Homework Assignment 9

Matthew Tiger

October 29, 2015

Problem 2.20. In the innovations algorithm, show that for each $n \geq 2$, the innovation $X_n - \hat{X}_n$ is uncorrelated with $X_1, X_2, \ldots, X_{n-1}$. Conclude that $X_n - \hat{X}_n$ is uncorrelated with the innovations $X_1 - \hat{X}_1, X_2 - \hat{X}_2, \ldots, X_{n-1} - \hat{X}_{n-1}$

Solution. Note that if $n \geq 2$, then $\hat{X}_n = P_{n-1}X_n = a_0 + a_1 X_{n-1} + \cdots + a_{n-1}X_1$ where $a_1, \ldots, a_{n-1}$ is the solution to the system of equations

$$\begin{bmatrix} \gamma(0) & \gamma(1) & \ldots & \gamma(n-2) \\ \gamma(1) & \gamma(0) & \ldots & \gamma(n-3) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-2) & \gamma(n-3) & \ldots & \gamma(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n-1) \end{bmatrix}.$$

Now, for $0 < i < n$,

$$\begin{aligned} \operatorname{Cov}\left(X_n - \hat{X}_n, X_i\right) &= \operatorname{Cov}\left(X_n, X_i\right) - \operatorname{Cov}\left(\sum_{j=1}^{n-1} a_j X_{n-j}, X_i\right) \\ &= \operatorname{Cov}\left(X_n, X_i\right) - \sum_{j=1}^{n-1} a_j \operatorname{Cov}\left(X_{n-j}, X_i\right) \\ &= \gamma(n-i) - \sum_{j=1}^{n-1} a_j \gamma(n-j-i) \end{aligned} \tag{1}$$

From the above system of equations it is clear that $\sum_{j=1}^{n-1} a_j \gamma(n-j-i) = \gamma(n-i)$ and that $\operatorname{Cov}\left(X_n - \hat{X}_n, X_i\right) = 0$. Therefore, $X_n - \hat{X}_n$ is uncorrelated with X_i for $n \geq 2$ and $0 < i < n$.

Now, say $\hat{X}_i = b_0 + b_1 X_{n-1} + \cdots + b_{n-1}X_1$for $0 < i < n$. Then

$$\begin{aligned} \operatorname{Cov}\left(X_n - \hat{X}_n, X_i - \hat{X}_i\right) &= \operatorname{Cov}\left(X_n - \hat{X}_n, X_i\right) - \operatorname{Cov}\left(X_n - \hat{X}_n, \hat{X}_i\right) \\ &= -\operatorname{Cov}\left(X_n - \hat{X}_n, \sum_{j=1}^{i-1} b_j X_{i-j}\right) \\ &= -\sum_{j=1}^{i-1} b_j \operatorname{Cov}\left(X_n - \hat{X}_n, X_{i-j}\right) = 0 \end{aligned}$$

from our above results since $i - j < n$ for $0 < i < n$, $0 < j < i$, and $n \geq 2$. Therefore $X_n - \hat{X}_n$ is uncorrelated with $X_i - \hat{X}_i$ for $n \geq 2$ and $0 < i < n$.

$\square$

Problem 2.21. Let X_1, X_2, X_3, X_4, X_5 be observations from the MA(1) model.

$$X_t = Z_t + \theta Z_{t-1}, \{Z_t\} \sim \mathrm{WN}(0, \sigma^2).$$

a. Find the best linear estimate of the missing value X_3 in terms of X_1 and X_2.

b. Find the best linear estimate of the missing value X_3 in terms of X_4 and X_5.

c. Find the best linear estimate of the missing value X_3 in terms of X_1, X_2, X_4, and X_5.

d. Compute the mean squared errors for each of the estimates in (a), (b), and (c).

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

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

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

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

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

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

and

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

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

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

Therefore, the best predictor of X_3 is

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

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

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

and

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

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

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

Therefore, the best predictor of X_3 is

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

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

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

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

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

d. The mean squared error of the predictor in terms of the known random variables is $\mathrm{E}\left[(Y - P(Y|\boldsymbol{W}))^2\right] = \mathrm{Var}(Y) - \boldsymbol{a}^\intercal\gamma$.

Therefore, the mean squared error for:

(a) is $\mathrm{E}\left[(X_3 - P(X_3|\boldsymbol{W}))^2\right] = \sigma^2(1+\theta^2) - \frac{\sigma^2\theta^2(1+\theta^2)}{1+\theta^2+\theta^4}$

(b) is $\mathrm{E}\left[(X_3 - P(X_3|\boldsymbol{W}))^2\right] = \sigma^2(1+\theta^2) - \frac{\sigma^2\theta^2(1+\theta^2)}{1+\theta^2+\theta^4}$

(c) is $\mathrm{E}\left[(X_3 - P(X_3|\boldsymbol{W}))^2\right] = \sigma^2(1+\theta^2) - \frac{2\sigma^2\theta^2(1+\theta^2)}{1+\theta^2+\theta^4}$

□

Problem 2.22. Repeat parts (a)-(d) of Problem 2.21 assuming now that the observations X_1, X_2, X_3, X_4, X_5 are from the causal AR(1) model

$$X_t = \phi X_{t-1} + Z_t, \{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$$

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

$$\gamma_X(h) = \frac{\sigma^2\phi^{|h|}}{1-\phi^2}$$

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

$$\Gamma = \mathrm{Cov}(\boldsymbol{W}, \boldsymbol{W}) = \begin{bmatrix} \gamma_X(0) & \gamma_X(1) \\ \gamma_X(1) & \gamma_X(0) \end{bmatrix} = \frac{\sigma^2}{1-\phi^2}\begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix}$$

and

$$\gamma = \begin{bmatrix} \gamma_X(1) \\ \gamma_X(2) \end{bmatrix} = \frac{\sigma^2}{1-\phi^2}\begin{bmatrix} \phi \\ \phi^2 \end{bmatrix}.$$

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

$$\boldsymbol{a} = \begin{bmatrix} \phi \\ 0 \end{bmatrix}.$$

Therefore, the best predictor of X_3 is

$$P(X_3|\boldsymbol{W}) = \mathrm{E}(X_3) + \boldsymbol{a}^\intercal(\boldsymbol{W} - \boldsymbol{\mu}_W) = \phi X_2$$

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

$$\Gamma = \operatorname{Cov}(\boldsymbol{W}, \boldsymbol{W}) = \begin{bmatrix} \gamma_X(0) & \gamma_X(1) \\ \gamma_X(1) & \gamma_X(0) \end{bmatrix} = \frac{\sigma^2}{1-\phi^2} \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix}$$

and

$$\gamma = \begin{bmatrix} \gamma_X(2) \\ \gamma_X(1) \end{bmatrix} = \frac{\sigma^2}{1-\phi^2} \begin{bmatrix} \phi^2 \\ \phi \end{bmatrix}.$$

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

$$\boldsymbol{a} = \begin{bmatrix} 0 \\ \phi \end{bmatrix}.$$

Therefore, the best predictor of X_3 is

$$P(X_3|\boldsymbol{W}) = \mathrm{E}(X_3) + \boldsymbol{a}^\intercal(\boldsymbol{W} - \boldsymbol{\mu}_W) = \phi X_4$$

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

$$\begin{aligned}\Gamma = \operatorname{Cov}(\boldsymbol{W}, \boldsymbol{W}) = & \begin{bmatrix} \gamma_X(0) & \gamma_X(1) & \gamma_X(3) & \gamma_X(4) \\ \gamma_X(1) & \gamma_X(0) & \gamma_X(2) & \gamma_X(3) \\ \gamma_X(3) & \gamma_X(2) & \gamma_X(0) & \gamma_X(1) \\ \gamma_X(4) & \gamma_X(3) & \gamma_X(1) & \gamma_X(0) \end{bmatrix} \\ = \frac{\sigma^2}{1-\phi^2} & \begin{bmatrix} 1 & \phi & \phi^3 & \phi^4 \\ \phi & 1 & \phi^2 & \phi^3 \\ \phi^3 & \phi^2 & 1 & \phi \\ \phi^4 & \phi^3 & \phi & 1 \end{bmatrix}\end{aligned}$$

and

$$\gamma = \begin{bmatrix} \gamma_X(2) \\ \gamma_X(1) \\ \gamma_X(1) \\ \gamma_X(2) \end{bmatrix} = \frac{\sigma^2}{1-\phi^2} \begin{bmatrix} \phi^2 \\ \phi \\ \phi \\ \phi^2 \end{bmatrix}.$$

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

$$\boldsymbol{a} = \phi \begin{bmatrix} 0 \\ \frac{1-\phi^2}{1-\phi^4} \\ \frac{1-\phi^2}{1-\phi^4} \\ 0 \end{bmatrix}.$$

Therefore, the best predictor of X_3 is

$$\begin{aligned} P(X_3|\boldsymbol{W}) &= \mathrm{E}(X_3) + \boldsymbol{a}^\intercal(\boldsymbol{W} - \boldsymbol{\mu}_W) \\ &= \frac{\phi - \phi^3}{1-\phi^4}(X_4 + X_2) \end{aligned}$$

d. The mean squared error of the predictor in terms of the known random variables is $\mathrm{E}\left[(Y - P(Y|\boldsymbol{W}))^2\right] = \mathrm{Var}(Y) - \boldsymbol{a}^\intercal \gamma$.

Therefore, the mean squared error for:

(a) is $\mathrm{E}\left[(X_3 - P(X_3|\boldsymbol{W}))^2\right] = \frac{\sigma^2}{1-\phi^2} - \frac{\sigma^2\phi^2}{1-\phi^2} = \sigma^2$

(b) is $\mathrm{E}\left[(X_3 - P(X_3|\boldsymbol{W}))^2\right] = \frac{\sigma^2}{1-\phi^2} - \frac{\sigma^2\phi^2}{1-\phi^2} = \sigma^2$

(c) is $\mathrm{E}\left[(X_3 - P(X_3|\boldsymbol{W}))^2\right] = \frac{\sigma^2}{1-\phi^2} - \frac{2\sigma^2\phi^2(1-\phi^2)}{(1-\phi^2)(1-\phi^4)} = \frac{\sigma^2(1-\phi^2)}{1-\phi^4}$

$\square$