6. MA & ARMA

author: ECONOMIC FORECASTING date: SUMMER 2017 autosize: true

Definition

Moving Average

A moving average process of order q, MA(q), is a linear combination of the most recent white noise term and the q most recent white noise terms, and is denoted as:

$$x_t = w_t + \beta_1 w_{t-1} + \cdots + \beta_q w_{t-q}$$

where $W \sim WN(0, \sigma_w^2)$

Using the backshift operator:

$$x_t = (1 + \beta_1 \mathbf{B} + \cdots + \beta_q \mathbf{B}^q) w_t = \phi_q(\mathbf{B}) w_t$$

Properties

Expected Value

$$\begin{aligned}
\mathbf{E}(x_t) &= \mathbf{E}(w_t + \beta_1 w_{t-1} + \cdots + \beta_q w_{t-q}) \\
&= \mathbf{E}(w_t) + \beta_1 \mathbf{E}(w_{t-1}) + \cdots + \beta_q \mathbf{E}(w_{t-q}) \\
&= 0
\end{aligned}$$

Variance

$$\begin{aligned}
\mathbf{V}(x_t) &= \mathbf{V}(w_t + \beta_1 w_{t-1} + \cdots + \beta_q w_{t-q}) \\
&= \mathbf{V}(w_t) + \beta_1^2 \mathbf{V}(w_{t-1}) + \cdots + \beta_q^2 \mathbf{V}(w_{t-q}) \\
&= \sigma_w^2 (1 + \beta_1^2 + \cdots + \beta_q^2)
\end{aligned}$$

The Correlation Coefficient

$$\rho(k) = \begin{cases} 1 & \text{for } k = 0 \\ \displaystyle\sum_{i=0}^{q-k} \beta_i \beta_{i+k} \Big/ \sum_{i=0}^{q} \beta_i^2 & \text{for } k = 1, ..., q \\ 0 & \text{for } k > q \end{cases}$$

where $\beta_0 = 1$

Proving this relies on the fact that

$$\mathbf{C}(\sum x_t, \sum y_t) = \sum \sum \mathbf{c}(x_t, y_t)$$

but, instead of proving this, we will shor how it is derived for MA(1) process.

Autocovariance for MA(1)

$$\begin{aligned}
\gamma(x_t, x_{t+k}) &= \mathbf{E}[(x_t - \mu_{x_t})(x_{t+k} - \mu_{x_{t+k}})] \\
&= \mathbf{E}[(x_t)(x_{t+k})] \\
&= \mathbf{E}[(w_t + \beta w_{t-1})(w_{t-1} + \beta w_{t-2})] \\
&= \mathbf{E}[(w_t w_{t-1}) + (\beta w_t w_{t-2}) + (\beta w_{t-1} w_{t-1}) + (\beta^2 w_{t-1} w_{t-2})] \\
&= \mathbf{E}(wtw_{t-1}) + \beta \mathbf{E}(w_t w_{t-2}) + \beta \mathbf{E}(w_{t-1}^2) + \beta^2 \mathbf{E}(w_{t-1} w_{t-2}) \\
&= 0 + 0 + \beta^2 \sigma_w^2 + 0 \\
&= \beta^2 \sigma_w^2
\end{aligned}$$

Correlation Coefficient for MA(1)

$$\begin{aligned}
\sqrt{\mathbf{V}(x_t)\mathbf{V}(x_{t+k})} &= \sqrt{\mathbf{V}(x_t)^2} \\
&= \mathbf{V}(x_t) \\
&= \sigma_w^2(1 + \beta^2)
\end{aligned}$$

Putting it all together…

$$\begin{aligned}
\rho(MA(1)) &= \frac{\beta \sigma_w^2}{\sigma_w^2(1 + \beta^2)} \\
&= \frac{\beta}{(1 + \beta^2)}
\end{aligned}$$

Theoretical ACF & PACF

MA(1) process with $\beta > 0$ - ACF: Positive spike at lag 1. $\rho_k = 0$ for $k \geq 2$ - PACF: Oscillating decay with $\phi_{11} > 0$

MA(1) process with $\beta < 0$ - ACF: Negative spike at lag 1. $\rho_k = 0$ for $k \geq 2$ - PACF: Geometric decay with $\phi_{11} > 0$

Graphically

Graphically

Graphically

Invertibility

Unlike an AR model, the MA process is not directly observable since the w_t's are not observable - To estimate the MA(q) process, we need to assume that the process is **invertible**.

Invertibility

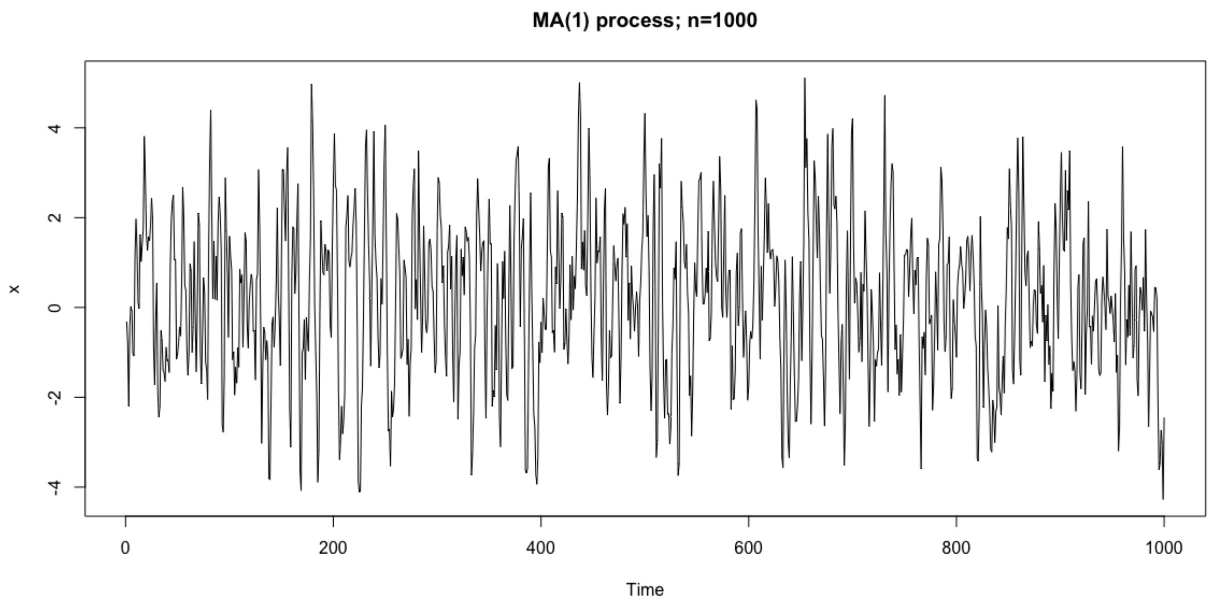

Figure 1:

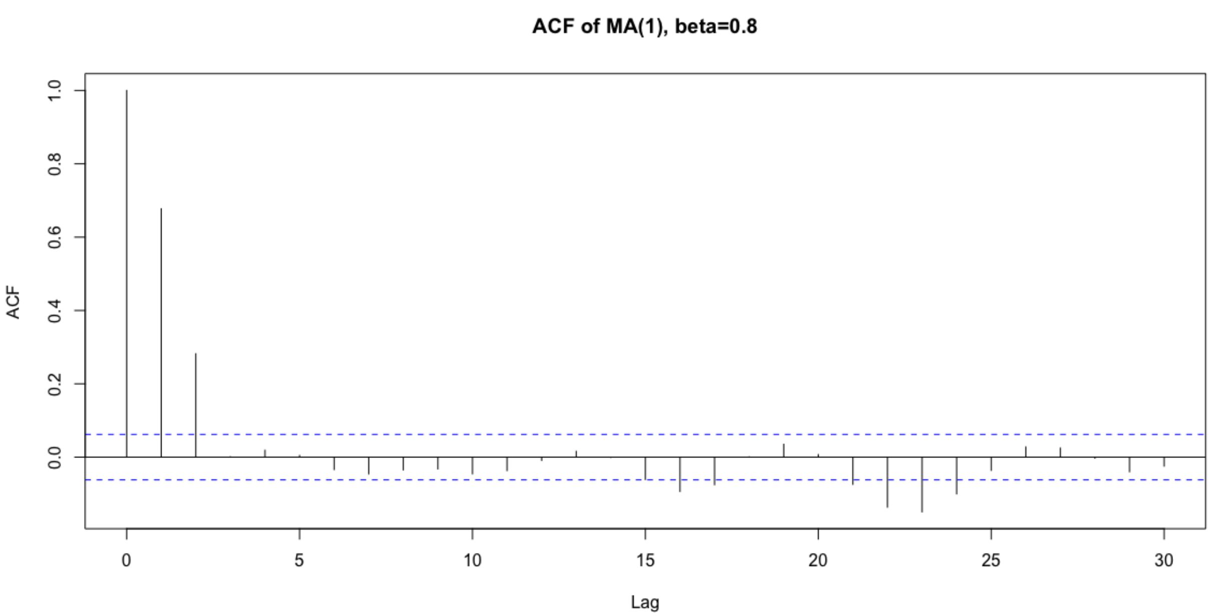

Figure 2:

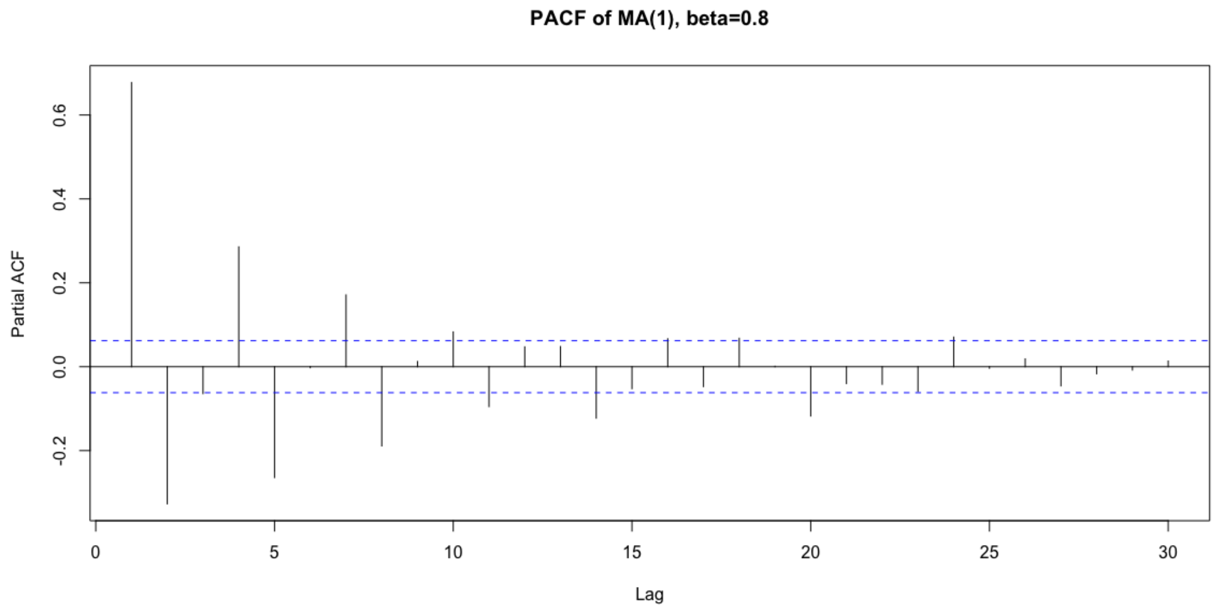

Figure 3:

An error term process is invertible if it can be written as an infinite autoregressive process without an error term:

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

assuming that $|\beta| < 1 \forall \beta$. That is, the roots of $\phi_q(\mathbf{B})$ all exceed unity in absolute value.

Continued...

Rewriting in backshift notation

$$w_t = (1 - \beta \mathbf{B})^{-1} x_t$$

This process will be invertible if $|\beta| < 1$ for any β. - If there exists a $|\beta| > 1$ then $\{w_t\}$ is not writable as an infinite autoregressive process

Assuming that the process is invertible will allow us to estimate and find a unique solution to a maximum likelihood estimation procedure - Again, this is due to the w_t's being unobservable.

Forecasting MA(q) with R

Use the `arima()` function

Example:

Suppose you have an MA(3) process. - `arima(<data object>, order=c(0,0,3))`

Of course, you typically won't know what order an MA process is... - Use the ACF and PACF functions to make an intitial guess. - Use the AIC and AIC_c to compare models

Use the `predict(<arima object>)` function to make predictions.

ARMA

Putting an AR(p) model together with an MA(q) model

Autoregressive Moving Average

A time series $\{x_t\}$ is an autoregressive moving average series (ARMA) of order (p, q) if the series can be expressed by

$$\begin{aligned} x_t &= \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p} \\ &\quad + w_t + \beta_1 w_{t-1} + \beta_2 w_{t-2} + \cdots + \beta_p w_{t-p} \end{aligned}$$

where $w_t \sim WN$

In backshift notation

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

Properties of ARMA(1,1)

$$\begin{aligned} x_t &= \alpha x_{t-1} + w_t + \beta w_{t-1} \\ (1 - \alpha \mathbf{B}) x_t &= (1 + \beta \mathbf{B}) w_t \\ x_t &= (1 - \alpha \mathbf{B})^{-1} (1 + \beta \mathbf{B}) w_t \\ &= 1 + \alpha \mathbf{B} + (\alpha \mathbf{B})^2 = \cdots)(1 + \beta \mathbf{B}) w_t \\ &= \left(\sum_{i=0}^{\infty} (\alpha \mathbf{B})^i \right) (1 + \beta \mathbf{B}) w_t \\ &= \left(1 + \sum_{i=0}^{\infty} \alpha^{i+1} \mathbf{B}^{i+1} + \sum_{i=0}^{\infty} \alpha^i \beta \mathbf{B}^{i+1} \right) w_t \\ &= w_t + (\alpha + \beta) \sum_{i=0}^{\infty} \alpha^{i-1} w_{t-1} \end{aligned}$$

Properties Continued...

Expected Value

$$\mathbf{E} \left[w_t + (\alpha + \beta) \sum_{i=0}^{\infty} \alpha^{i-1} w_{t-1} \right] = 0$$

and

$$\rho_k = \lambda_1 \rho_{k-1} + \lambda_2 \rho_{k-2} + \cdots + \lambda_p \rho_{k-p}$$

for the AMRA(p, q) process

ARMA(1,1) process with $\lambda_1 > 0$ - ACF: Geometric decay beginning after lag 1. The sign $(\rho_1) =$ sign $(\lambda_1 + \beta)$ - PACF: Oscillating after lag 1. $\phi_{11} = \rho_1$

Properties Continued...

ARMA(1,1) process with $\lambda_1 > 0$ - ACF: Oscillating decay beginning after lag 1. The sign $(\rho_1) =$ sign $(\lambda_1 + \beta)$ - PACF: Geometric decay beginning after lag 1. $\phi_{11} = \rho_1$ and sign $(\phi_{kk}) =$ sign (ϕ_{11})

ARMA(p,q) - ACF: Decay (either oscillating or geometrically) beginning after lag q. - PACF: Decay (either oscillating or geometrically) beginning after lag p.

Forecasting the ARMA(p,q) Model

Same strategy as the MA(q) model... `- arima(<data object>, order=c(p,0,q))` - Use fit statistics to choose order of the model - Parsimony is desired. High order AR processes can be modeled with a lower order ARMA model - Use the `predict()` function to forecast

Notes

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

- ARMA(p,0) is equivalent to AR(p)
- ARMA(0,q) is equivalent to MA(q)
- Time series data is stationary when all roots of θ are greater than 1
- TIme series data is invertible when all roots of θ are greater than 1

Parameter Redundancy

ARMA(2,1)

$$\begin{aligned}
x_t &= \frac{5}{6}x_{t-1} - \frac{1}{6}x_{t-2} + w_t - \frac{1}{2}x_{t-1} \\
x_t\left(1 - \frac{5}{6}\mathbf{B} + \frac{1}{6}\mathbf{B}^2\right) &= \left(1 - \frac{1}{2}\mathbf{B}\right)w_t \\
x_t\left(1 - \frac{1}{2}\mathbf{B}\right)\left(1 - \frac{1}{3}\mathbf{B}\right) &= \left(1 - \frac{1}{2}\mathbf{B}\right)w_t \\
x_t\left(1 - \frac{1}{3}\mathbf{B}\right) &= w_t \\
x_t &= \frac{1}{3}x_{t-1} + w_t
\end{aligned}$$

which is an ARMA(1,0) $\equiv$ AR(1)

===

Much like confidence intervals for the estimation of parameters, we would like an interval for forecasted values.

One period ahead for an AR(1) process:

$$[\hat{y}_{t+1} \pm z^* s]$$

where s is the standard error and z^* is the desired confidence level.

Two periods ahead for an AR(1) process:

$$\left[\hat{y}_{t+2} \pm z^* s\sqrt{1 + (\hat{\alpha})^2}\right]$$

where $\hat{\alpha}$ is the estimation from $x_t = \hat{\alpha} x_{t-1} + w_t$

For the τ period ahead for an AR(1) process:

$$\left[\hat{y}_{t+2} \pm z^* s\sqrt{1 + (\hat{\alpha})^2 + \cdots + (\hat{\alpha})^{2(\tau-1)}}\right]$$

==

As you can guess, prediction intervals grow more complicated as more terms are added - little consensus on prediction intervals - software package "black box" - R does not have a calculation method in its default packages - could try downloading the "forecast" package

Overall, all of the moethods discussed in this course are good for short-run (i.e. one period ahead forecasting)

Long-run Forecasting

- Estimate seasonal data
- Split data into "training" and "testing" portions
  - check long-run usefulness
  - estimate and compare parameters for stability
- Bootstrapping prediction intervals