On Linear Innovations Sequence

Yuchao Li

August 03, 2017

1 PROBLEM STATEMENT

Suppose X, Y_1, Y_2,..., Y_n are random vectors and 1, Y_1, Y_2,..., Y_n have finite second moments but not orthogonal. According to [1], by orthogonalizing this sequences we can obtain a new sequence 1. $\widetilde{Y}_1$, $\widetilde{Y}_2$,..., $\widetilde{Y}_n$ used for linear estimating of X. Let $\widetilde{Y}_1 = Y_1 - E[Y_1]$, and for $i \geq 2$, let

$$\widetilde{Y}_i = Y_i - \widehat{E}[Y_i|Y_1, Y_2, \ldots, Y_{i-1}] \tag{1.1}$$

It is claimed in the following elaboration (online version of [1], page 95 and 96) that

1. $E[\widetilde{Y}_i] = 0$ for all i and $E[\widetilde{Y}_i \widetilde{Y}_j^T] = 0$ for all $i \neq j$.
2. $\widehat{E}[X|Y_1, Y_2, \ldots, Y_n] = \widehat{E}[X|\widetilde{Y}_1, \widetilde{Y}_2, \ldots, \widetilde{Y}_n]$.

2 ELABORATION

First, prove $E[\widetilde{Y}_i] = 0$ for all i. $E[\widetilde{Y}_1] = E[Y_1] - E[E[Y_1]] = 0$. Denote $Z_i^T = [Y_1^T, Y_2^T, \ldots, Y_{i-1}^T]$. For $i \geq 2$,

$$\widehat{E}[Y_i|Y_1, Y_2, \ldots, Y_{i-1}] = E[Y_i] + \text{Cov}[Y_i, Z_i]\text{Cov}[Z_i]^{-1}(Z_i - E[Z_i]). \tag{2.1}$$

Apparently, $E[\widehat{E}[Y_i|Y_1, Y_2, \ldots, Y_{i-1}]] = E[Y_i]$. Therefore, $E[\widetilde{Y}_i] = 0$ for $i \geq 2$ also holds. Then prove $E[\widetilde{Y}_i \widetilde{Y}_j^T] = 0$ for all $i \neq j$. Without loss of generality, suppose that $i > j$. According to [1] page 86, $\widetilde{Y}_i = e_i$ and is orthogonal to all linear combination of $1, Y_1, Y_2, \ldots, Y_{i-1}$. $\widetilde{Y}_j$ is a linear combination of $1, Y_1, Y_2, \ldots, Y_j$. Since $i > j$, $\widetilde{Y}_j$ is also a linear combination of $1, Y_1, Y_2, \ldots, Y_{i-1}$. Therefore $E[\widetilde{Y}_i \widetilde{Y}_j^T] = 0$ for all $i \neq j$.

Second, according to Eq. (1.1), $\widetilde{Y}_i$ is a linear function of 1, Y_1, $Y_2, \ldots, Y_i$. Therefore we can claim that

$$\widetilde{Y} = BY + c \tag{2.2}$$

where matrix B is lower triangular matrix with all diagonal elements to be 1. Therefore B is invertible. By [1],

$$\widehat{E}[X|Y_1, Y_2, \ldots, Y_n] = E[X] + \mathrm{Cov}[X, Y]\mathrm{Cov}[Y]^{-1}(Y - E[Y]) \tag{2.3}$$

and similarly

$$\widehat{E}[X|\widetilde{Y}_1, \widetilde{Y}_2, \ldots, \widetilde{Y}_n] = E[X] + \mathrm{Cov}[X, \widetilde{Y}]\mathrm{Cov}[\widetilde{Y}]^{-1}(\widetilde{Y} - E[\widetilde{Y}]) \tag{2.4}$$

Since $\mathrm{Cov}[\widetilde{Y}] = B\mathrm{Cov}[Y]B^T$ which leads to

$$\mathrm{Cov}[\widetilde{Y}]^{-1} = (B^T)^{-1}\mathrm{Cov}[Y]^{-1}B^{-1}$$

and

$$\mathrm{Cov}[X, \widetilde{Y}] = \mathrm{Cov}[X, BY + c] = \mathrm{Cov}[X, Y]B^T,$$

Eq. (2.4) can be reformulated as

$$\widehat{E}[X|\widetilde{Y}_1, \widetilde{Y}_2, \ldots, \widetilde{Y}_n] = E[X] + \mathrm{Cov}[X, Y]B^T(B^T)^{-1}\mathrm{Cov}[Y]^{-1}B^{-1}B(Y - E[Y]) = \widehat{E}[X|Y_1, Y_2, \ldots, Y_n] \tag{2.5}$$

References

[1] Bruce Hajek, *Random processes for engineers*, Cambridge University Press, 2015.