1} + u + w_t \\
x_t - x_{t-1} &= x_{t-1} + u + w_t - x_{t-1} \\
x_t - x_{t-1} &= u + w_t
\end{aligned}
$$

Differencing a biased random walk



$x_t = x_{t-1} + 1 + w_t; w_t \sim N(0, 1)$

LINEAR STATIONARY MODELS

Linear stationary models

We saw last week that linear filters are a useful way of modeling time series

Here we extend those ideas to a general class of models call *autoregressive moving average* (ARMA) models

Autoregressive (AR) models

An *autoregressive* model of order *p*, or AR(*p*), is defined as

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t$$

where we assume

1. w_t is white noise
2. $\phi_p \neq 0$ for an order-*p* process

Examples of AR(p) models

AR(1)

$x_t = 0.5x_{t-1} + w_t$

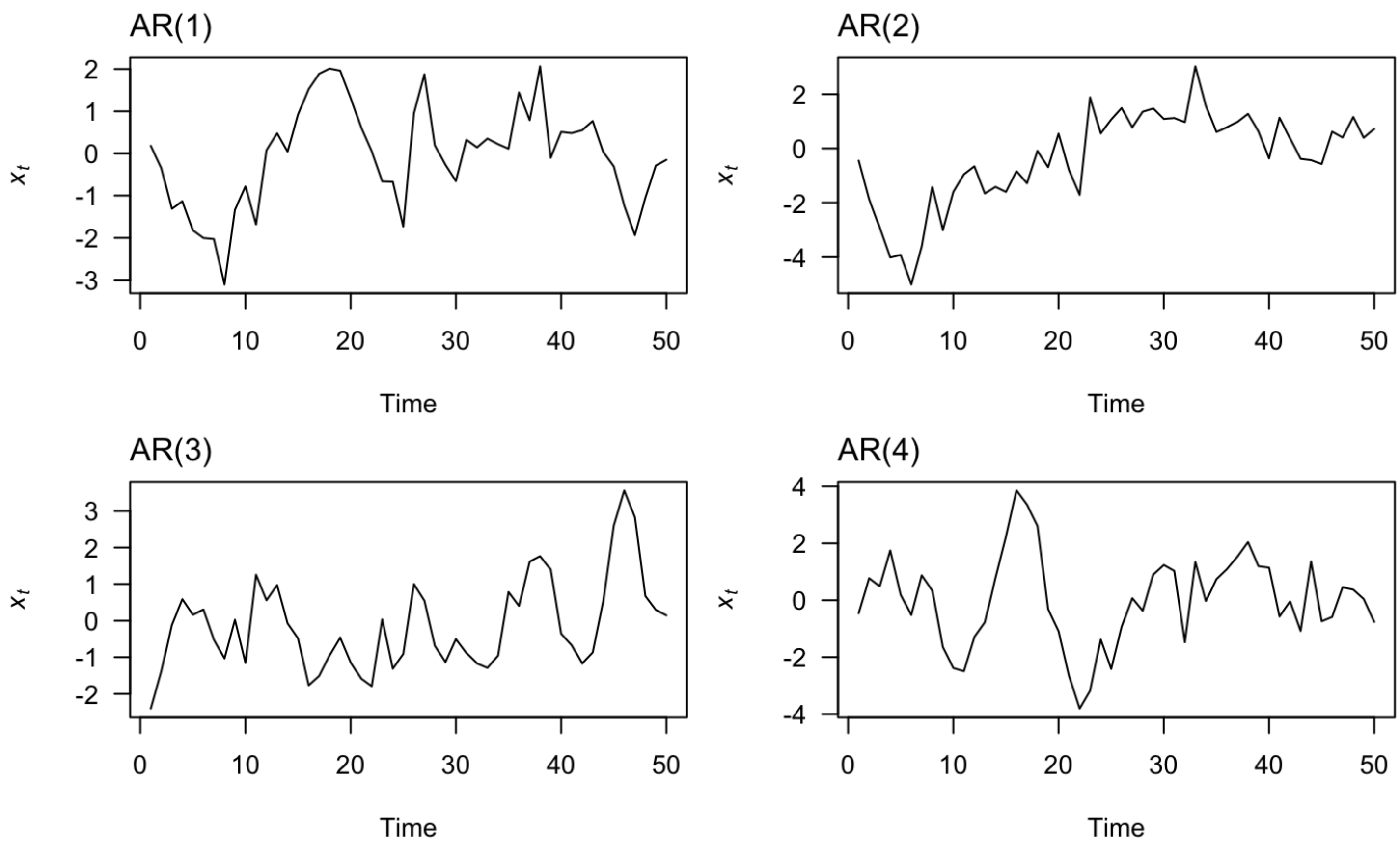
AR(1) with $\phi_1 = 1$ (random walk)

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

AR(2)

$x_t = -0.2x_{t-1} + 0.4x_{t-2} + w_t$

Examples of AR(p) models



Stationary AR(p) models

Recall that *stationary* processes have the following properties

1. no systematic change in the mean or variance
2. no systematic trend
3. no periodic variations or seasonality

We seek a means for identifying whether our AR(p) models are also stationary

Stationary AR(p) models

We can write out an AR(p) model using the backshift operator

$$
\begin{aligned}
x_t &= \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t \\
&\Downarrow \\
x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \dots - \phi_p x_{t-p} &= w_t \\
(1 - \phi_1 \mathbf{B} - \phi_2 \mathbf{B}^2 - \dots - \phi_p \mathbf{B}^p) x_t &= w_t \\
\phi_p(\mathbf{B}) x_t &= w_t
\end{aligned}
$$

Stationary AR(p) models

If we treat $\mathbf{B}$ as a number (or numbers), we can out write the *characteristic equation* as

$$
\begin{gathered}
\phi_p(\mathbf{B}) x_t = w_t \\
\Downarrow \\
\phi_p(\mathbf{B}) = 0
\end{gathered}
$$

To be stationary, **all roots** of the characteristic equation **must exceed 1 in absolute value**

Stationary AR(p) models

For example, consider this AR(1) model from earlier

$$\begin{aligned} x_t &= 0.5 x_{t-1} + w_t \\ x_t - 0.5 x_{t-1} &= w_t \\ (1 - 0.5 \mathbf{B}) x_t &= w_t \end{aligned}$$

Stationary AR(p) models

For example, consider this AR(1) model from earlier

$$
\begin{aligned}
x_t &= 0.5 x_{t-1} + w_t \\
x_t - 0.5 x_{t-1} &= w_t \\
(1 - 0.5 \mathbf{B}) x_t &= w_t \\
&\Downarrow \\
1 - 0.5 \mathbf{B} &= 0 \\
-0.5 \mathbf{B} &= -1 \\
\mathbf{B} &= 2
\end{aligned}
$$

This model is indeed stationary because $\mathbf{B} > 1$

Stationary AR(p) models

What about this AR(2) model from earlier?

$$\begin{aligned} x_t &= -0.2x_{t-1} + 0.4x_{t-2} + w_t \\ x_t + 0.2x_{t-1} - 0.4x_{t-2} &= w_t \\ (1 + 0.2\mathbf{B} - 0.4\mathbf{B}^2)x_t &= w_t \end{aligned}$$

Stationary AR(p) models

What about this AR(2) model from earlier?

$$
\begin{aligned}
x_t &= -0.2 x_{t-1} + 0.4 x_{t-2} + w_t \\
x_t + 0.2 x_{t-1} - 0.4 x_{t-2} &= w_t \\
(1 + 0.2 \mathbf{B} - 0.4 \mathbf{B}^2) x_t &= w_t \\
&\Downarrow \\
1 + 0.2 \mathbf{B} - 0.4 \mathbf{B}^2 &= 0 \\
&\Downarrow \\
\mathbf{B} \approx -1.35 \text{ and } \mathbf{B} &\approx 1.85
\end{aligned}
$$

This model is *not* stationary because only one $\mathbf{B} > 1$

What about random walks?

Consider our random walk model

$$
\begin{aligned}
x_t &= x_{t-1} + w_t \\
x_t - x_{t-1} &= w_t \\
(1 - 1\mathbf{B})x_t &= w_t
\end{aligned}
$$

What about random walks?

Consider our random walk model

$$
\begin{aligned}
x_t &= x_{t-1} + w_t \\
x_t - x_{t-1} &= w_t \\
(1 - 1\mathbf{B})x_t &= w_t \\
&\Downarrow \\
1 - 1\mathbf{B} &= 0 \\
-1\mathbf{B} &= -1 \\
\mathbf{B} &= 1
\end{aligned}
$$

Random walks are **not** stationary because $\mathbf{B} = 1 \ngtr 1$

Stationary AR(1) models

We can define a space over which all AR(1) models are stationary

Stationary AR(1) models

For $x_t = \phi x_{t-1} + w_t$, we have

$$
\begin{aligned}
1 - \phi \mathbf{B} &= 0 \\
-\phi \mathbf{B} &= -1 \\
\mathbf{B} &= \frac{1}{\phi} > 1 \Rightarrow 0 < \phi < 1
\end{aligned}
$$

Stationary AR(1) models

For $x_t = \; x_{t-1} + w_t$, we have

$$
\begin{aligned}
1 - \; \mathbf{B} &= 0 \\
- \; \mathbf{B} &= -1 \\
\mathbf{B} &= \frac{1}{\;} > 1 \Rightarrow 0 < \; < 1
\end{aligned}
$$

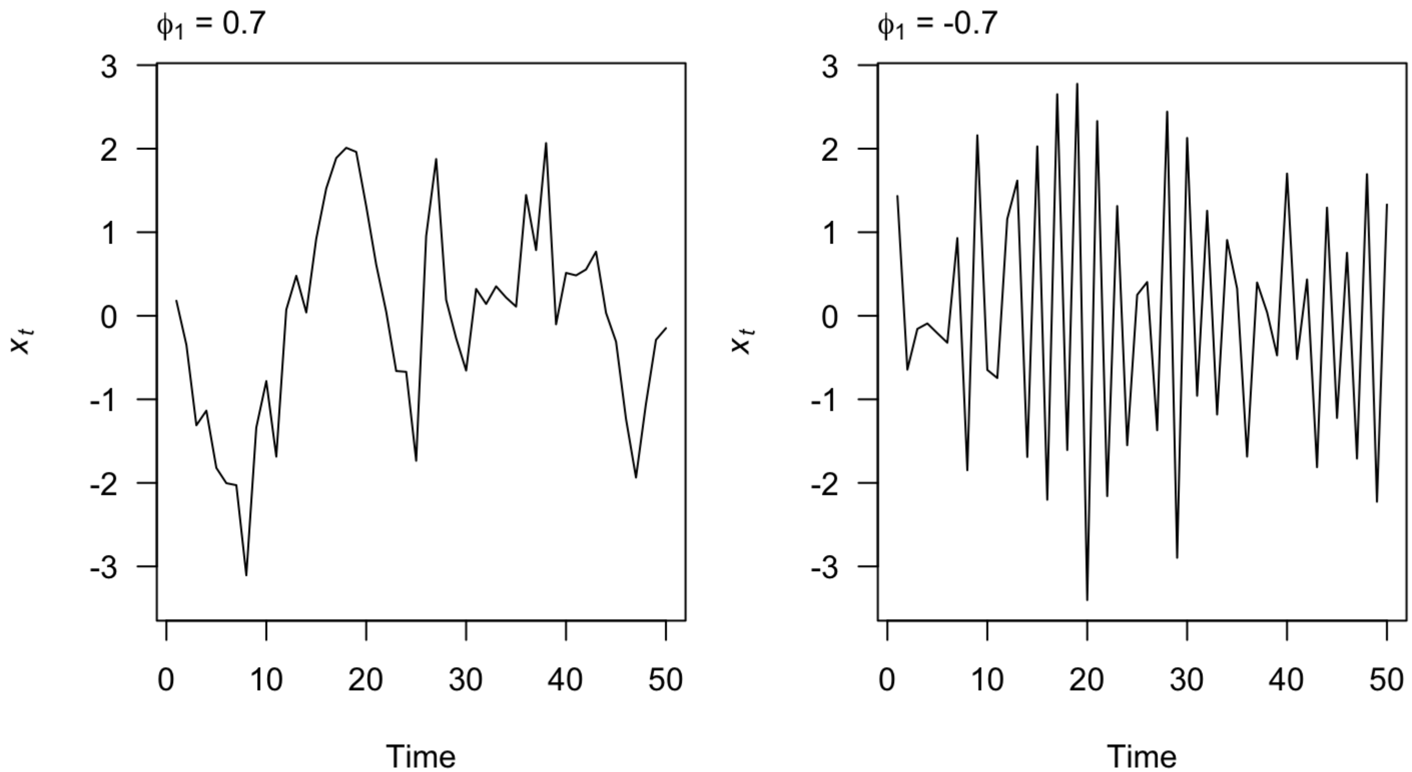
For $x_t = - \; x_{t-1} + w_t$, we have

$$
\begin{aligned}
1 + \; \mathbf{B} &= 0 \\
\; \mathbf{B} &= -1 \\
\mathbf{B} &= \frac{-1}{\;} > 1 \Rightarrow -1 < \; < 0
\end{aligned}
$$

Stationary AR(1) models

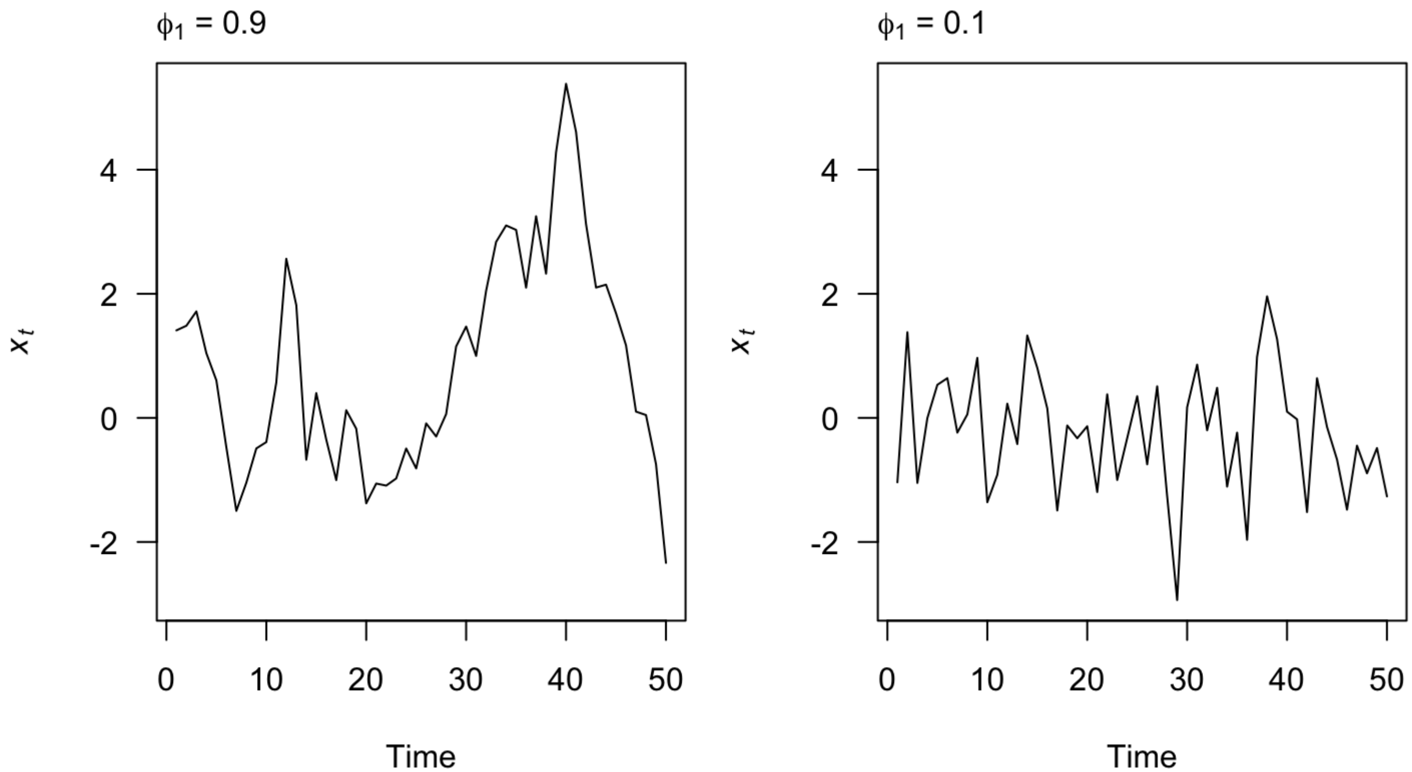
Thus, AR(1) models are stationary if and only if $| \quad | < 1$

Coefficients of AR(1) models



Same value, but different sign

Coefficients of AR(1) models

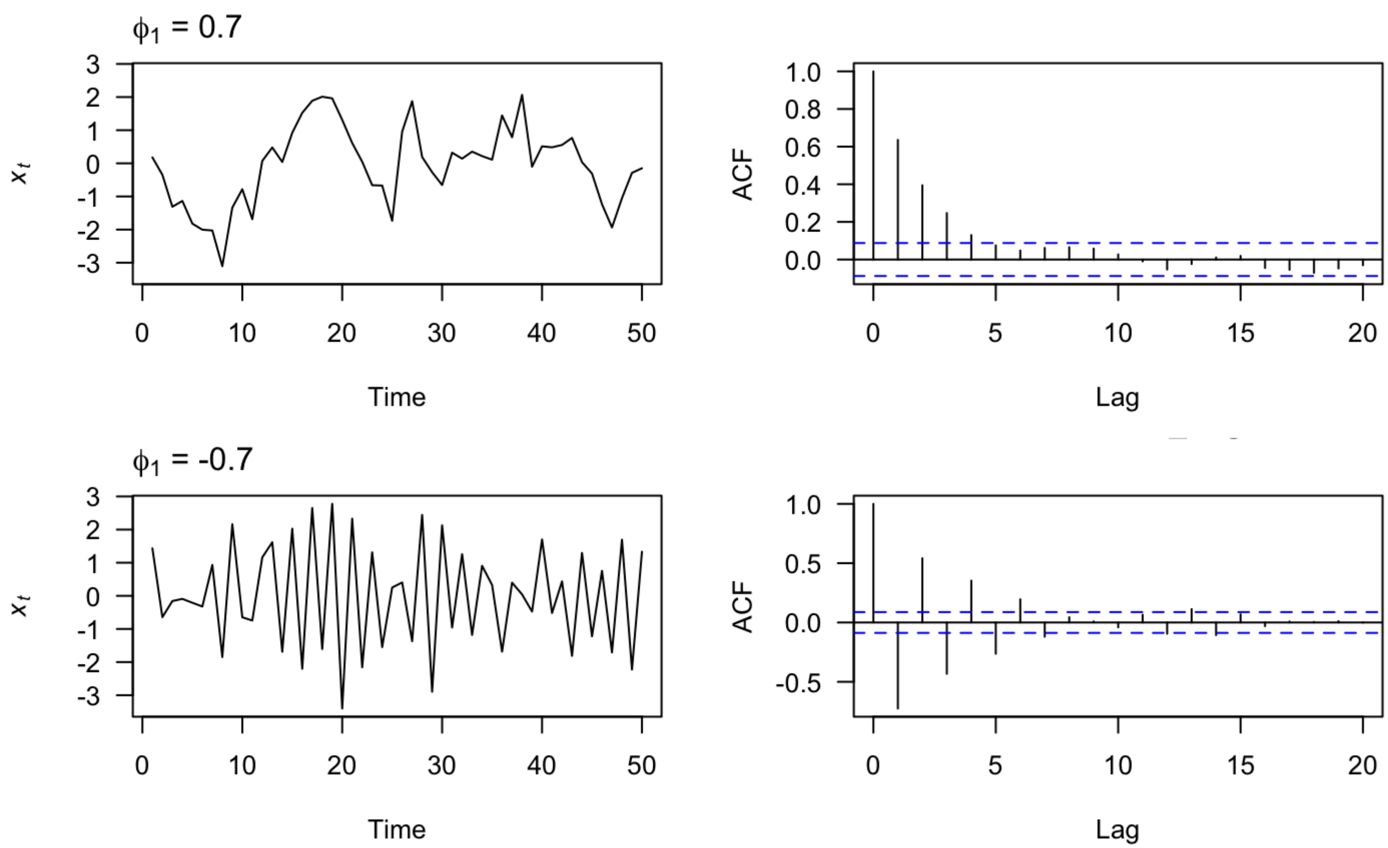


Both positive, but different magnitude

Autocorrelation function (ACF)

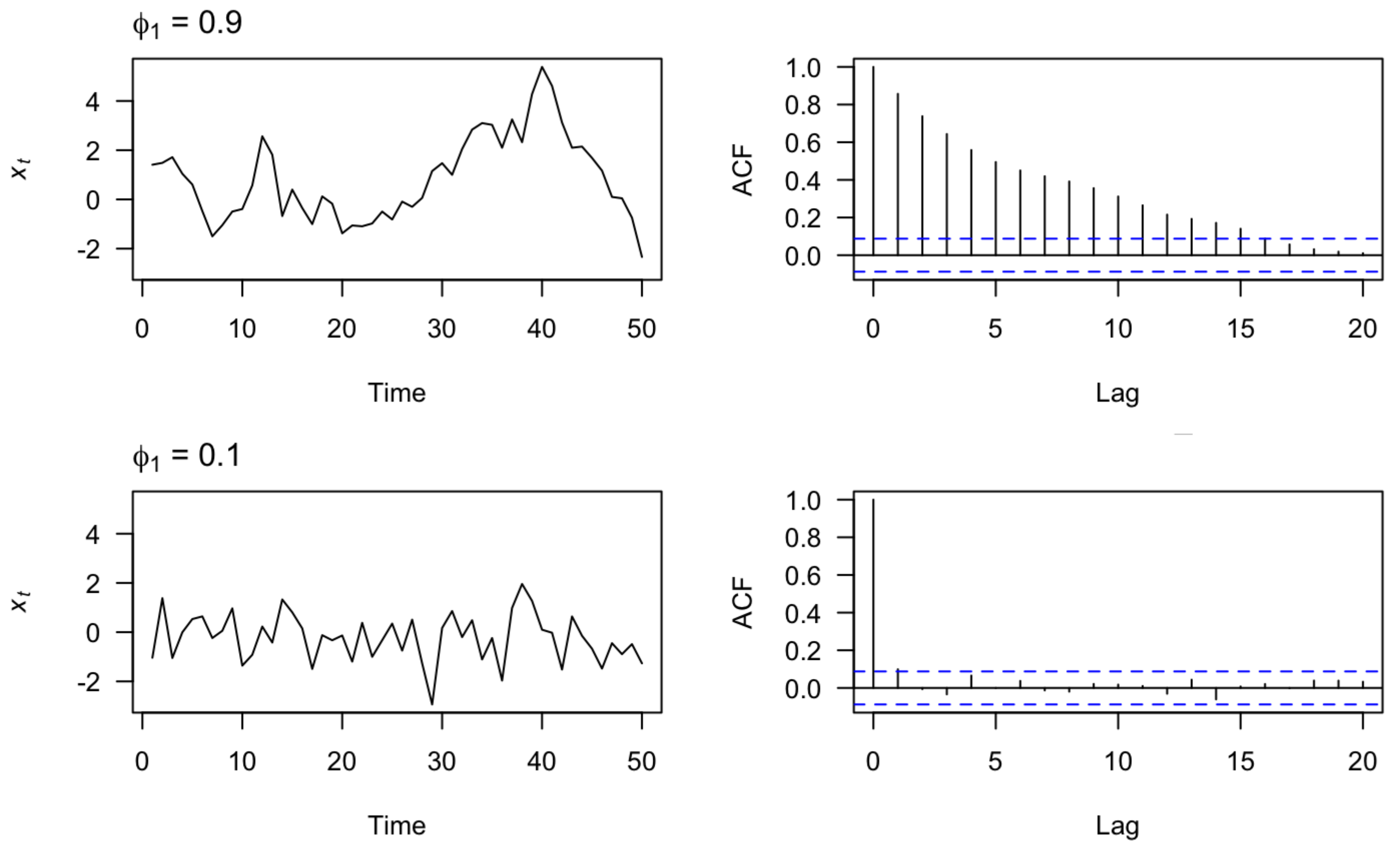
Recall that the *autocorrelation function* ($\ _k$) measures the correlation between $\{x_t\}$ and a shifted version of itself $\{x_{t+k}\}$

ACF for AR(1) models



ACF oscillates for model with –.

ACF for AR(1) models

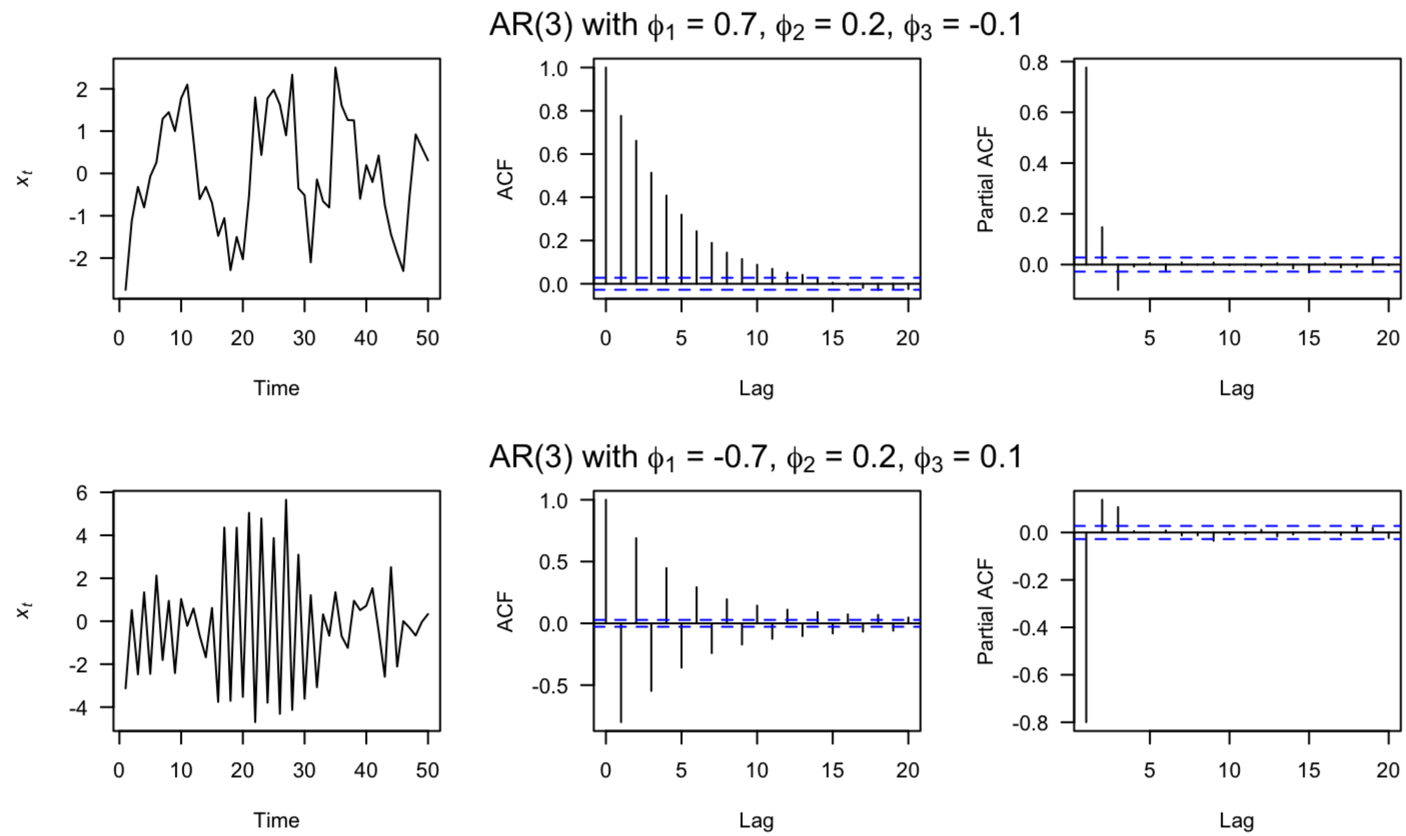


For model with large , ACF has longer tail

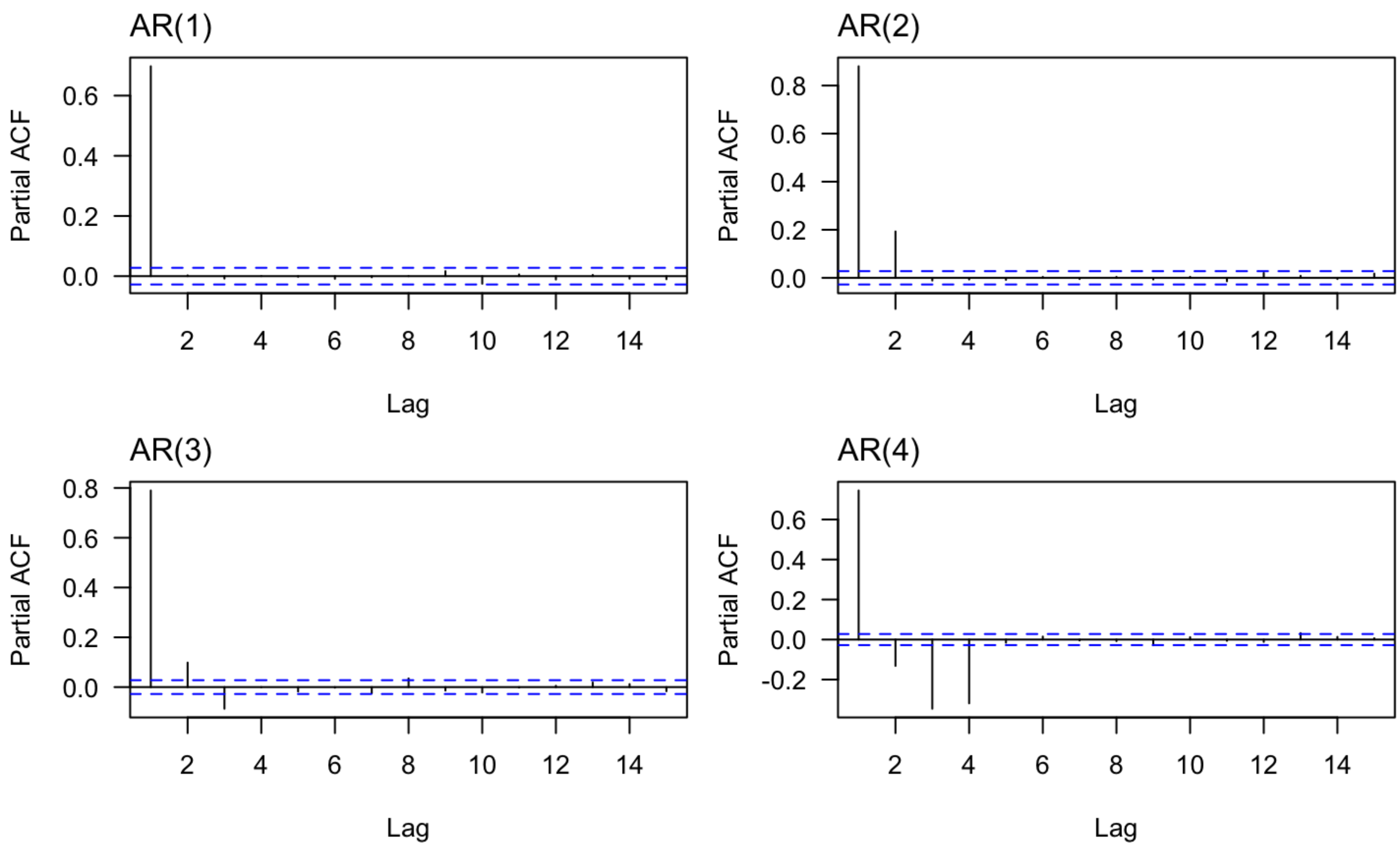
Partial autocorrelation funcion (PACF)

Recall that the *partial autocorrelation function* ($_{k}$) measures the correlation between $\{x_t\}$ and a shifted version of itself $\{x_{t+k}\}$, with the linear dependence of $\{x_{t-1}, x_{t-2}, \ldots, x_{t-k-1}\}$ removed

ACF & PACF for AR(p) models



PACF for AR(p) models



Do you see the link between the order p and lag k?

Using ACF & PACF for model ID

Model	ACF	PACF
AR(p)	Tails off slowly	Cuts off after lag p

Moving average (MA) models

A moving average model of order *q*, or MA(*q*), is defined as

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

where w_t is white noise

Each of the x_t is a sum of the most recent error terms

Moving average (MA) models

A moving average model of order *q*, or MA(*q*), is defined as

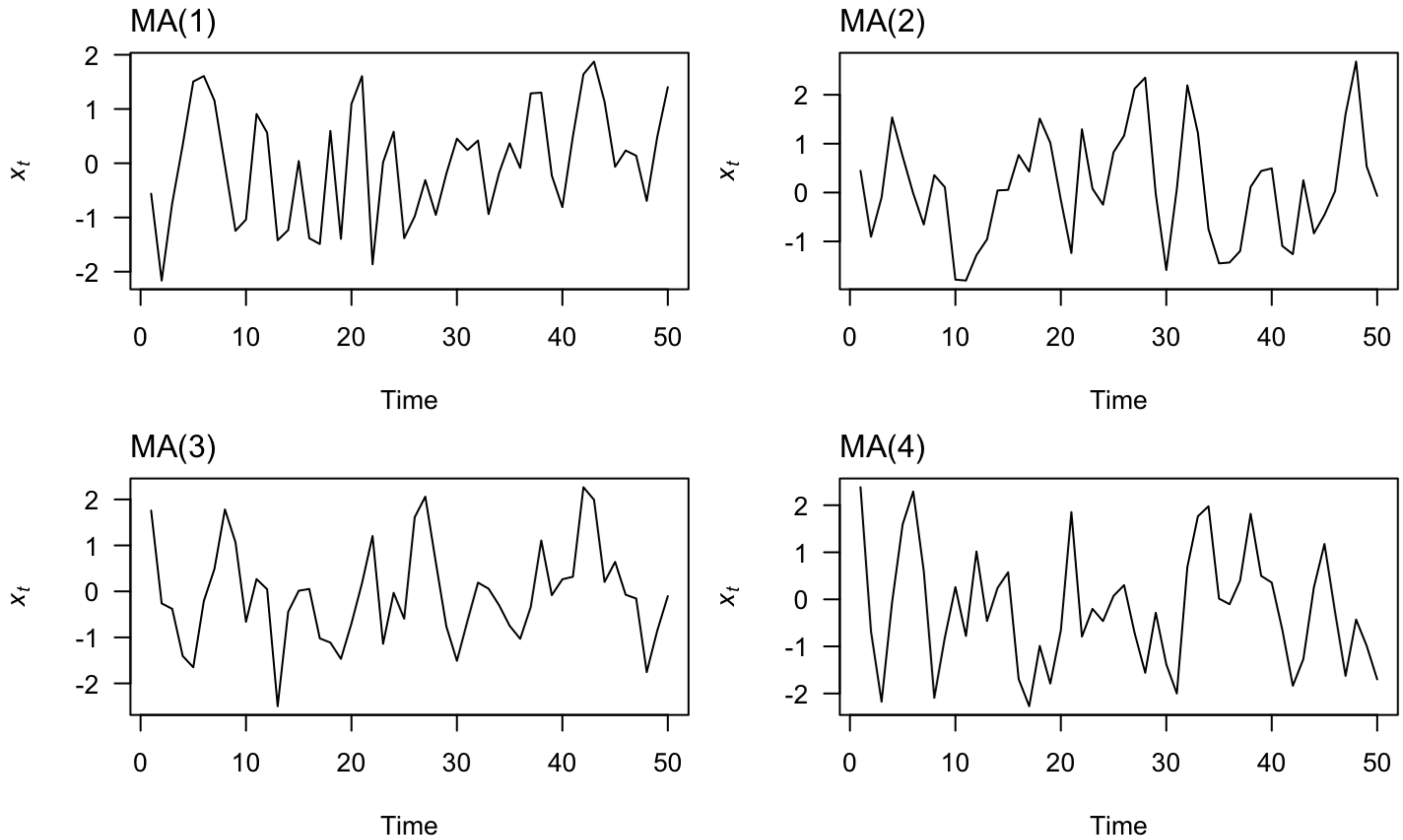
$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

where w_t is white noise

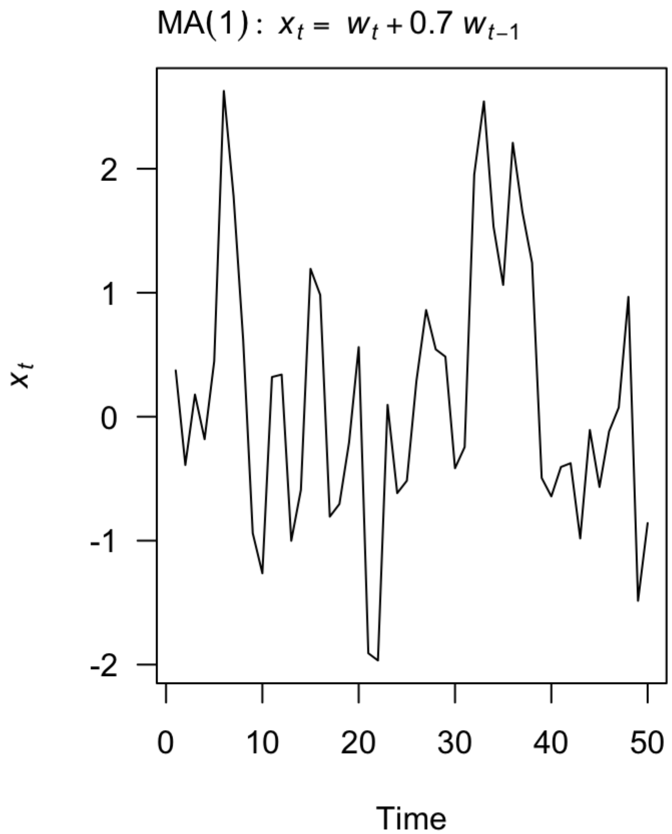
Each of the x_t is a sum of the most recent error terms

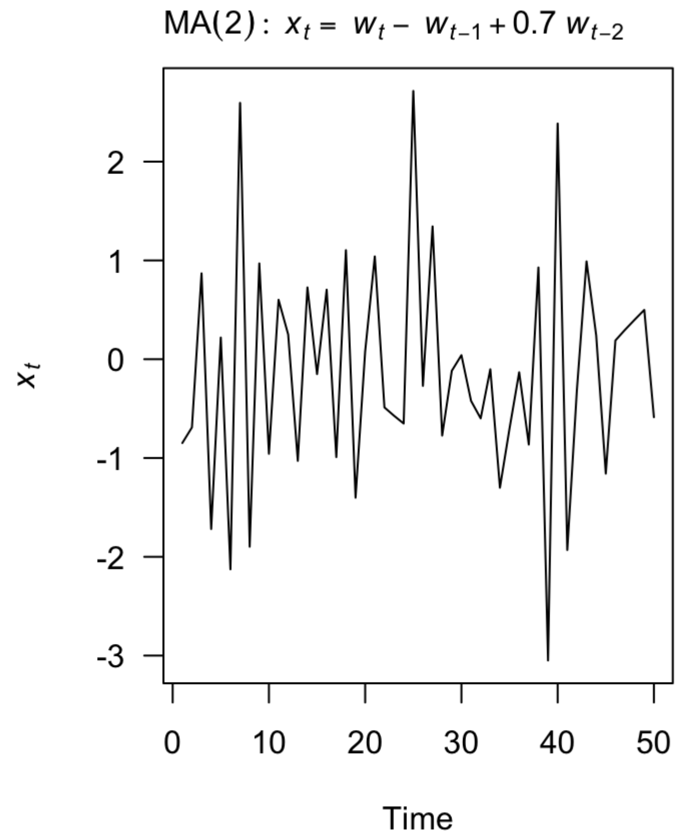
Thus, *all* MA processes are stationary because they are finite sums of stationary WN processes

Examples of MA(q) models



Examples of MA(q) models





AR(p) model as an MA(∞) model

It is possible to write an AR(p) model as an MA(∞) model

AR(1) model as an MA(∞) model

For example, consider an AR(1) model

$$
\begin{aligned}
x_t &= \phi x_{t-1} + w_t \\
x_t &= \phi(\phi x_{t-2} + w_{t-1}) + w_t \\
x_t &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\
x_t &= \phi^3 x_{t-3} + \phi^2 w_{t-2} + \phi w_{t-1} + w_t \\
&\Downarrow \\
x_t &= w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \cdots + \phi^k w_{t-k} + \phi^{k+1} x_{t-k-1}
\end{aligned}
$$

AR(1) model as an MA(∞) model

If our AR(1) model is stationary, then

$$|\phi| < 1 \;\Rightarrow\; \lim_{k \to \infty} \phi^{k+1} = 0$$

so

$$
\begin{aligned}
x_t &= w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \cdots + \phi^k w_{t-k} + \phi^{k+1} x_{t-k-1} \\
&\Downarrow \\
x_t &= w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \cdots + \phi^k w_{t-k}
\end{aligned}
$$

Invertible MA(q) models

An MA(q) process is invertible if it can be written as a *stationary autoregressive process of infinite order without an error term*

Invertible MA(1) model

For example, consider an MA(1) model

$$
\begin{aligned}
x_t &= w_t + \theta\, w_{t-1} \\
&\Downarrow \\
w_t &= x_t - \theta\, w_{t-1} \\
w_t &= x_t - \theta\, (x_{t-1} - \theta\, w_{t-2}) \\
w_t &= x_t - \theta\, x_{t-1} - \theta^{2} w_{t-2} \\
&\vdots \\
w_t &= x_t - \theta\, x_{t-1} + \cdots + (-\theta)^{k} x_{t-k} + (-\theta)^{k+1} w_{t-k-1}
\end{aligned}
$$

Invertible MA(1) model

If we constrain $|\theta| < 1$, then

$$\lim_{k\to\infty}(-\theta)^{k+1}w_{t-k-1} = 0$$

and

$$\begin{aligned}
w_t &= x_t - \theta x_{t-1} + \cdots + (-\theta)^k x_{t-k} + (-\theta)^{k+1} w_{t-k-1} \\
&\Downarrow \\
w_t &= x_t - \theta x_{t-1} + \cdots + (-\theta)^k x_{t-k} \\
w_t &= x_t + \sum_{k=1}^{\infty}(-\theta)^k x_{t-k}
\end{aligned}$$

Invertible MA(1) model

We can also show this by writing an MA(q) model with the backshift operator

$$x_t = (1 + \theta_1 \mathbf{B} + \theta_2 \mathbf{B}^2 + \cdots + \theta_q \mathbf{B}^q) w_t$$

Invertible MA(1) model

For example, consider an MA(1) model

$$
\begin{aligned}
x_t &= (1 - \theta \mathbf{B}) w_t \\
&\Downarrow \\
w_t &= \frac{1}{1 - \theta \mathbf{B}} x_t \\
&\Downarrow \text{ via Taylor Series} \\
w_t &= (1 + \theta \mathbf{B} + \theta^2 \mathbf{B}^2 + \cdots + \theta^\infty \mathbf{B}^\infty) x_t \\
w_t &= x_t + \theta x_{t-1} + \theta^2 x_{t-2} + \cdots + \theta^\infty x_{t-\infty}
\end{aligned}
$$

Invertible MA(q) models

Q: Why do we care if an MA(q) model is invertible?

A: It helps us identify the model's parameters

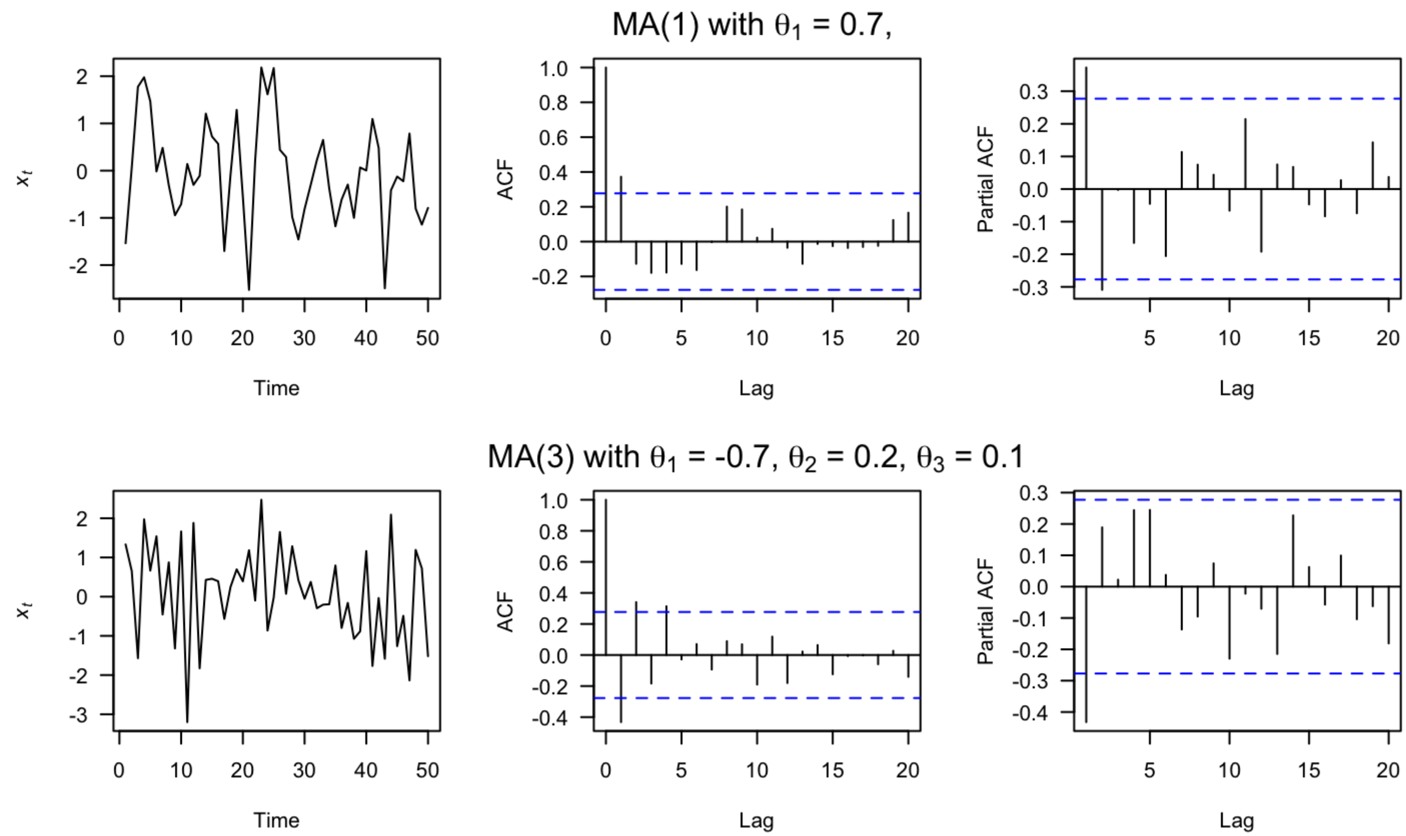
Invertible MA(q) models

For example, these MA(1) models are equivalent

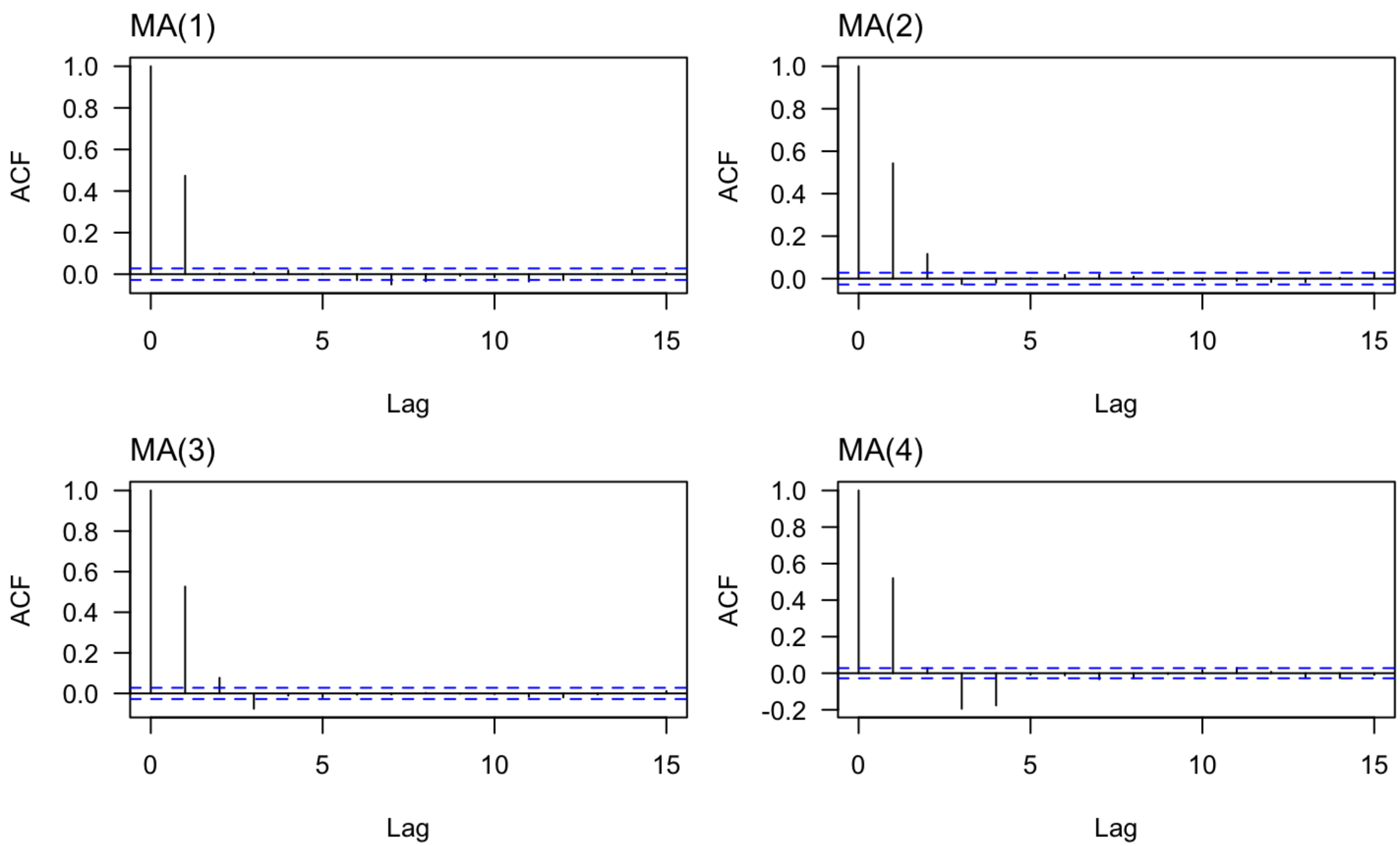
$$x_t = w_t + \frac{1}{5} w_{t-1}, \text{ with } w_t \sim \text{N}(0, 25)$$

$$x_t = w_t + 5 w_{t-1}, \text{ with } w_t \sim \text{N}(0, 1)$$

ACF & PACF for MA(q) models



ACF for MA(q) models



Do you see the link between the order q and lag k?

Using ACF & PACF for model ID

Model	ACF	PACF
AR(p)	Tails off slowly	Cuts off after lag p
MA(q)	Cuts off after lag q	Tails off slowly

Using ACF & PACF for model ID

Autoregressive moving average models

An autoregressive moving average, or ARMA(p,q), model is written as

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

Autoregressive moving average models

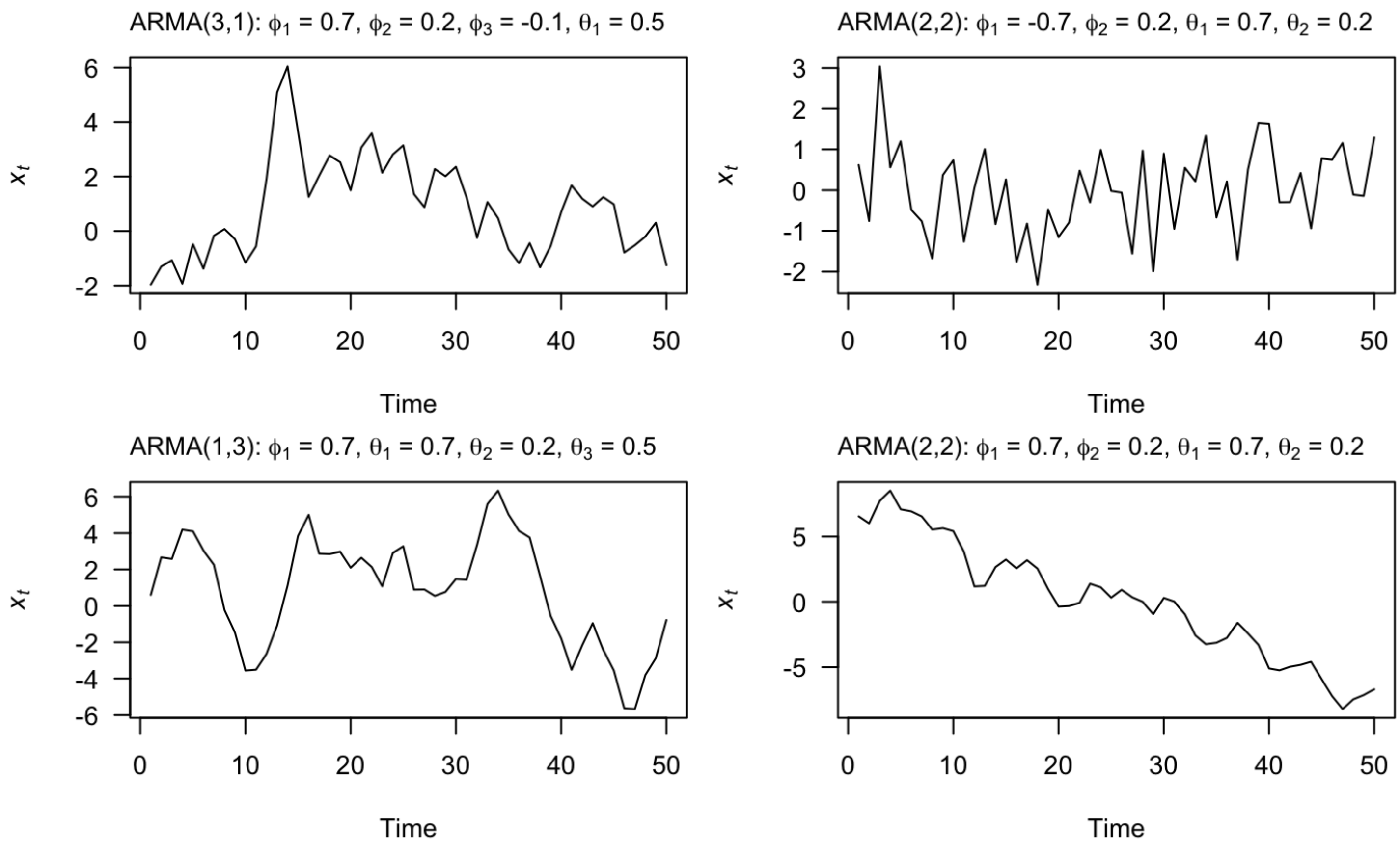
We can write an ARMA(*p*,*q*) model using the backshift operator

$$\phi_p(\mathbf{B})x_t = \theta_q(\mathbf{B})w_t$$

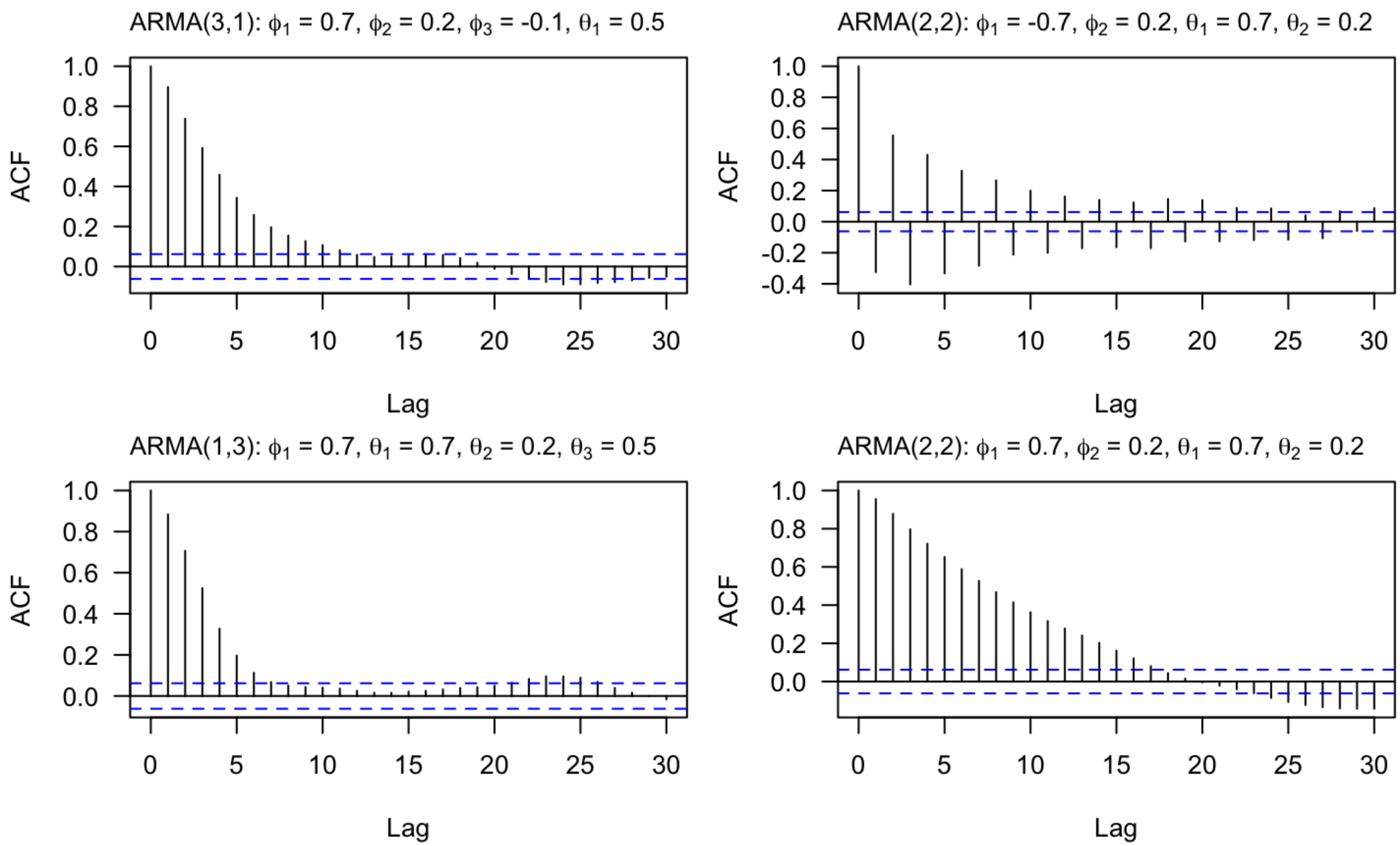
ARMA models are *stationary* if all roots of $\phi_p(\mathbf{B}) > 1$

ARMA models are *invertible* if all roots of $\theta_q(\mathbf{B}) > 1$

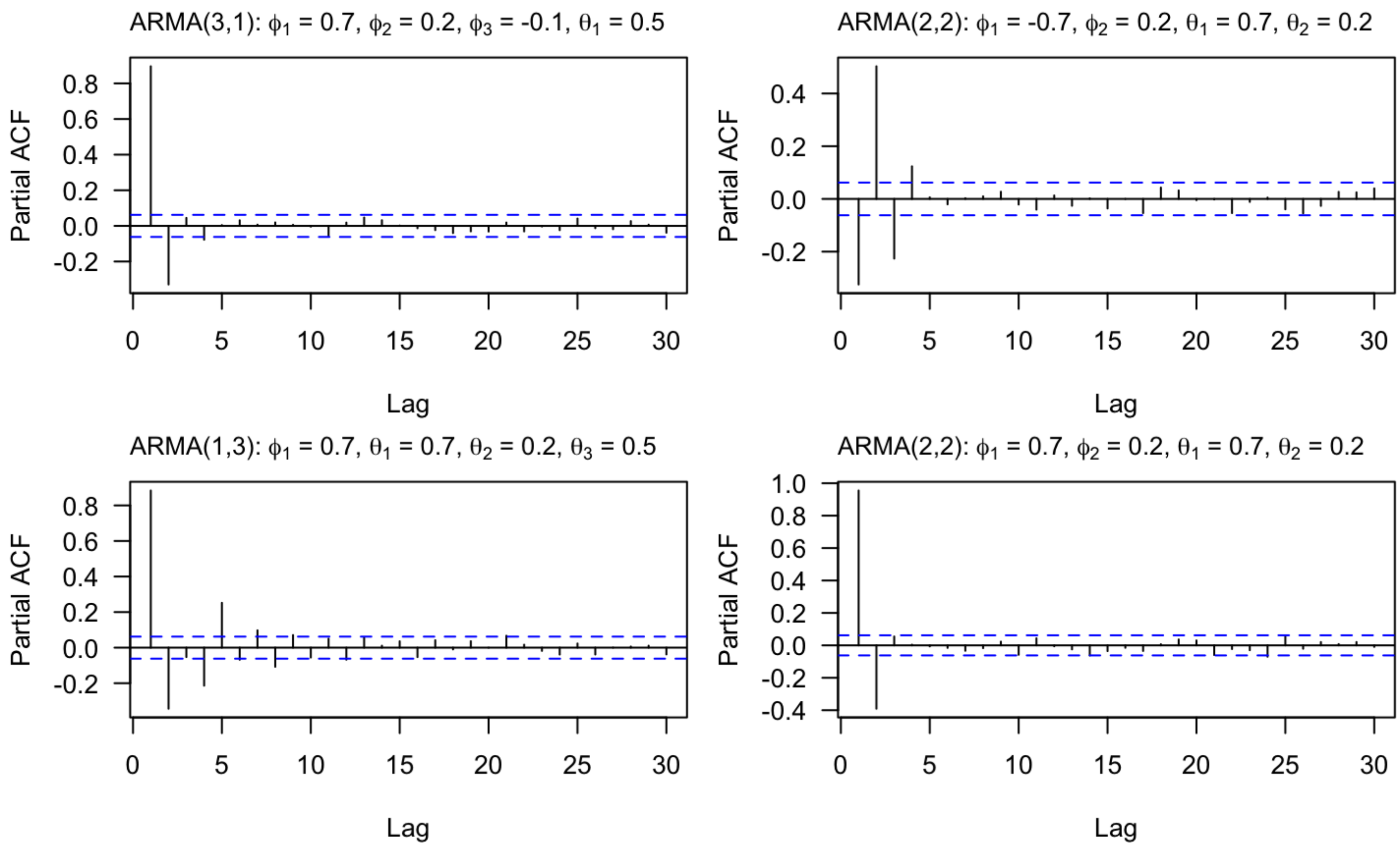
Examples of ARMA(p,q) models



ACF for ARMA(*p*,*q*) models



PACF for ARMA(p,q) models



Using ACF & PACF for model ID

Model	ACF	PACF
AR(*p*)	Tails off slowly	Cuts off after lag *p*
MA(*q*)	Cuts off after lag *q*	Tails off slowly
ARMA(*p*,*q*)	Tails off slowly	Tails off slowly

NONSTATIONARY MODELS

Autoregressive integrated moving average (ARIMA) models

If the data do not appear stationary, differencing can help

This leads to the class of *autoregressive integrated moving average* (ARIMA) models

ARIMA models are indexed with orders (p,d,q) where d indicates the order of differencing

ARIMA(*p*,*d*,*q*) models

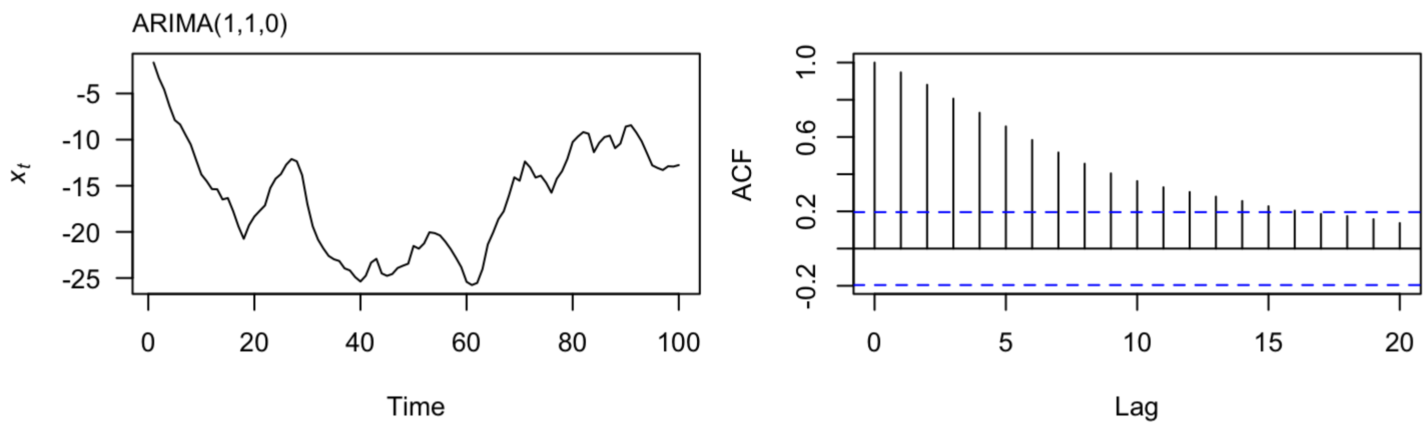
For $d > 0$, $\{x_t\}$ is an ARIMA(*p*,*d*,*q*) process if $(1 - \mathbf{B})^d x_t$ is an ARMA(*p*,*q*) process

ARIMA(*p*,*d*,*q*) models

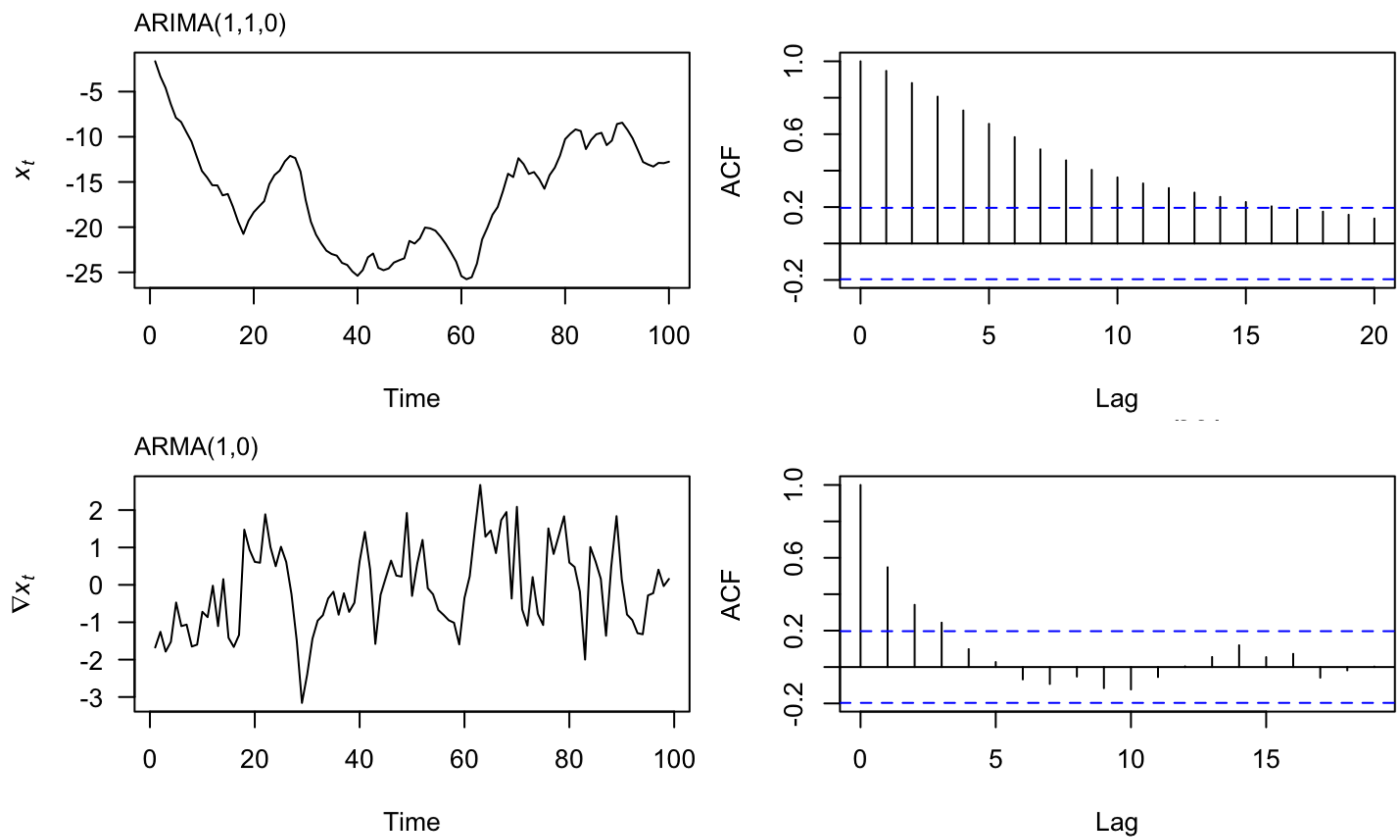
For $d > 0$, $\{x_t\}$ is an ARIMA(*p*,*d*,*q*) process if $(1 - \mathbf{B})^d x_t$ is an ARMA(*p*,*q*) process

For example, if $\{x_t\}$ is an ARIMA(1,1,0) process then $\nabla\{x_t\}$ is an ARMA(1,0) = AR(1) process

ARIMA(*p*,*d*,*q*) models



ARIMA(*p*,*d*,*q*) models



Topics for today

Review

- White noise
- Random walks

Autoregressive (AR) models

Moving average (MA) models

Autoregressive moving average (ARMA) models

Using ACF & PACF for model ID