MSA 8200 Predictive Analytics
Week 3: ARMA model

Yichen Cheng

Spring 2020

Outline

- Moving average (MA) model.
- PACF
- Estimation
- ARMA

Moving Average Model

MA(q) model:

$$x_t = w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q}$$

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

Equivalently, a MA(q) model can be written as:

- $x_t = \theta(B) w_t$

The moving average operator is

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

Moving Average Model

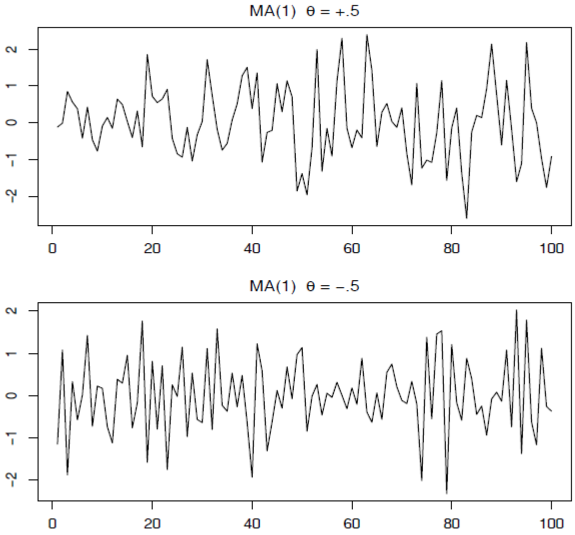


Fig. 3.2. Simulated MA(1) models: $\theta = .5$ (top); $\theta = -.5$ (bottom).

ACF of a MA(1) model:

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

- $\gamma(h) = cov(x_{t+h}, x_t) = cov(w_{t+h} + \theta w_{t+h-1}, w_t + \theta w_{t-1})$.

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

- ACF

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

ACF of MA(1)

Discussion:

- $|\rho(1)| \leq .5$.
- x_t correlated with x_{t-1} but not correlated with x_{t-2}, x_{t-3}, ...
- $\rho(h)$ is the same for θ and $1/\theta$. Which one to choose?

Choose the model with an infinite AR representation. – Invertible process.

ACF of MA(q)

ACF of a MA(q) model:

$$x_t = w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q}$$

- $\gamma(h) = cov(x_{t+h}, x_t) = cov(\sum_{j=0}^{q} \theta_j w_{t+h-j}, \sum_{k=0}^{q} \theta_k w_{t-k})$.

$$\gamma(h) = \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h} & 0 \le h \le q. \\ 0 & h > q. \end{cases}$$

- ACF

$$\rho(h) = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{1+\theta_1^2+\ldots+\theta_q^2} & 0 \le h \le q. \\ 0 & h > q. \end{cases}$$

Partial Correlation

Partial correlation of x and y given z:

- $\rho_{xy|z} = cor(x - \hat{x}, y - \hat{y})$.

measures the strength of relationship between two varibles, while controlling for the effect of other variables.

Example:

- correlation between amount of food eaten and blood pressure, while controlling for weight, BMI and amout of exercise.
- correlation between log(wage) and IQ, while controlling for *Edu*, *Ability*, *Expr*, $Expr^2$

Partial Correlation

- $\rho_{xy|z} = cov(x - \hat{x}, y - \hat{y})$.
- $\hat{x} = P_z(x)$, $\hat{y} = P_z(y)$.

Example:

- $x = z + x_1$, $z \sim \mathcal{N}(0, 4)$, $x_1 \sim \mathcal{N}(0, 1)$, $z \perp x_1$.
- $y = z + y_1$, $z \sim \mathcal{N}(0, 4)$, $y_1 \sim \mathcal{N}(0, 1)$, $z \perp y_1$.

$$cor(x, y) = \frac{Var(z)}{se(x)se(y)}.$$

- $\rho_{xy|z} = 0$
- So: x and y are correlated but not partially correlated given z.

Definition 3.9: Partial autocorrelation function (PACF)

- $\phi_{11} = cor(x_{t+1}, x_t)$.
- $\phi_{hh} = cor(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t)$.

- $\hat{x}_{t+h}$: predicted x_{t+h} using $x_{t+h-1}, \ldots, x_{t+1}$.
- $\hat{x}_t$: predicted x_t using $x_{t+h-1}, \ldots, x_{t+1}$.

Example 3.15: PACF of AR(1)

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

- $\phi_{11} = \rho(1) = \phi$.
- $\phi_{22} = cor(x_{t+2} - \hat{x}_{t+2}, x_t - \hat{x}_t)$.

- Question: what is $\hat{x}_{t+2}$?
- Set $\hat{x}_{t+2} = \beta_1 x_{t+1}$ and choose β_1 to minimize

$$E(x_{t+2} - \hat{x}_{t+2})^2 = E(x_{t+2} - \beta_1 x_{t+1})^2$$

Minimize:

$$E(x_{t+2} - \hat{x}_{t+2})^2 = E(x_{t+2} - \beta_1 x_{t+1})^2 = \gamma(0) - 2\beta_1\gamma(1) + \beta_1^2\gamma(0)$$
$$\Rightarrow \beta_1 = \gamma(1)/\gamma(0) = \rho(1) = \phi$$

Similarly, $\hat{x}_t = \beta_2 x_{t+1} \Rightarrow \beta_2 = \phi$.

So, we have:

$$\begin{aligned}
\phi_{22} &= cor(x_{t+2} - \hat{x}_{t+2}, x_t - \hat{x}_t) \\
&= cor(x_{t+2} - \phi x_{t+1}, x_t - \phi x_{t+1}) \\
&= cor(w_{t+2}, x_t - \phi x_{t+1}) = 0
\end{aligned}$$

Example 3.16: AR(p): $x_{t+h} = \sum_{j=1}^{p} \phi_j x_{t+h-j} + w_{t+h}$

- Roots of $\phi(z)$ outside unit circle.
- For $h > p$, regress x_{t+h} on $\{x_{t+1}, x_{t+2}, \ldots, x_{t+h}\}$.

$$\hat{x}_{t+h} = \sum_{j=1}^{p} \phi_j x_{t+h-j}$$

- Thus, $\phi_{hh} = cor(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t) = cor(w_{t+h}, x_t - \hat{x}_t) = 0.$

Example 3.17 PACF for MA(q)

An invertible MA(q):

$$x_t = -\sum_{j=1}^{\infty} \pi_j x_{t-j} + w_t$$

- No finite representation exists.
- The PACF will never cut off.

Example 3.17 PACF for MA(q)

An invertible MA(1):

$$x_t = w_t + \theta w_{t-1}, |\theta| < 1.$$

Following calculations similar to Example 3.15:

- $\phi_{22} = -\theta^2/(1 + \theta^2 + \theta^4)$.
- $\phi_{hh} = -\frac{(-\theta)^h(1-\theta^2)}{1-\theta^{2(h+1)}}$.

Estimation

Example 3.29 Method of moments estimation for an MA(1)

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

with $|\theta| < 1$.
Use the same idea as for Yule-Walker Equation:

- $\gamma(0) = \sigma_w^2(1 + \theta^2)$
- $\gamma(0) = \sigma_w^2\theta$

$\hat{\theta}$ can be found by solving

$$\hat{\rho}(1) = \frac{\hat{\gamma}(0)}{\hat{\gamma}(0)} = \frac{\hat{\theta}}{1 + \hat{\theta}^2} \tag{1}$$

Estimation

Example 3.29 Method of moments estimation for an MA(1)

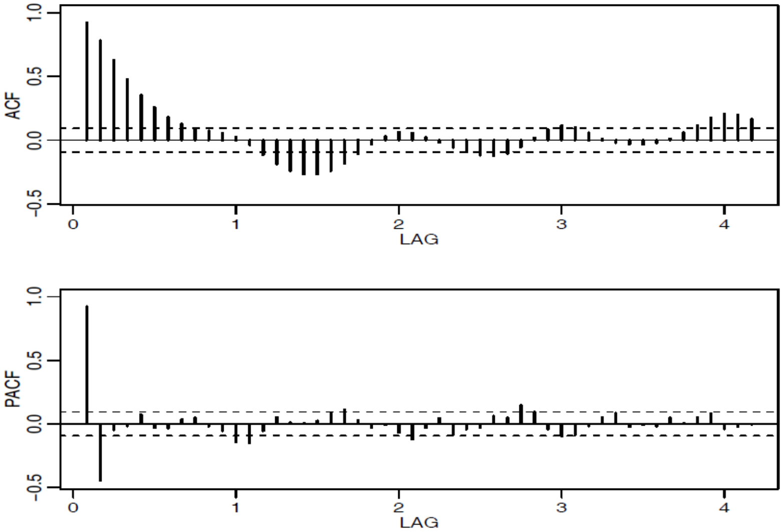
$$x_t = w_t + \theta w_{t-1}$$

with $|\theta| < 1$.
$\hat{\theta}$ can be found by solving

$$\hat{\rho}(1) = \frac{\hat{\gamma}(0)}{\hat{\gamma}(0)} = \frac{\hat{\theta}}{1 + \hat{\theta}^2} \tag{2}$$

- Two solutions exist, pick the invertible one.
- If $|\hat{\rho}(1)| \leq 1/2$, the solutions are real. Otherwise, no real solution exists.
- Example: $\hat{\rho}(1) = .507$ while $\rho(1) = .9/(1 + .9^2) = .497$.

Examples 3.28 and 3.31 Preliminary analysis of the recruitment series.



Estimation

Example 3.33: Fitting the glacial varve series

- Paleocolimatic glacial varve thicknesses from Massachusetts for $n = 634$ years.
- Melting glaciers deposit yearly layers of sand and silt during the spring melting seasons.
- Begining 11,834 years ago.

To do:

- Estimation.
- Model diagnosis.

ARMA model – Definition

A time series $\{x_t; t = 0, \pm 1, \ldots\}$ is ARMA(p,q) if:

- $x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q}$.
- $\phi_p \neq 0$, $\theta_q \neq 0$, $\sigma_w^2 > 0$.
- Stationary.
- p is the AR order and q is the MA order.

ARMA model – Definition

Definition: AR and MA polynomials:

- AR polynomials: $\Phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p$, $\phi_p \neq 0$.
- MA polynomials: $\Theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$, $\theta_q \neq 0$.
- z is a complex number.

ARMA model – Redundancy

Parameter redundancy:

$$\Phi(B)x_t = \Theta(B)w_t$$
$$\Rightarrow \eta(B)\Phi(B)x_t = \eta(B)\Theta(B)w_t$$

Example:

- Ture model: $x_t = w_t$.
- Let $\eta(B) = 1 - 0.5B$, then we have

$$(1 - 0.5B)x_t = (1 - 0.5B)w_t$$
$$\Rightarrow x_t = 0.5x_{t-1} - 0.5w_{t-1} + w_t$$

- It looks like an ARMA model.

ARMA model – Redundancy

We use ARMA(p,q) model to refer to its simplest form.

- That is $\Phi(z)$ and $\theta(z)$ have no common factors.

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

- $\Phi(z) = 1 - 0.5z.$
- $\Theta(z) = 1 - 0.5z.$
- $\Phi(z)$ and $\Theta(z)$ have common factor.
- $\Phi(z) = 1$ and $\Theta(z) = 1.$
- It is a ARMA(0,0) process.

ARMA model – Causality

Definition 3.7: Concept of causality

- An ARMA(p,q) model for x_t is said to be causal if x_t can be written as a one-sided linear process:

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \Psi(B) w_t$$

- where $\Psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$.
- $\psi_0 = 1$.
- It is also called the MA representation.

ARMA model – Causality

An ARMA(p,q) model is causal if and only if:

- $\Phi(z) \neq 0$ for $|z| \leq 1$.
- Then $x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$.
- $\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \Theta(z)/\Phi(z)$, $|z| \leq 1$.
- Equivalently $\Phi(z) = 0$ only when $|z| > 1$.

ARMA model – Invertibility

Definition 3.8: An ARMA(p,q) model is invertible, if the time series $\{x_t : t = 0, \pm 1, \pm 2, \ldots\}$ can be written as

$$\Pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t$$

- $\Pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$;
- $\sum_{j=0}^{\infty} |\pi_j| < \infty$;
- $\pi_0 = 1$;
- It is called the invertible representation.

ARMA model – Invertibility

Prop 3.2: Invertibility of an ARMA(p,q) process.

An ARMA(p,q) model is invertible if and only if

- $\Theta(z) \neq 0$ for $|z| \leq 1$.
- Then, $\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \Phi(z)/\Theta(z)$, $|z| \leq 1$.
- In other words, roots of $\Theta(z)$ lie outside the unit circle.
- $\Theta(z) = 0$ only when $|z| > 1$.

ARMA model – Example

Example 3.7: Parameter redundancy, causality and invertibility

$$x_t = 0.4x_{t-1} + 0.45x_{t-2} + w_t + w_{t-1} + 0.25w_{t-2}$$

- $(1 - 0.4B - 0.45B^2)x_t = (1 + B + 0.25B^2)w_t$.
- It appears to be an ARMA(2,2) process.

- $\Phi(B) = (1 - 0.4B - 0.45B^2) = (1 - 0.9B)(1 + 0.5B)$.
- $\Theta(B) = (1 + B + 0.25B^2) = (1 + 0.5B)^2$.
- Common factor: $(1 + 0.5B)$. ARMA(1,1)!

Exemple 3.7 continued:

$$(1 - 0.9B)x_t = (1 + 0.5B)w_t$$

- $x_t = 0.9x_t + 0.5w_{t-1} + w_t$.

- Causal: $\Phi(z) = (1 - 0.9z) = 0 \Rightarrow z = 10/9$.
- Invertible: $\Theta(z) = (1 + 0.5z) = 0 \Rightarrow z = -2$.

ARMA model – Example

Exemple 3.7 continued:

$$(1 - 0.9B)x_t = (1 + 0.5B)w_t$$

- The MA representation:

$$x_t = (1 + \psi_1 B + \psi_2 B^2 + \ldots)w_t.$$

- $(1 - 0.9z)(1 + \psi_1 z + \psi_2 z^2 + \ldots) = 1 + 0.5z$
- $1+(\psi_1-0.9)z+(\psi_2-0.9\psi_1)z^2+\ldots+(\psi_j-0.9\psi_{j-1})z^2+\ldots = 1+0.5z.$

- $\psi_1 - 0.9 = 0.5 \Rightarrow \psi_1 = 1.4.$
- $\psi_j - 0.9\psi_{j-1} = 0 \Rightarrow \psi_j = 1.4 \times 0.9^{(j-1)}$

How about the invertible representation?

Reading Materials

- Ch 3.1 p.81-83
- Ch 3.3 p.97-99
- Ch 3.5 p.113-120