

Homework Assignment 9

Matthew Tiger

November 5, 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, \dots, X_{n-1} . Conclude that $X_n - \hat{X}_n$ is uncorrelated with the innovations $X_1 - \hat{X}_1, X_2 - \hat{X}_2, \dots, X_{n-1} - \hat{X}_{n-1}$

Solution. Note that if $n \geq 2$, then $\hat{X}_n = P_{n-1}X_n = a_0 + a_1X_{n-1} + \dots + a_{n-1}X_1$ where a_1, \dots, a_{n-1} is the solution to the system of equations

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

Now, for $0 < i < n$,

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

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 $\text{Cov}(X_n - \hat{X}_n, X_i) = 0$. Therefore, $X_n - \hat{X}_n$ is uncorrelated with X_i for $n \geq 2$ and $0 < i < n$.

Now, say $\hat{X}_i = b_0 + b_1X_{i-1} + \dots + b_{i-1}X_1$ for $0 < i < n$. Then

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

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 $0 < i < n$. □

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

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

- Find the best linear estimate of the missing value X_3 in terms of X_1 and X_2 .
- Find the best linear estimate of the missing value X_3 in terms of X_4 and X_5 .
- Find the best linear estimate of the missing value X_3 in terms of X_1, X_2, X_4 , and X_5 .
- Compute the mean squared errors for each of the estimates in (a), (b), and (c).

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

$$P(Y|\mathbf{W}) = E(Y) + \mathbf{a}^\top(\mathbf{W} - \boldsymbol{\mu}_W)$$

where \mathbf{a} is the solution of $\Gamma \mathbf{a} = \gamma$ for $\Gamma = \text{Cov}(\mathbf{W}, \mathbf{W})$ and $\gamma = \text{Cov}(Y, \mathbf{W})$.

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 $W = (X_2, X_1)^\top$. Then

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

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 $\Gamma \mathbf{a} = \gamma$ is

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

Therefore, the best predictor of X_3 is

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

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

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

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 $\Gamma \mathbf{a} = \gamma$ is

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

Therefore, the best predictor of X_3 is

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

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

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

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 $\Gamma \mathbf{a} = \gamma$ is

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

Therefore, the best predictor of X_3 is

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

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

Therefore, the mean squared error for:

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

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

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

□

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

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

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

$$P(Y|\mathbf{W}) = E(Y) + \mathbf{a}^\top (\mathbf{W} - \boldsymbol{\mu}_W)$$

where \mathbf{a} is the solution of $\Gamma \mathbf{a} = \gamma$ for $\Gamma = \text{Cov}(\mathbf{W}, \mathbf{W})$ and $\gamma = \text{Cov}(Y, \mathbf{W})$.

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 $W = (X_2, X_1)^\top$. Then

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

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 $\Gamma \mathbf{a} = \gamma$ is

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

Therefore, the best predictor of X_3 is

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

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

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

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 $\Gamma \mathbf{a} = \gamma$ is

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

Therefore, the best predictor of X_3 is

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

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

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

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 $\Gamma \mathbf{a} = \gamma$ is

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

Therefore, the best predictor of X_3 is

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

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

Therefore, the mean squared error for:

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

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

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

□