Time Series Exercise Sheet 2

Sofia Olhede

February 27, 2025

Exercise 2.1

When we defined the MA(q) process we specified $\theta_0 = -1$. In terms of the covariance structure for moving average processes, why do we not gain more generality by letting θ_0 be an arbitrary number?

Solution 2.1

If we change the first coefficient, then we can divide through the specifying equation by that number and obtain a new process with the same autocorrelation. So for identifiability purpose, we force the coefficient to be $\theta_0 \equiv -1$.

For example, let's consider the first order moving average process, MA(1), when $\mu = 0$. In other words, the process is determined by $\theta_0 = -1$ and θ_1:

$$X_t = -\theta_0 \epsilon_t - \theta_1 \epsilon_{t-1} = \epsilon_t - \theta_1 \epsilon_{t-1}. \tag{1}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. Assuming there is another parametrization for the process using $\theta_0' = c\theta_0 = c$ and θ_1' for some $c \neq 0$.

$$X_t = -\theta_0' \epsilon_t' - \theta_1' \epsilon_{t-1}' = c\epsilon_t' - \theta_1' \epsilon_{t-1}'. \tag{2}$$

where $\epsilon_t' \sim N(0, \sigma_{\epsilon'}^2)$. If we rewrite (1) as

$$\begin{aligned} X_t &= \epsilon_t - \theta_1 \epsilon_{t-1} \\ &= c\frac{\epsilon_t}{c} - (c\theta_1)\frac{\epsilon_{t-1}}{c}. \end{aligned}$$

Hence, $\theta_1' = c\theta_1$ and $\epsilon_t' = c^{-1}\epsilon_t$ for all t. Moreover, $\sigma_{\epsilon'}^2 = c^{-2}\sigma_\epsilon^2$. Note that the two representations have the same autocovariance (and autocorrelation):

$$\begin{aligned} \mathrm{Var}\,(X_t) &= (1+\theta_1^2)\sigma_\epsilon^2 && \text{from (1)} \\ \mathrm{Var}\,(X_t) &= \{(\theta_0')^2 + (\theta_1')^2\}\sigma_{\epsilon'}^2 \\ &= c^2(1+\theta_1^2)c^{-2}\sigma_\epsilon^2 = (1+\theta_1^2)\sigma_\epsilon^2 && \text{from (2)} \\ \mathrm{Cov}\,(X_t, X_{t-1}) &= \theta_0\theta_1\sigma_\epsilon^2 \end{aligned}$$

This shows us that there is an identifiability issue if we do not constrain $\theta_0 = -1$. Specifically, there are multiple MA(1) representations for X_t, making the representation of X_t non-unique.

Using the same argument to the general MA(q) process, we will reach the same conclusion as the MA(1) example.

Exercise 2.2

Determine that the moving average process defined by

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

can be written as

$$X_t = \varepsilon_t - \sum_{j=1}^{p} \theta^j X_{t-j} - \theta^{p+1}\varepsilon_{t-p-1}$$

for any positive integer value of p.

Solution 2.2

Note that for any integer value of $p \geqslant 1$, we have $X_{t-p} = \varepsilon_{t-p} - \theta\varepsilon_{t-p-1}$, so

$$\varepsilon_{t-p} = X_{t-p} + \theta\varepsilon_{t-p-1}$$

Applying the latter recursively we have

$$\begin{aligned} X_t = \varepsilon_t - \theta\varepsilon_{t-1} = \varepsilon_t - \theta\left(X_{t-1} + \theta\varepsilon_{t-2}\right) &= \varepsilon_t - \theta X_{t-1} - \theta^2\varepsilon_{t-2} \\ &= \varepsilon_t - \theta X_{t-1} - \theta^2\left(X_{t-2} + \theta\varepsilon_{t-3}\right) \\ &= \varepsilon_t - \theta X_{t-1} - \theta^2 X_{t-2} - \theta^3\varepsilon_{t-3} \\ &= \cdots \\ &= \varepsilon_t - \sum_{j=1}^{p} \theta^j X_{t-j} - \theta^{p+1}\varepsilon_{t-p-1} \end{aligned}$$

Exercise 2.3

Assume that $Y_t = \sum_{s=-p}^{p} g_{s-t}\epsilon_s$ and $Z_t = \sum_{s=-p}^{p} h_{s-t}\epsilon_s$ where ϵ_t is zero-mean white noise. We define X_t pointwise by $Y_t + Z_t$. Determine the first and second moments of X_t.

Solution 2.3

Let $\mathrm{Var}\left(\epsilon_s\right) = \sigma^2$ and rewrite $X_t = \sum_{s=-p}^{p}\left(g_{s-t} + h_{s-t}\right)\epsilon_s$. As ϵ_s are uncorrelated, we have for all $t, \tau \in T$

$$\mathbb{E}\left[X_t\right] = \sum_{s=-p}^{p}\left(g_{s-t} + h_{s-t}\right)\mathrm{E}\left(\epsilon_s\right) = \sum_{s=-p}^{p}\left(g_{s-t} + h_{s-t}\right)\cdot 0 = 0$$

$$\begin{aligned} \mathbb{E}\left[X_t X_{t+\tau}\right] = \mathrm{Cov}\left(X_t, X_{t+\tau}\right) &= \mathrm{Cov}\left(\sum_{s=-p}^{p}\left(g_{s-t} + h_{s-t}\right)\epsilon_s, \sum_{l=-p}^{p}\left(g_{l-t-\tau} + h_{l-t-\tau}\right)\epsilon_l\right) \\ &= \sum_{s=-p}^{p}\sum_{l=-p}^{p}\left(g_{s-t} + h_{s-t}\right)\left(g_{l-t-\tau} + h_{l-t-\tau}\right)\mathrm{Cov}\left(\epsilon_s, \epsilon_l\right) \\ &= \sigma^2\sum_{s=-p}^{p}\left(g_{s-t} + h_{s-t}\right)\left(g_{s-t-\tau} + h_{s-t-\tau}\right) \end{aligned}$$

Exercise 2.4

In the lectures we defined two different estimators for the ACVS:

$$\tilde{\gamma}_\tau = \frac{1}{n-|\tau|}\sum_{t=1}^{n-|\tau|}\left\{X_t - \bar{X}\right\}\left\{X_{t+|\tau|} - \bar{X}\right\},$$

$$\hat{\gamma}_\tau = \frac{1}{n}\sum_{t=1}^{n-|\tau|}\left\{X_t - \bar{X}\right\}\left\{X_{t+|\tau|} - \bar{X}\right\}.$$

If X_t has independent Gaussian realizations at each time point please calculate the mean, variance and mean-square error of these estimators. Repeat the calculation for an MA(1) process. As in class to do the analysis replace $\bar{X}$ by $\mu = \mathrm{E}\left(X_t\right)$, arguing that for large samples this will be appropriate.

(Hint): There is a known result (Isserlis' theorem) that if $\mathbf{X} = (X_1, X_2, X_3, X_4)$ is a multivariate normal random vector, then:

$$\mathbb{E}\left[X_1X_2X_3X_4\right] = \mathbb{E}\left[X_1X_2\right]\mathbb{E}\left[X_3X_4\right] + \mathbb{E}\left[X_1X_3\right]\mathbb{E}\left[X_2X_4\right] + \mathbb{E}\left[X_1X_4\right]\mathbb{E}\left[X_2X_3\right].$$

Solution 2.4

For a white noise process we have (informally replacing $\bar{X}$ with μ like in class)

$$\mathbb{E}\left[\hat{\gamma}_\tau\right] = \frac{1}{n}\sum_{t=1}^{n-|\tau|}\mathbb{E}\left[\left(X_t-\mu\right)\left(X_{t+|\tau|}-\mu\right)\right]$$

$$= \begin{cases}\sigma_X^2 & \text{if } \tau=0\\ 0 & \text{otherwise.}\end{cases} \quad \mathbb{E}\left[\tilde{\gamma}_\tau\right] \quad = \begin{cases}\sigma_X^2 & \text{if } \tau=0\\ 0 & \text{otherwise.}\end{cases}$$

Furthermore,

$$\begin{aligned}\operatorname{Var}\left(\hat{\gamma}_\tau\right) &= \frac{1}{n^2}\sum_{t_1=1}^{n-|\tau|}\sum_{t_2=1}^{n-|\tau|}\operatorname{Cov}\left(\left(X_{t_1}-\mu\right)\left(X_{t_1+|\tau|}-\mu\right),\left(X_{t_2}-\mu\right)\left(X_{t_2+|\tau|}-\mu\right)\right)\\ &= \frac{1}{n^2}\sum_{t_1=1}^{n-|\tau|}\sum_{t_2=1}^{n-|\tau|}\left\{\mathbb{E}\left[\left(X_{t_1}-\mu\right)\left(X_{t_1+|\tau|}-\mu\right)\left(X_{t_2}-\mu\right)\left(X_{t_2+|\tau|}-\mu\right)\right]\right.\\ &\qquad\left.-\mathbb{E}\left[\left(X_{t_1}-\mu\right)\left(X_{t_1+|\tau|}-\mu\right)\right]\mathbb{E}\left[\left(X_{t_2}-\mu\right)\left(X_{t_2+|\tau|}-\mu\right)\right]\right\}\end{aligned}$$

Using Isserlis' theorem, we find

$$\begin{aligned}&\mathbb{E}\left[\left(X_{t_1}-\mu\right)\left(X_{t_1+|\tau|}-\mu\right)\left(X_{t_2}-\mu\right)\left(X_{t_2+|\tau|}-\mu\right)\right]\\ &\quad=\gamma_X(|\tau|)\gamma_X(|\tau|)+\gamma_X\left(t_1-t_2\right)\gamma_X\left(t_1-t_2\right)+\gamma_X\left(t_1-t_2-|\tau|\right)\gamma_X\left(t_1-t_2+|\tau|\right)\end{aligned}$$

For white noise we have $\gamma_X(\tau)=\sigma_X^2\delta_\tau$ where

$$\delta_\tau = \begin{cases}1 & \tau=0,\\ 0 & \text{otherwise.}\end{cases}$$

Thus we get

$$\operatorname{Cov}\left(\left(X_{t_1}-\mu\right)\left(X_{t_1+|r|}-\mu\right),\left(X_{t_2}-\mu\right)\left(X_{t_2+|\tau|}-\mu\right)\right)=\sigma_X^4\delta_{t_1-t_2}+\sigma_X^4\delta_{t_1-t_2}\delta_\tau.$$

We therefore find that

$$\operatorname{Var}\left(\hat{\gamma}_\tau\right)=\frac{1+\delta_\tau}{n^2}\sigma_X^4(n-|\tau|).$$

It follows directly that

$$\operatorname{Var}\left(\tilde{\gamma}_\tau\right)=\frac{1+\delta_\tau}{n-|\tau|}\sigma_X^4.$$

The mean square error is then (for white noise)

$$\operatorname{MSE}\left(\hat{\gamma}_\tau\right)=(0)^2+\frac{n-|\tau|}{n^2}\left(1+\delta_\tau\right)\sigma_X^4,$$

whilst

$$\operatorname{MSE}\left(\tilde{\gamma}_\tau\right)=(0)^2+\frac{1}{n-|\tau|}\left(1+\delta_\tau\right)\sigma_X^4.$$

Thus is this case the biased estimator $\hat{\gamma}$ wins out.

For the MA(1) process, $\operatorname{Cov}\left(X_t,X_{t+|\tau|}\right)=\delta_\tau\left(1+\theta^2\right)\sigma_X^2-\delta_{|\tau|-1}\theta\sigma_X^2$ and this goes into the calculations.

Exercise 2.5

Assume we know the mean of X_t and it is μ. Form the estimator

$$\bar{\gamma}_\tau^{(\alpha)}=\alpha\tilde{\gamma}_\tau+(1-\alpha)\hat{\gamma}_\tau$$

For X_t Gaussian white noise determine the MSE of $\bar{\gamma}_\tau^{(\alpha)}$. For $\sigma_X^2=1$, $\tau=1$ and $n=10$ plot it as a function of α. What value of α is appropriate?

Solution 2.5

Wlog we may assume that $\mu = 0$, else we may define a new process, say $X_t^* = X_t - \mu$, s.t. $\gamma_{X^*} = \gamma_X$. For a white noise process, we note that both $\tilde{\gamma}$ and $\hat{\gamma}$ are unbiased, and hence $\bar{\gamma}^{(u)}$ is also unbiased, i.e. $\mathbb{E}[\bar{\gamma}_\tau^{(\alpha)}] = \gamma(\tau)$; see Exercise 2.4. Thus, for a white noise process, we may re-write the $\text{MSE}(\bar{\gamma}_\tau^{(\alpha)})$ as:

$$\begin{aligned}\text{MSE}(\bar{\gamma}_\tau^{(\alpha)}) &= \mathbb{E}[(\bar{\gamma}_\tau^{(\alpha)} - \gamma(\tau))^2] = \mathbb{E}[(\bar{\gamma}_\tau^{(\alpha)} - \mathbb{E}[\bar{\gamma}_\tau^{(\alpha)}])^2] = \text{Cov}\left(\bar{\gamma}_\tau^{(\alpha)}, \bar{\gamma}_\tau^{(\alpha)}\right) \\ &= \text{Cov}\left(\alpha\tilde{\gamma}_\tau + (1-\alpha)\hat{\gamma}_\tau, \alpha\tilde{\gamma}_\tau + (1-\alpha)\hat{\gamma}_\tau\right).\end{aligned}$$

We continue by expanding the terms inside the covariance

$$\text{MSE}(\bar{\gamma}_\tau^{(\alpha)}) = \alpha^2\text{Cov}(\tilde{\gamma}_\tau, \tilde{\gamma}_\tau) + 2\alpha(1-\alpha)\text{Cov}(\tilde{\gamma}_\tau, \hat{\gamma}_\tau) + (1-\alpha)^2\text{Cov}(\hat{\gamma}_\tau, \hat{\gamma}_\tau). \tag{3}$$

We use the fact that $\hat{\gamma}_\tau = \frac{n-|\tau|}{n}\tilde{\gamma}_\tau$ to compute

$$\begin{aligned}\text{Cov}\left(\tilde{\gamma}_\tau, \hat{\gamma}_\tau)\right) &= \frac{n-|\tau|}{n}\text{Var}(\tilde{\gamma}_\tau), \\ \text{Var}(\hat{\gamma}_\tau) &= \left\{\frac{n-|\tau|}{n}\right\}^2\text{Var}(\tilde{\gamma}_\tau).\end{aligned}$$

For the time being, assume that $\text{Var}(\tilde{\gamma}_\tau) = V(\tau)$, and let $C(\tau) = \frac{n-|\tau|}{n}$. Then we note

$$\text{MSE}(\bar{\gamma}_\tau^{(\alpha)}) = (\alpha - \alpha C(\tau) + C(\tau))^2 V(\tau).$$

A good choice of α can be computed by minimizing $\text{MSE}(\bar{\gamma}_\tau^{(\alpha)})$. Since $V(\tau)$ is not dependent on α for a given τ, we simply minimize $(\alpha - \alpha C(\tau) + C(\tau))^2$. We have for $\tau > 0$ that $C(\tau) < 1$ hence the expression in brackets is always positive for $|\tau| < n$. Differentiation wrt. α yields

$$(\alpha - \alpha C(\tau) + C(\tau))(1 - C(\tau)) \stackrel{!}{=} 0,$$

from where, for $\tau > 0$, we can pick $\alpha = C(\tau)/(C(\tau) - 1)$. For $\tau = 0$ any choice of α leaves the MSE unchanged.

We demonstrate this for the case when X_t is a Gaussian white noise. From Exercise 2.4, we already know $\text{Var}(\tilde{\gamma}_\tau) = \frac{1+\delta_\tau}{n-|\tau|}\sigma_X^4$. Finally, for $\tau > 0$, replacing $n = 10$, $\sigma_X = 1$ we obtain

$$\text{MSE}(\bar{\gamma}_\tau^{(\alpha)}) = \left\{\alpha - \alpha\frac{n-|\tau|}{n} + \frac{n-|\tau|}{n}\right\}^2 \frac{1}{n-|\tau|} \tag{4}$$

In Figure 1 we plot the MSE for $\tau = 1$ as a function of α.

Exercise 2.6

Show that for any series $\{x_1, \ldots, x_n\}$ the sample autocovariance satisfies $\sum_{|\tau|<n} \hat{\gamma}_\tau = 0$.

Solution 2.6

Let $\bar{X} = (\sum_{t=1}^n X_t)/n$. We compute:

$$\begin{aligned}\sum_{\tau=-n+1}^{n-1} \hat{\gamma}_\tau &= \frac{1}{n}\sum_{\tau=-n+1}^{n-1}\sum_{t=1}^{n-|\tau|}(X_t - \bar{X})(X_{t+|\tau|} - \bar{X}) \\ &= \frac{1}{n}\sum_{t_1=1}^{n}\sum_{t_2=1}^{n}(X_{t_1} - \bar{X})(X_{t_2} - \bar{X}) \\ &= \frac{1}{n}(\sum_{t=1}^{n}(X_t - \bar{X}))^2 \\ &= 0.\end{aligned}$$

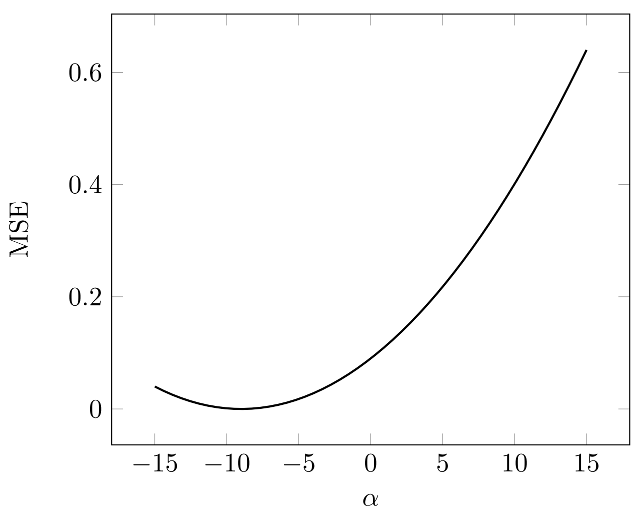

Figure 1: MSE of the autocovariance estimate for different values of α.