4.3 Moving Average Process MA(q)

Definition 4.5. $\{X_t\}$ *is a* ***moving-average process of order q*** *if*

$$X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}, \tag{4.9}$$

where

$$Z_t \sim WN(0, \sigma^2)$$

and $\theta_1, \ldots, \theta_q$ *are constants.*

Remark 4.6. X_t is a linear combination of $q+1$ white noise variables and we say that it is q**-correlated**, that is X_t and $X_{t+\tau}$ are uncorrelated for all lags $\tau > q$.

Remark 4.7. If Z_t is an i.i.d process then X_t is a strictly stationary TS since

$$(Z_t, \ldots, Z_{t-q})^{\mathrm{T}} \stackrel{d}{=} (Z_{t+\tau}, \ldots, Z_{t-q+\tau})^{\mathrm{T}}$$

for all τ. Then it is called q**-dependent**, that is X_t and $X_{t+\tau}$ are independent for all lags $\tau > q$.

Remark 4.8. Obviously,

- IID noise is a 0-dependent TS.
- White noise is a 0-correlated TS.
- MA(1) is 1-correlated TS if it is a combination of WN r.vs, 1-dependent if it is a combination of IID r.vs.

Remark 4.9. The MA(q) process can also be written in the following equivalent form

$$X_t = \theta(B) Z_t, \tag{4.10}$$

where the **moving average operator**

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q \tag{4.11}$$

defines a linear combination of values in the shift operator $B^k Z_t = Z_{t-k}$.

Example 4.4. MA(2) process.
This process is written as

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} = (1 + \theta_1 B + \theta_2 B^2) Z_t. \tag{4.12}$$

What are the properties of MA(2)? As it is a combination of a zero mean white noise, it also has zero mean, i.e.,

$$\mathrm{E}\, X_t = \mathrm{E}(Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}) = 0.$$

It is easy to calculate the covariance of X_t and $X_{t+\tau}$. We get

$$\gamma(\tau) = \operatorname{cov}(X_t, X_{t+\tau}) = \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma^2 & \text{for } \tau = 0, \\ (\theta_1 + \theta_1\theta_2)\sigma^2 & \text{for } \tau = \pm 1, \\ \theta_2 \sigma^2 & \text{for } \tau = \pm 2, \\ 0 & \text{for } |\tau| > 2, \end{cases}$$

which shows that the autocovariances depend on lag, but not on time. Dividing $\gamma(\tau)$ by $\gamma(0)$ we obtain the autocorrelation function,

$$\rho(\tau) = \begin{cases} 1 & \text{for } \tau = 0, \\ \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} & \text{for } \tau = \pm 1, \\ \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2} & \text{for } \tau = \pm 2 \\ 0 & \text{for } |\tau| > 2. \end{cases}$$

MA(2) process is a weakly stationary, 2-correlated TS. □

Figure 4.5 shows MA(2) processes obtained from the simulated Gaussian white noise shown in Figure 4.1 for various values of the parameters (θ_1, θ_2).
The blue series is

$$x_t = z_t + 0.5 z_{t-1} + 0.5 z_{t-2},$$

while the purple series is

$$x_t = z_t + 5 z_{t-1} + 5 z_{t-2},$$

where z_t are realizations of an i.i.d. Gaussian noise.
As you can see very different processes can be obtained for different sets of the parameters. This is an important property of MA(q) processes, which is a very large family of models. This property is reinforced by the following Proposition.

Proposition 4.2. *If $\{X_t\}$ is a stationary q-correlated time series with mean zero, then it can be represented as an MA(q) process.* □

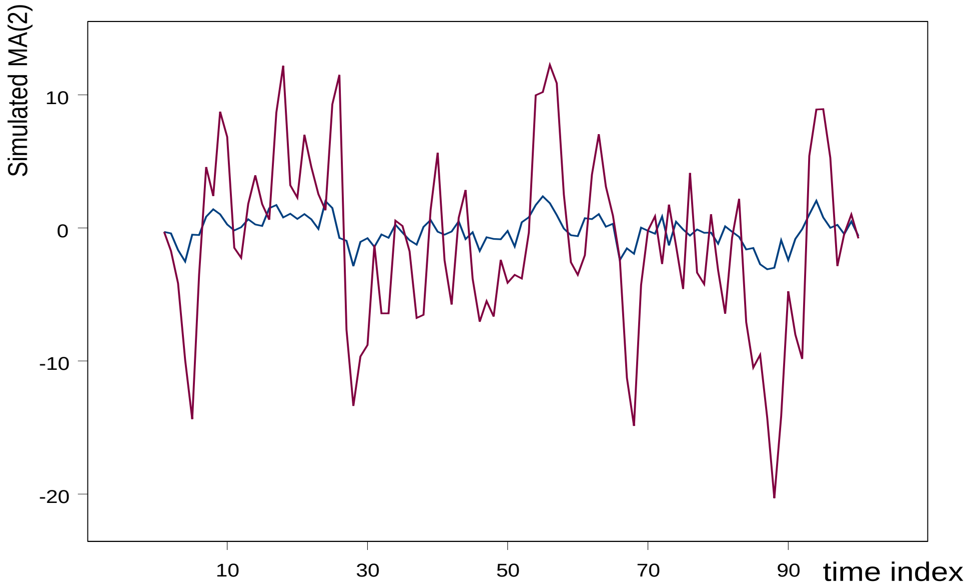
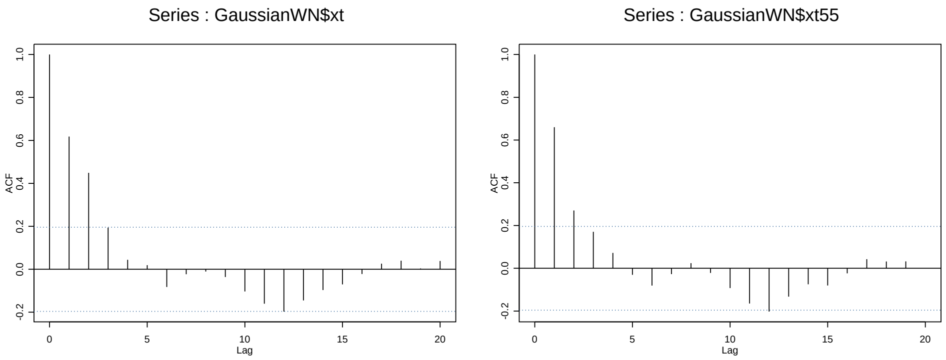


Figure 4.5: Two simulated MA(2) processes, both from the white noise shown in Figure 4.1, but for different sets of parameters: $(\theta_1, \theta_2) = (0.5, 0.5)$ and $(\theta_1, \theta_2) = (5, 5)$.



(a) (b)

Figure 4.6: (a) Sample ACF for $x_t = z_t + 0.5z_{t-1} + 0.5z_{t-2}$ and (b) for $x_t = z_t + 5z_{t-1} + 5z_{t-2}$.

Also, the following theorem gives the form of ACF for a general MA(q).

Theorem 4.2. *An MA(q) process (as in Definition 4.5) is a weakly stationary TS with the ACVF*

$$\gamma(\tau) = \begin{cases} \sigma^2 \sum_{j=0}^{q-|\tau|} \theta_j \theta_{j+|\tau|}, & \textit{if } |\tau| \leq q, \\ 0, & \textit{if } |\tau| > q, \end{cases} \tag{4.13}$$

where θ_0 is defined to be 1.

□

The ACF of an MA(q) has a distinct "cut-off" at lag $\tau = q$. Furthermore, if q is small the maximum value of $|\rho(1)|$ is well below unity. It can be shown that

$$|\rho(1)| \leq \cos\left(\frac{\pi}{q+2}\right). \tag{4.14}$$

4.3.1 Non-uniqueness of MA Models

Consider an example of MA(1)

$$X_t = Z_t + \theta Z_{t-1}$$

whose ACF is

$$\rho(\tau) = \begin{cases} 1, & \text{if } \tau = 0, \\ \frac{\theta}{1+\theta^2} & \text{if } \tau = \pm 1, \\ 0, & \text{if } |\tau| > 1. \end{cases}$$

For $q = 1$, formula (4.14) means that the maximum value of $|\rho(1)|$ is 0.5. It can be verified directly from the formula for the ACF above. Treating $\rho(1)$ as a function of θ we can calculate its extrema. Denote

$$f(\theta) = \frac{\theta}{1+\theta^2}.$$

Then

$$f'(\theta) = \frac{1-\theta^2}{(1+\theta^2)^2}.$$

The derivative is equal to 0 at $\theta = \pm 1$ and the function f attains maximum at $\theta = 1$ and minimum at $\theta = -1$. We have

$$f(1) = \frac{1}{2}, \quad f(-1) = -\frac{1}{2}.$$

This fact can be helpful in recognizing MA(1) processes. In fact, MA(1) with $|\theta| = 1$ may be uniquely identified from the autocorrelation function.

However, it is easy to see that the form of the ACF stays the same for θ and for $\frac{1}{\theta}$. Take for example 5 and $\frac{1}{5}$. In both cases

$$\rho(\tau) = \begin{cases} 1 & \text{if } \tau = 0, \\ \frac{5}{26} & \text{if } \tau = \pm 1, \\ 0 & \text{if } |\tau| > 1. \end{cases}$$

Also, the pair $\sigma^2 = 1, \theta = 5$ gives the same ACVF as the pair $\sigma^2 = 25, \theta = \frac{1}{5}$, namely

$$\gamma(\tau) = \begin{cases} (1+\theta^2)\sigma^2 = 26, & \text{if } \tau = 0, \\ \theta\sigma^2 = 5, & \text{if } \tau = \pm 1, \\ 0, & \text{if } |\tau| > 1. \end{cases}$$

Hence, the MA(1) processes

$$X_t = Z_t + \frac{1}{5} Z_{t-1}, \quad Z_t \underset{iid}{\sim} \mathcal{N}(0, 25)$$

and

$$X_t = Y_t + 5Y_{t-1}, \quad Y_t \underset{iid}{\sim} \mathcal{N}(0, 1)$$

are the same. We can observe the variable X_t, not the noise variable, so we can not distinguish between these two models. Except for the case of $|\theta| = 1$, a particular autocorrelation function will be compatible with two models.

To which of the two models we should restrict our attention?

4.3.2 Invertibility of MA Processes

The MA(1) process can be expressed in terms of lagged values of X_t by substituting repeatedly for lagged values of Z_t. We have

$$Z_t = X_t - \theta Z_{t-1}.$$

The substitution yields

$$\begin{aligned} Z_t &= X_t - \theta Z_{t-1} \\ &= X_t - \theta(X_{t-1} - \theta Z_{t-2}) \\ &= X_t - \theta X_{t-1} + \theta^2 Z_{t-2} \\ &= X_t - \theta X_{t-1} + \theta^2 (X_{t-2} - \theta Z_{t-3}) \\ &= X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \theta^3 Z_{t-3} \\ &= \dots \\ &= X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \theta^3 X_{t-3} + \theta^4 X_{t-4} + \dots + (-\theta)^n Z_{t-n}. \end{aligned}$$

This can be rewritten as

$$(-\theta)^n Z_{t-n} = Z_t - \sum_{j=0}^{n-1} (-\theta)^j X_{t-j}.$$

However, if $|\theta| < 1$, then

$$\mathrm{E}\left(Z_t - \sum_{j=0}^{n-1} (-\theta)^j X_{t-j}\right)^2 = \mathrm{E}\left(\theta^{2n} Z_{t-n}^2\right) \underset{n\to\infty}{\longrightarrow} 0$$

and we say that the sum is convergent in the mean square sense. Hence, we obtain another representation of the model

$$Z_t = \sum_{j=0}^{\infty} (-\theta)^j X_{t-j}.$$

This is a representation of another class of models, called infinite autoregressive (AR) models. So we inverted MA(1) to an infinite AR. It was possible due to the assumption that $|\theta| < 1$. Such a process is called an **invertible process**. This is a desired property of TS, so in the example we would choose the model with $\sigma^2 = 25, \theta = \frac{1}{5}$.

4.4 Linear Processes

Definition 4.6. *The TS $\{X_t\}$ is called a* ***linear process*** *if it has the representation*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \tag{4.15}$$

for all t, where $Z_t \sim WN(0, \sigma^2)$ and $\{\psi_j\}$ is a sequence of constants such that $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$. □

Remark 4.10. The condition $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ ensures that the process converges in the mean square sense, that is

$$\mathrm{E}\left(X_t - \sum_{j=-n}^{n} \psi_j Z_{t-j}\right)^2 \to 0 \quad \text{as } n \to \infty.$$

Remark 4.11. MA(∞) is a linear process with $\psi_j = 0$ for $j < 0$ and $\psi_j = \theta_j$ for $j \geq 0$, that is MA(∞) has the representation

$$X_t = \sum_{j=0}^{\infty} \theta_j Z_{t-j},$$

where $\theta_0 = 1$.

Note that the formula (4.15) can be written using the backward shift operator B. We have

$$Z_{t-j} = B^j Z_t.$$

Hence

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} = \sum_{j=-\infty}^{\infty} \psi_j B^j Z_t.$$

Denoting

$$\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j, \tag{4.16}$$

we can write the linear process in a neat way

$$X_t = \psi(B)Z_t.$$

The operator $\psi(B)$ is a linear filter, which when applied to a stationary process produces a stationary process. This fact is proved in the following proposition.

Proposition 4.3. *Let $\{Y_t\}$ be a stationary TS with mean zero and autocovariance function γ_Y. If $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, then the process*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j} = \psi(B)Y_t \tag{4.17}$$

is stationary with mean zero and autocovariance function

$$\gamma_X(\tau) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(\tau - k + j). \tag{4.18}$$

Proof. The assumption $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ assures convergence of the series. Now, since $\mathrm{E}\, Y_t = 0$, we have

$$\mathrm{E}\, X_t = \mathrm{E}\left(\sum_{j=-\infty}^{\infty} \psi_j Y_{t-j}\right) = \sum_{j=-\infty}^{\infty} \psi_j \,\mathrm{E}(Y_{t-j}) = 0$$

and

$$\begin{aligned}
\mathrm{E}(X_t X_{t+\tau}) &= \mathrm{E}\left[\left(\sum_{j=-\infty}^{\infty} \psi_j Y_{t-j}\right)\left(\sum_{k=-\infty}^{\infty} \psi_k Y_{t+\tau-k}\right)\right] \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \,\mathrm{E}(Y_{t-j} Y_{t+\tau-k}) \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(\tau - k + j).
\end{aligned}$$

It means that $\{X_t\}$ is a stationary TS with the autocavarianxe function given by formula (4.18).

□

Corrolary 4.1. *If $\{Y_t\}$ is a white noise process, then $\{X_t\}$ given by* (4.17) *is a stationary linear process with zero mean and the ACVF*

$$\gamma_X(\tau) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+\tau} \sigma^2. \tag{4.19}$$

□