Midterm 1

Matthew Tiger

November 11, 2015

Problem 1.a. Consider the process

$$X_t + 0.4X_{t-1} - 0.32X_{t-2} = Z_t - 0.8Z_{t-1} + 0.16Z_{t-2}. \tag{1}$$

Determine whether the model is a stationary process.

Solution. The model $\{X_t\}$ is a stationary process if $\{X_t\}$ is a stationary solution of the equations (1). By the existence and uniqueness theorem of ARMA(p, q) processes, a stationary solution $\{X_t\}$ of the equations

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

that define the model exists if and only if

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \quad \text{for all } |z| = 1,$$

i.e. if and only if the roots of $\phi(z)$ do not lie on the unit circle.

For our model, we have $\phi_1 = -0.4$ and $\phi_2 = 0.32$ so that $\phi(z) = 1 + 0.4z - 0.32z^2$. Note that the roots of $\phi(z)$ are $z_1 = -1.25$ and $z_2 = 2.5$. As $|z_i| \neq 1$ for $i = 1, 2$, we conclude that the roots of $\phi(z)$ do not lie on the unit circle and that the model $\{X_t\}$ is a stationary process assuming that $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$. □

Problem 1.b. Considering the model in problem 1.a, what is R_3, i.e. the correlation matrix of size 3?

Solution. The covariance matrix of size 3 for our model $\{X_t\}$ is given by

$$\Gamma_3 = \begin{bmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{bmatrix}$$

where $\gamma(h)$ is the autocovariance function of the process $\{X_t\}$. For an ARMA(p, q) process $X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$, the autocovariance function $\gamma(h)$ satisfies the equations

$$\gamma(k) - \phi_1\gamma(k-1) - \cdots - \phi_p\gamma(k-p) = \sigma^2 \sum_{j=0}^{\infty} \theta_{k+j}\psi_j \quad \text{for } 0 \leq k < \max(p, q+1)$$

where $\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = \theta_j$ for $j \geq 0$ and $\psi_j = 0$ for $j < 0$. For our process, this corresponds to the system of equations

$$\begin{aligned} \gamma(0) - \phi_1\gamma(1) - \phi_2\gamma(2) &= \sigma^2(\psi_0 + \theta_1\psi_1 + \theta_2\psi_2) \\ \gamma(1) - \phi_1\gamma(0) - \phi_2\gamma(1) &= \sigma^2(\theta_1\psi_0 + \theta_2\psi_1) \\ \gamma(2) - \phi_1\gamma(1) - \phi_2\gamma(0) &= \sigma^2\theta_2\psi_0 \end{aligned} \tag{2}$$

where $\psi_0 = 1$, $\psi_1 = \theta_1 + \phi_1$, and $\psi_2 = \theta_2 + \phi_1^2 + \phi_1\theta_1 + \phi_2$. Using the parameters ϕ_j and θ_k defining our model, the system of equations (2) becomes

$$\begin{aligned} \gamma(0) + 0.4\gamma(1) - 0.32\gamma(2) &= 2.1136\sigma^2 \\ \gamma(1) + 0.4\gamma(0) - 0.32\gamma(1) &= -0.992\sigma^2 \\ \gamma(2) + 0.4\gamma(1) - 0.32\gamma(0) &= 0.16\sigma^2 \end{aligned}$$

the solution of which is $\gamma(0) = 5\sigma^2$, $\gamma(1) = -4.4\sigma^2$, and $\gamma(2) = 3.52\sigma^2$. Thus, the covariance matrix Γ_3 is given by

$$\Gamma_3 = \sigma^2 \begin{bmatrix} 5.00 & -4.40 & 3.52 \\ -4.40 & 5.00 & -4.40 \\ 3.52 & -4.40 & 5.00 \end{bmatrix}.$$

Note that the correlation matrix R_3 is given by $(1/\gamma(0))\Gamma_3$. Therefore,

$$R_3 = \begin{bmatrix} 1.000 & -0.880 & 0.704 \\ -0.880 & 1.000 & -0.880 \\ 0.704 & -0.880 & 1.000 \end{bmatrix}.$$

$\square$

Problem 1.c. Express the process in problem 1.a as a pure MA process in the form of $X_t = \sum_{j=0}^{\infty} \psi_j Z_t$.

Solution. For our process, the roots of the equation $\phi(z) = 1 + 0.4z - 0.32z^2 = 0$ are $z_1 = -1.25$ and $z_2 = 2.5$. As $|z_i| > 1$ for $i = 1, 2$, this process is causal and can be represented as an MA(∞) process, i.e. $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$, where the coefficients ψ_j are determined by the equations $\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = \theta_j$ for $j \geq 0$ and $\psi_j = 0$ for $j < 0$.

Note that for an ARMA(p, q) process, as $\theta_j = 0$ for $j > q$, the equations determining the coefficients are difference equations determined by the boundary conditions

$$\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = \theta_j \quad \text{for } 0 \leq j < \max(p, q+1)$$

and the homogeneous equation

$$\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = 0 \quad \text{for } j \geq \max(p, q+1).$$

For our process, the characteristic equation of these difference equations is $\phi(z)$. The roots of this characteristic equation are, as shown above, $z_1 = -1.25$ and $z_2 = 2.5$. As these roots are distinct, the solution to the homogeneous difference equation is

$$\psi_j = \alpha_1 z_1^{-j} + \alpha_2 z_2^{-j} = \alpha_1(-1.25)^{-j} + \alpha_2(2.5)^{-j} \quad \text{for } j \geq 1$$

where the coefficients are determined by the boundary conditions $\psi_0 = 1$, $\psi_1 = \theta_1 + \phi_1 = -1.2$, and $\psi_2 = \theta_2 + \phi_1^2 + \phi_1\theta_1 + \phi_2 = 0.96$. Using the method of undetermined coefficients, we can see that $\alpha_1 = 1.5$ and $\alpha_2 = 0$. Therefore $\psi_j = 1.5(-1.25)^{-j}$ for $j \geq 1$, $\psi_0 = 1$, and

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} = Z_t + 1.5\sum_{j=1}^{\infty}(-1.25)^{-j} Z_{t-j}.$$

□

Problem 2.a. Let X_t be the AR(2) process such that $X_t = 0.8X_{t-2} + Z_t$ where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$. Find the autocorrelation function of X_t.

Solution. This AR(2) process is defined by the parameters $\phi_1 = 0$ and $\phi_2 = 0.8$. This process has characteristic equation $\phi(z) = 1 - 0.8z^2 = 0$ of which the roots are $z_1 = 1.11803$ and $z_2 = -1.11803$. As these roots lie outside the unit circle this process is causal.

Note that $\{X_t\}$ can be represented as $(1 - \xi_1^{-1}B)(1 - \xi_2^{-1}B)X_t = Z_t$ where $0 = \phi_1 = \xi_1^{-1} + \xi_2^{-1}$ and $0.8 = \phi_2 = -\xi_1^{-1}\xi_2^{-1}$. Thus, $\xi_1^{-1} = -\frac{2}{\sqrt{5}}$ and $\xi_2^{-1} = \frac{2}{\sqrt{5}}$ so

$$X_t - 0.8X_{t-2} = \left(1 + \frac{2}{\sqrt{5}}B\right)\left(1 - \frac{2}{\sqrt{5}}B\right)X_t = Z_t.$$

The covariance function of this AR(2) process is given by

$$\gamma(h) = \frac{\sigma^2 \xi_1^2 \xi_2^2}{(\xi_1\xi_2 - 1)(\xi_2 - \xi_1)}\left[\frac{\xi_1^{1-|h|}}{\xi_1^2 - 1} - \frac{\xi_2^{1-|h|}}{\xi_2^2 - 1}\right].$$

Using $\xi_1 = -\frac{\sqrt{5}}{2}$ and $\xi_2 = \frac{\sqrt{5}}{2}$, we see that for our process,

$$\gamma(h) = \frac{5\sqrt{5}\sigma^2}{9}\left[\left(\frac{\sqrt{5}}{2}\right)^{1-|h|} - \left(\frac{-\sqrt{5}}{2}\right)^{1-|h|}\right].$$

As $\gamma(0) = \frac{25\sigma^2}{9}$, the autocorrelation function of this process is given by

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\sqrt{5}}{5}\left[\left(\frac{\sqrt{5}}{2}\right)^{1-|h|} - \left(\frac{-\sqrt{5}}{2}\right)^{1-|h|}\right].$$

□

Problem 2.b. Let X_t be the AR(2) process such that $X_t = 0.8X_{t-2} + Z_t$ where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$. Find the partial autocorrelation function of X_t.

Solution. The partial autocorrelation function $\alpha(h)$ is defined as $\alpha(0) = 1$, and for $h > 0$, $\alpha(h) = \phi_{hh}$ where ϕ_{hh} is the last component of

$$\phi_h = \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(h-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(h-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(h-1) & \gamma(h-2) & \dots & \gamma(0) \end{bmatrix}^{-1} \begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(h) \end{bmatrix}.$$

Note for an $AR(p)$ process that $\alpha(h) = 0$ if $h > p$ and $\alpha(p) = \phi_p$. So for our process, we need only determine $\alpha(1)$. From the above,

$$\alpha(1) = \frac{\gamma(1)}{\gamma(0)} = 0.$$

Therefore, for our AR(2) process, the partial autocorrelation function is

$$\alpha(h) = \begin{cases} 1 & \text{if } h = 0 \\ 0 & \text{if } |h| = 1 \\ 0.8 & \text{if } |h| = 2 \\ 0 & \text{if } |h| > 2 \end{cases}.$$

□

Problem 4.a. Let X_1, X_2, X_3, X_4, X_5 be observations from the MA(1) model. Find the best linear estimate of the missing value X_3.

Solution. If Y and $W_n, \dots, W_1$ are random variables, then for $\boldsymbol{W} = (W_n, \dots, W_1)^\intercal$ and $\boldsymbol{\mu}_W = (\mathrm{E}(W_n), \dots, \mathrm{E}(W_1))^\intercal$, the best linear predictor of Y in terms of $\{1, W_n, \dots, W_1\}$ is

$$P(Y|\boldsymbol{W}) = \mathrm{E}(Y) + \boldsymbol{a}^\intercal(\boldsymbol{W} - \boldsymbol{\mu_W})$$

where $\boldsymbol{a}$ is the solution of $\Gamma\boldsymbol{a} = \gamma$ for $\Gamma = \mathrm{Cov}\,(\boldsymbol{W}, \boldsymbol{W})$ and $\gamma = \mathrm{Cov}\,(Y, \boldsymbol{W})$. Also, note for an MA(1) process, the autocovariance function is defined as

$$\gamma_X(h) = \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0 \\ \sigma^2\theta & \text{if } |h| = 1 \\ 0 & \text{if } |h| > 1 \end{cases}$$

Using the above, set $Y = X_3$ and $W = (X_5, X_4, X_2, X_1)^\intercal$. Then

$$\begin{aligned} \Gamma = \mathrm{Cov}\,(\boldsymbol{W}, \boldsymbol{W}) &= \begin{bmatrix} \gamma_X(0) & \gamma_X(1) & \gamma_X(3) & \gamma_X(4) \\ \gamma_X(1) & \gamma_X(0) & \gamma_X(2) & \gamma_X(3) \\ \gamma_X(3) & \gamma_X(2) & \gamma_X(0) & \gamma_X(1) \\ \gamma_X(4) & \gamma_X(3) & \gamma_X(1) & \gamma_X(0) \end{bmatrix} \\ &= \sigma^2 \begin{bmatrix} 1+\theta^2 & \theta & 0 & 0 \\ \theta & 1+\theta^2 & 0 & 0 \\ 0 & 0 & 1+\theta^2 & \theta \\ 0 & 0 & \theta & 1+\theta^2 \end{bmatrix} \end{aligned}$$

and

$$\gamma = \begin{bmatrix} \gamma_X(2) \\ \gamma_X(1) \\ \gamma_X(1) \\ \gamma_X(2) \end{bmatrix} = \sigma^2 \begin{bmatrix} 0 \\ \theta \\ \theta \\ 0 \end{bmatrix}.$$

The solution to the system of equations $\Gamma \boldsymbol{a} = \gamma$ is

$$\boldsymbol{a} = \frac{\theta}{1+\theta^2+\theta^4} \begin{bmatrix} -\theta \\ 1+\theta^2 \\ 1+\theta^2 \\ -\theta \end{bmatrix}.$$

Therefore, the best predictor of X_3 is

$$\begin{aligned} P(X_3|\boldsymbol{W}) &= \mathrm{E}(X_3) + \boldsymbol{a}^\intercal(\boldsymbol{W} - \boldsymbol{\mu}_W) \\ &= \frac{\theta}{1+\theta^2+\theta^4}(-\theta X_5 + (1+\theta^2)X_4 + (1+\theta^2)X_2 - \theta X_1). \end{aligned}$$

□

Problem 4.b. Let X_1, X_2, X_3, X_4, X_5 be observations from the MA(1) model. Find the mean square error of the best linear estimate of the missing value X_3.

Solution. The mean squared error of the predictor in terms of the known random variables is $\mathrm{E}\left[(Y - P(Y|\boldsymbol{W}))^2\right] = \mathrm{Var}(Y) - \boldsymbol{a}^\intercal\gamma$ where Y, $\boldsymbol{W}$, $\boldsymbol{a}$, and γ are defined as in problem 4.a.

As $\mathrm{Var}(X_3) = \gamma_X(0) = \sigma^2(1+\theta^2)$, the mean squared error is given by

$$\mathrm{E}\left[(Y - P(Y|\boldsymbol{W}))^2\right] = \sigma^2(1+\theta^2) - \frac{2\sigma^2\theta^2(1+\theta^2)}{1+\theta^2+\theta^4}.$$

□