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

\[ w_t \sim N(0; \sigma^2) \]

and the following apply as well

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

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

\[ x_t = x_{t-1} + 1 + w_t; \quad w_t \sim N(0, 1) \]
Differencing a biased random walk

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

\[
\nabla x_t = x_{t-1} + u + w_t
\]

\[
x_t - x_{t-1} = x_{t-1} + u + w_t - x_{t-1}
\]

\[
x_t - x_{t-1} = u + w_t
\]
Differencing a biased random walk

\[ x_t = x_{t-1} + 1 + w_t; w_t \sim \mathcal{N}(0, 1) \]
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 called *autoregressive moving average* (ARMA) models.
Autoregressive models are widely used in ecology to treat a current state of nature as a function of its past state(s).
Autoregressive (AR) models

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

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

where we assume

1. $w_t$ is white noise
2. $p \neq 0$ 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

\[
x_t = \cdot 1 x_{t-1} + 2 x_{t-2} + \cdots + p x_{t-p} + w_t
\]

\[
\downarrow
\]

\[
x_t - 1 x_{t-1} - 2 x_{t-2} - \cdots - p x_{t-p} = w_t
\]

\[
(1 - 1 B - 2 B^2 - \cdots - p B^p) x_t = w_t
\]

\[
. \ p (B) x_t = w_t
\]
Stationary AR\((p)\) models

If we treat \( B \) as a number (or numbers), we can write the characteristic equation as

\[
p(B)x_t = w_t
\]

\[
\downarrow
\]

\[
p(B) = 0
\]

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

\[ x_t = 0.5x_{t-1} + w_t \]
\[ x_t - 0.5x_{t-1} = w_t \]
\[ (1 - 0.5B)x_t = w_t \]
Stationary AR($p$) models

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

\[ x_t = 0.5x_{t-1} + w_t \]
\[ x_t - 0.5x_{t-1} = w_t \]
\[ (1 - 0.5B)x_t = w_t \]
\[ \downarrow \]
\[ 1 - 0.5B = 0 \]
\[ -0.5B = -1 \]
\[ B = 2 \]

This model is indeed stationary because $B > 1$
Stationary AR($p$) models

What about this AR(2) model from earlier?

\[ x_t = -0.2x_{t-1} + 0.4x_{t-2} + w_t \]
\[ x_t + 0.2x_{t-1} - 0.4x_{t-2} = w_t \]
\[ (1 + 0.2B - 0.4B^2)x_t = w_t \]
Stationary AR(\(p\)) models

What about this AR(2) model from earlier?

\[ x_t = -0.2x_{t-1} + 0.4x_{t-2} + \epsilon_t \]

\[ x_t + 0.2x_{t-1} - 0.4x_{t-2} = \epsilon_t \]

\[ (1 + 0.2B - 0.4B^2)x_t = \epsilon_t \]

\[ \downarrow \]

\[ 1 + 0.2B - 0.4B^2 = 0 \]

\[ \downarrow \]

\[ B \approx -1.35 \text{ and } B \approx 1.85 \]

This model is *not* stationary because only one \(B > 1\)
What about random walks?

Consider our random walk model

\[ x_t = x_{t-1} + w_t \]

\[ x_t - x_{t-1} = w_t \]

\[ (1 - 1B)x_t = w_t \]
What about random walks?

Consider our random walk model

\[ x_t = x_{t-1} + w_t \]
\[ x_t - x_{t-1} = w_t \]
\[ (1 - 1B)x_t = w_t \]
\[ \downarrow \]
\[ 1 - 1B = 0 \]
\[ -1B = -1 \]
\[ B = 1 \]

Random walks are not stationary because \( B = 1 \neq 1 \)
Stationary AR(1) models

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

For \( x_t = x_{t-1} + w_t \), we have

\[
\begin{align*}
1 - \beta &= 0 \\
\therefore \beta &= -1 \\
\frac{1}{1 - \beta} &= \frac{1}{-1} \Rightarrow 0 < \beta < 1
\end{align*}
\]
Stationary AR(1) models

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

$$1 - B = 0$$
$$1 - B = -1$$
$$B = \frac{1}{-1} > 1 \Rightarrow 0 < \cdot < 1$$

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

$$1 + B = 0$$
$$1 - B = -1$$
$$B = \frac{-1}{-1} > 1 \Rightarrow -1 < \cdot < 0$$
Stationary AR(1) models

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

Same value, but different sign
Coefficients of AR(1) models

Both positive, but different magnitude

\[ \phi_1 = 0.9 \]

\[ \phi_1 = 0.1 \]
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 $\phi_1 = -0.7$. 

$\phi_1 = 0.7$
ACF for AR(1) models

For model with large $\phi_1$, ACF has longer tail
Partial autocorrelation funcion (PACF)

Recall that the *partial autocorrelation function* \( (\phi_k) \) 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($\rho$) models

AR(3) with $\phi_1 = 0.7$, $\phi_2 = 0.2$, $\phi_3 = -0.1$

AR(3) with $\phi_1 = -0.7$, $\phi_2 = 0.2$, $\phi_3 = 0.1$
PACF for AR($p$) models

Do you see the link between the order $p$ and lag $k$?
Using ACF & PACF for model ID

<table>
<thead>
<tr>
<th>Model</th>
<th>ACF</th>
<th>PACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR($p$)</td>
<td>Tails off slowly</td>
<td>Cuts off after lag $p$</td>
</tr>
</tbody>
</table>
Moving average (MA) models

Moving average models are most commonly used for forecasting a future state.
Moving average (MA) models

A moving average model of order $q$, or $\text{MA}(q)$, is defined as

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

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

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

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

- MA(1)
- MA(2)
- MA(3)
- MA(4)
Examples of MA($q$) models

MA(1): $x_t = w_t + 0.7 \, w_{t-1}$

MA(2): $x_t = w_t - w_{t-1} + 0.7 \, w_{t-2}$
AR($p$) model as an MA($\infty$) model

It is possible to write an AR($p$) model as an MA($\infty$) model
AR(1) model as an MA(∞) model

For example, consider an AR(1) model

\[ x_t = x_{t-1} + w_t \]
\[ x_t = (x_{t-2} + w_{t-1}) + w_t \]
\[ x_t = 2x_{t-2} + w_{t-1} + w_t \]
\[ x_t = 3x_{t-3} + 2w_{t-2} + w_{t-1} + w_t \]
\[ \downarrow \]
\[ x_t = w_t + w_{t-1} + 2w_{t-2} + \cdots + k_w t-k + 1 x_{t-k-1} \]
AR(1) model as an MA(∞) model

If our AR(1) model is stationary, then

\[ | \phi | < 1 \Rightarrow \lim_{k \to \infty} w_{t+k+1} = 0 \]

so

\[
\begin{align*}
x_t &= w_t + w_{t-1} + 2w_{t-2} + \cdots + k w_{t-k} + x_{t-k-1} \\
\downarrow & \quad \downarrow \\
x_t &= w_t + w_{t-1} + 2w_{t-2} + \cdots + k w_{t-k}
\end{align*}
\]
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

\[ x_t = w_t + w_{t-1} \]

\[ \downarrow \]

\[ w_t = x_t - w_{t-1} \]

\[ w_t = x_t - (x_{t-1} - w_{t-2}) \]

\[ w_t = x_t - x_{t-1} - 2w_{t-2} \]

\[ \vdots \]

\[ w_t = x_t - x_{t-1} + \cdots + (-2)^k x_{t-k} + (-2)^{k+1} w_{t-k-1} \]
Invertible MA(1) model

If we constrain \(|\phi| < 1\), then

\[
\lim_{k \to \infty} (\phi^2)^{k+1} w_{t-k-1} = 0
\]

and

\[
\begin{align*}
w_t &= x_t - \phi^2 x_{t-1} + \cdots + (\phi^2)^k x_{t-k} + (\phi^2)^{k+1} w_{t-k-1} \\
&\downarrow \\
w_t &= x_t - \phi^2 x_{t-1} + \cdots + (\phi^2)^k x_{t-k} \\
w_t &= x_t + \sum_{k=1}^{\infty} (\phi^2)^k x_{t-k}
\end{align*}
\]
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

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

\[ x_t = w_t + 5w_{t-1}, \text{ with } w_t \sim N(0, 1) \]
ACF & PACF for MA(\(q\)) models

MA(1) with \(\theta_1 = 0.7\),

MA(3) with \(\theta_1 = -0.7, \theta_2 = 0.2, \theta_3 = 0.1\)
ACF for MA($q$) models

Do you see the link between the order $q$ and lag $k$?
Using ACF & PACF for model ID

<table>
<thead>
<tr>
<th>Model</th>
<th>ACF</th>
<th>PACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR($p$)</td>
<td>Tails off slowly</td>
<td>Cuts off after lag $p$</td>
</tr>
<tr>
<td>MA($q$)</td>
<td>Cuts off after lag $q$</td>
<td>Tails off slowly</td>
</tr>
</tbody>
</table>
Using ACF & PACF for model ID
Autoregressive moving average models

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

$$x_t = \sum_{i=1}^{p} a_i x_{t-i} + \sum_{j=1}^{q} \gamma_j w_{t-j}$$
Autoregressive moving average models

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

\[ p(B)x_t = q(B)w_t \]

ARMA models are *stationary* if all roots of $p(B) > 1$

ARMA models are *invertible* if all roots of $q(B) > 1$
Examples of ARMA($p, q$) models

ARMA(3,1): $\phi_1 = 0.7$, $\phi_2 = 0.2$, $\phi_3 = -0.1$, $\theta_1 = 0.5$

ARMA(2,2): $\phi_1 = -0.7$, $\phi_2 = 0.2$, $\theta_1 = 0.7$, $\theta_2 = 0.2$

ARMA(1,3): $\phi_1 = 0.7$, $\theta_1 = 0.7$, $\theta_2 = 0.2$, $\theta_3 = 0.5$

ARMA(2,2): $\phi_1 = 0.7$, $\phi_2 = 0.2$, $\theta_1 = 0.7$, $\theta_2 = 0.2$
ACF for ARMA($p,q$) models

ARMA(3,1): $\phi_1 = 0.7$, $\phi_2 = 0.2$, $\phi_3 = -0.1$, $\theta_1 = 0.5$

ARMA(2,2): $\phi_1 = -0.7$, $\phi_2 = 0.2$, $\theta_1 = 0.7$, $\theta_2 = 0.2$

ARMA(1,3): $\phi_1 = 0.7$, $\theta_1 = 0.7$, $\theta_2 = 0.2$, $\theta_3 = 0.5$

ARMA(2,2): $\phi_1 = 0.7$, $\phi_2 = 0.2$, $\theta_1 = 0.7$, $\theta_2 = 0.2$
PACF for ARMA\((p,q)\) models

ARMA(3,1): \(\phi_1 = 0.7, \phi_2 = 0.2, \phi_3 = -0.1, \theta_1 = 0.5\)

ARMA(2,2): \(\phi_1 = -0.7, \phi_2 = 0.2, \theta_1 = 0.7, \theta_2 = 0.2\)

ARMA(1,3): \(\phi_1 = 0.7, \theta_1 = 0.7, \theta_2 = 0.2, \theta_3 = 0.5\)

ARMA(2,2): \(\phi_1 = 0.7, \phi_2 = 0.2, \theta_1 = 0.7, \theta_2 = 0.2\)
# Using ACF & PACF for model ID

<table>
<thead>
<tr>
<th>Model</th>
<th>ACF</th>
<th>PACF</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR($p$)</td>
<td>Tails off slowly</td>
<td>Cuts off after lag $p$</td>
</tr>
<tr>
<td>MA($q$)</td>
<td>Cuts off after lag $q$</td>
<td>Tails off slowly</td>
</tr>
<tr>
<td>ARMA($p,q$)</td>
<td>Tails off slowly</td>
<td>Tails off slowly</td>
</tr>
</tbody>
</table>
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 $(1 - B)^d x_t$ is an ARMA($p,q$) process
ARIMA($p,d,q$) models

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

ARIMA(1,1,0)

$X_t$

ACF

ARMA(1,0)

$\nabla X_t$

ACF
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