On the Asympotic Mean Square Error of Multi-step Ahead Prediction for the Bilinear Time Series Model

Yamada, Kazutoshi

https://doi.org/10.15017/2920672

出版情報：経済論究. 63, pp.149-159, 1985-12-25. 九州大学大学院経済学会

On the Asymptotic Mean Square Error of Multi-step Ahead Prediction for the Bilinear Time Series Model

By Kazutoshi Yamada

1. Introduction

In this paper we shall consider the asymptotic mean square error (A.M.S.E.) of multi-step ahead prediction for the bilinear time series $BL(p, 0, p, 1)$ model, defined by (1.1) below, with estimated parameters, and, in the same way as in Yamamoto[11] and Baillie[1], we shall derive an explicit formula for the asymtotic mean square multi-step ahead prediction error of the $BL(p, 0, p, 1)$ model. The autoregressive model ($AR(p)$), which has been considered in Yamamoto[11] and Baillie[1], is linear time series model. On the contrary, the $BL(p, 0, p, 1)$ model is an non-linear time series model and includes the $AR(p)$ process as a special case. In this sense, the $BL(p, 0, p, 1)$ model which is the object of our consideration may be somewhat a generalized model.

Box and Jenkins[3] have derived the A.M.S.E. of multi-step ahead prediction for the univariate $AR(1)$ process. As for the univariate autoregressive-moving average model ($ARMA(p, q)$), Bloomfield[2] has derived the A.M.S.E. of one-step ahead prediction, but his approach is too intractable to apply to a more general model. Applicable approach has been developed by Yamamoto[11], for the univariate $AR(p)$ process. This approach using a theorem by Neudecker[7] and the Taylor expansion of the predictor is simple and tractable. In fact, by directly applying the approach of Yamamoto, Baillie[1] has successfully derived the A.M.S.E. of predicting more than one-step ahead for the general multivariate $AR(p)$ process. For the multivariate $AR(p)$ process, Fotopoulos and Ray[5] and Reinsel[8], have recently offered alternative approaches and respectively derived the same result as Baillie's. Yamamoto[12] has also succeeded in deriving the A.M.S.E. of multi-step ahead prediction for the general multivariate $ARMA(p, q)$ process.

Adopting the notations in the manner of Subba Rao[9], the bilinear time series model $BL(p, r, m, k)$ is defined by the following difference equation:

$$x_t + \sum_{j=1}^{p} a_j x_{t-j} = \sum_{j=0}^{r} c_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} x_{t-i} e_{t-j}, \quad c_0 = 1, \tag{1.1}$$

where $\{e_t\}$ is an independent white noise process. In this paper, we restrict ourselves to the $BL(p, 0, p, 1)$ model, and further suppose that each e_t is distributed $N(0, \sigma^2)$, where σ^2 is a known constant variance. In section 2, the $BL(p, 0, p, 1)$ model is represented in the state space and some main assumptions are introduced. Section 3 presents the multi-step ahead prediction of the $BL(p, 0, p, 1)$ model with the estimated parameters. Finally, in section 4, an explicit expression of the A.M.S.E. for the multi-step ahead prediction is derived.

2. The state space representation of the $BL(p, 0, p, 1)$ model and some assumptions

We shall consider the $BL(p, 0, p, 1)$ model

$$x_t + \sum_{j=1}^{p} a_j x_{t-j} = e_t + \sum_{j=1}^{p} b_j x_{t-j} e_{t-1}. \tag{2.1}$$

It will be convienient to express the model (2.1) in the state space form, therefore we will introduce the matrices and the vectors as follows:

$$A=\begin{bmatrix}-a_1-a_3\cdots-a_p\\ I_{p-1}\qquad O\end{bmatrix},\qquad B=\begin{bmatrix}b_1\ b_2\ \cdots\ b_p\\ O\end{bmatrix},\tag{2.2}$$

$$H'=(1,0,\dots,0),\ C'=(1,0,\dots\ 0),\ \boldsymbol{x}'_t=(x_t,x_{t-1},\dots,x_{t-p+1}),$$

where I_{p-1} is the identity matrix of rank p-1. Thus the state space representation of the model (2.1) is given by

$$\boldsymbol{x}_t=A\boldsymbol{x}_{t-1}+B\boldsymbol{x}_{t-1}e_{t-1}+Ce_t,\tag{2.3}$$

and we can get x_t by multiplying $\boldsymbol{x}_t$ by H', i. e. $x_t=H'\boldsymbol{x}_t$.

We shall here assume the following three conditions to be satisfied. They are

(i) $\rho(A)<1$,

(ii) $\rho(A\otimes A+B\otimes B\,\sigma^2)<1$,

(iii) $H'BE(\boldsymbol{x}_t\boldsymbol{x}'_t)B'H<(H'C)^2$,

where $\rho(A)=\max_i[|\lambda_i(A)|]$, $\lambda_i(A)$ is the ith eigenvalue of A, and $\otimes$ denotes the Kronecker product. As shown by Subba Rao and Gabr[10], condition (i) is the sufficient condition for asymptotic first-order stationary; condition (ii) is the sufficient condition for asymptotic second-order stationary; and condition (iii) is the sufficient condition for invertibility of the $BL(p,0,p,1)$ model. Hence, the condition (i) implies that the mean value of the time series process $\{\boldsymbol{x}_t\}$ satisfying (2.1) is asymptopically independent from time t, and the condition (ii) means that variance-covariance of $\boldsymbol{x}_t$ depend on only difference in time. The model (2.1) said to be invertible if $\lim_{t\to\infty} E\ \{e_t-\hat{e}_t\}^2=0$, where $\hat{e}_t$ is an estimate of e_t (see Granger and Andersen[6]). Thus the condition (iii) ensures that an estimate of e_t obtained from the model (2.1) tends to true value as $t\to\infty$.

For the convenience in the following sections, let us consider some expectations of $\boldsymbol{x}_t$ and e_t under the above conditions, and let

$$m\equiv E(\boldsymbol{x}_t),\ \Gamma\equiv E(\boldsymbol{x}_t\boldsymbol{x}'_t),\ S\equiv E(\boldsymbol{x}_t\boldsymbol{x}'_te_t),\ W\equiv E(\boldsymbol{x}_t\boldsymbol{x}'_te_t^2).\tag{2.4}$$

Noting $E(\boldsymbol{x}_te_t)=\sigma^2C$ and $E(e_t^4)=3\sigma^4$ (e. g. see Cramér[4], chap. 17);

we obtain from (2.3)

$$\boldsymbol{S}=\sigma^2\boldsymbol{C}\boldsymbol{m}'\boldsymbol{A}'+\sigma^2\boldsymbol{A}\boldsymbol{m}\boldsymbol{C}'+\sigma^4\boldsymbol{B}\boldsymbol{C}\boldsymbol{C}'+\sigma^4\boldsymbol{C}\boldsymbol{C}'\boldsymbol{B}', \tag{2.5}$$

$$\begin{aligned}\boldsymbol{W}&=\sigma^2\boldsymbol{A}\boldsymbol{\Gamma}\boldsymbol{A}'+\sigma^2\boldsymbol{A}\boldsymbol{S}\boldsymbol{B}'+\sigma^2\boldsymbol{B}\boldsymbol{S}\boldsymbol{A}'+\sigma^2\boldsymbol{B}\boldsymbol{W}\boldsymbol{B}'+3\sigma^4\boldsymbol{C}\boldsymbol{C}'\\&=\sigma^2\boldsymbol{\Gamma}+2\sigma^4\boldsymbol{C}\boldsymbol{C}'.\end{aligned} \tag{2.6}$$

3. The multi-step ahead prediction

We are concerned with forecasting a future value x_{n+h}, $h>0$, based on the given n observations $x_1, x_2, \dots, x_n$, at time n. It is well known that the best of such predictions is the conditional mean of x_{n+h} given all information of x's up to time n. From (2.3), x_{n+h} can be written as

$$x_{n+h}=\begin{cases}\boldsymbol{H}'(\boldsymbol{A}+\boldsymbol{B}e_n)\boldsymbol{x}_n+\boldsymbol{H}'\boldsymbol{C}e_{n+1} & \text{for } h=1,\\ \boldsymbol{H}'\left\{\prod_{j=1}^{h}(\boldsymbol{A}+\boldsymbol{B}e_{n+h-j})\right\}\boldsymbol{x}_n+\boldsymbol{H}'\boldsymbol{C}e_{n+h} & \\ \qquad +\sum_{i=1}^{h-1}\boldsymbol{H}'\left\{\prod_{j=1}^{i}(\boldsymbol{A}+\boldsymbol{B}e_{n+h-j})\right\}\boldsymbol{C}e_{n+h-i} & \text{for } h\geq 2,\end{cases} \tag{3.1}$$

where $n>p$. Hence, denoting by $x_{n,h}$ the conditional expectation of x_{n+h} at time n, from (3.1) we have

$$\begin{aligned}x_{n,h}&=E(x_{n+h}\,|\,x_1, x_2, \dots, x_n, e_n)\\&=\begin{cases}\boldsymbol{H}'\boldsymbol{A}\boldsymbol{x}_n+\boldsymbol{H}'\boldsymbol{B}\boldsymbol{x}_n e_n & \text{for } h=1,\\ \boldsymbol{H}'\boldsymbol{A}^h\boldsymbol{x}_n+\boldsymbol{H}'\boldsymbol{A}^{h-1}\boldsymbol{B}\boldsymbol{x}_n e_n+\sigma^2\boldsymbol{H}'\sum_{j=0}^{h-2}\boldsymbol{A}^j\boldsymbol{B}\boldsymbol{C} & \text{for } h\geq 2.\end{cases}\end{aligned} \tag{3.2}$$

$x_{n,h}$ is the multi-step ahead predictor when all parameters are known. However, the parameters are generally not given, therefore, they must be inferred. Fortunately, we can estimate the parameters of the model (2.3) by the maximum likelihood estimation method (see Subba Rao [9]), and let us denote the estimators of $\boldsymbol{A}$ and $\boldsymbol{B}$ as $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$. Thus, an estimate of $x_{n,h}$, say $\hat{x}_{n,h}$, are obtained by substituting $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ in place of $\boldsymbol{A}$ and $\boldsymbol{B}$ in (3.2), i.e.

$$\hat{x}_{n,h}=\begin{cases}\boldsymbol{H}'\hat{\boldsymbol{A}}\boldsymbol{x}_n+\boldsymbol{H}'\hat{\boldsymbol{B}}\boldsymbol{x}_n e_n & \text{for } h=1,\\ \boldsymbol{H}'\hat{\boldsymbol{A}}^h\boldsymbol{x}_n+\boldsymbol{H}'\hat{\boldsymbol{A}}^{h-1}\hat{\boldsymbol{B}}\boldsymbol{x}_n e_n+\sigma^2\boldsymbol{H}'\sum_{j=0}^{h-2}\hat{\boldsymbol{A}}^j\hat{\boldsymbol{B}}\boldsymbol{C} & \text{for } h\geq 2.\end{cases} \tag{3.3}$$

Here, note that σ^2 is assumed as a known parameter for simplicity, but this assumption is not substantial.

4. The A. M. S. E. of the multi-step ahead prediction

In this section, we will derive an explicit formula of the A.M.S.E. of the multi-step ahead prediction. We will begin by presenting the following important result, which was proved by using a theorem of Neudecker[7], and this result will be used of our advantage in our problem.

$$\boldsymbol{M}(h)\equiv(\partial\boldsymbol{H}'\boldsymbol{A}^h/\partial\boldsymbol{a})=\boldsymbol{L}_p\left\{\sum_{s=0}^{h-1}(\boldsymbol{A}^{s\prime}\otimes\boldsymbol{A}^{h-1-s})\right\}\boldsymbol{L}_p', \tag{4.1}$$

where $\boldsymbol{L}_p\equiv(\boldsymbol{I}_p, \boldsymbol{O}, \ldots, \boldsymbol{O}): p\times p^2$, $\boldsymbol{a}\equiv(-a_1, -a_2, \ldots, -a_p)'$. The proof of (4.1) is found in Yamamoto[11].

Let

$$\boldsymbol{V}(\hat{\boldsymbol{a}})\equiv E\{(\hat{\boldsymbol{a}}-\boldsymbol{a})(\hat{\boldsymbol{a}}-\boldsymbol{a})'\},\ \boldsymbol{V}(\hat{\boldsymbol{b}})\equiv E\{(\hat{\boldsymbol{b}}-\boldsymbol{b})(\hat{\boldsymbol{b}}-\boldsymbol{b})'\},$$
$$Cov(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}})\equiv E\{(\hat{\boldsymbol{a}}-\boldsymbol{a})(\hat{\boldsymbol{b}}-\boldsymbol{b})'\}, \tag{4.2}$$

where $\hat{\boldsymbol{a}}\equiv(-\hat{a}_1, -\hat{a}_2, \ldots, -\hat{a}_p)'$, $\hat{\boldsymbol{b}}\equiv(\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p)'$, $\boldsymbol{b}\equiv(b_1, b_2, \ldots, b_p)'$. By deriving the information matrix from the second differentiation of the likelihood, the variance covariance matrix of $(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}})$ is obtained (e. g. see Subba Rao[9]; Box and Jenkins[3], etc.). By the Taylor expansion of $\boldsymbol{H}'\hat{\boldsymbol{A}}^h\boldsymbol{x}_n$ about $\hat{\boldsymbol{a}}=\boldsymbol{a}$, we have

$$\boldsymbol{H}'\hat{\boldsymbol{A}}^h\boldsymbol{x}_n\simeq\boldsymbol{H}'\boldsymbol{A}^h\boldsymbol{x}_n+(\hat{\boldsymbol{a}}-\boldsymbol{a})'\boldsymbol{M}(h)\boldsymbol{x}_n. \tag{4.3}$$

Similarly, we have also

$$\boldsymbol{H}'\hat{\boldsymbol{A}}^{h-1}\hat{\boldsymbol{B}}\boldsymbol{x}_n e_n\simeq\boldsymbol{H}'\boldsymbol{A}^{h-1}\boldsymbol{B}\boldsymbol{x}_n e_n+(\hat{\boldsymbol{a}}-\boldsymbol{a})'\boldsymbol{M}(h-1)\boldsymbol{B}\boldsymbol{x}_n e_n$$
$$+\boldsymbol{H}'\boldsymbol{A}^{h-1}\boldsymbol{H}(\hat{\boldsymbol{b}}-\boldsymbol{b})'\boldsymbol{x}_n e_n, \tag{4.4}$$

$$\boldsymbol{H}'\hat{\boldsymbol{A}}^j\hat{\boldsymbol{B}}\boldsymbol{C}\simeq\boldsymbol{H}'\boldsymbol{A}^j\boldsymbol{B}\boldsymbol{C}+(\hat{\boldsymbol{a}}-\boldsymbol{a})'\boldsymbol{M}(j)\boldsymbol{B}\boldsymbol{C}+\boldsymbol{H}'\boldsymbol{A}^j\boldsymbol{H}(\hat{\boldsymbol{b}}-\boldsymbol{b})'\boldsymbol{C}. \tag{4.5}$$

Here, we shall consider a particular case $h=1$, i. e., the A.M.S.E. of the one-step ahead prediction. Now there is a need to introduce an assumption that is often made in the A.M.S.E. of the prediction (e.g. Yamamoto[11]; Reinsel[8]), and the assumption is that $\hat{\boldsymbol{a}}-\boldsymbol{a}$ and $\hat{\boldsymbol{b}}-\boldsymbol{b}$ are independent of $\boldsymbol{x}_n$. From (3.2), (3.3), (4.3), (4.4) and the above assumption, the mean squared prediction error due to the use of the estimated parameters is approximately as follows:

$$
\begin{aligned}
E(\hat{x}_{n,1}-x_{n,1})^2 &= E(\boldsymbol{H}'\hat{\boldsymbol{A}}\boldsymbol{x}_n+\boldsymbol{H}'\hat{\boldsymbol{B}}\boldsymbol{x}_n e_n-\boldsymbol{H}'\boldsymbol{A}\boldsymbol{x}_n-\boldsymbol{H}'\boldsymbol{B}\boldsymbol{x}_n e_n)^2\\
&\simeq E((\hat{\boldsymbol{a}}-\boldsymbol{a})'\boldsymbol{x}_n+(\hat{\boldsymbol{b}}-\boldsymbol{b})'\boldsymbol{x}_n e_n)^2\\
&=\mathrm{tr}[\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{\Gamma}]+\mathrm{tr}[\boldsymbol{V}(\hat{\boldsymbol{b}})\boldsymbol{W}]+2\mathrm{tr}[Cov(\hat{\boldsymbol{a}},\hat{\boldsymbol{b}})\boldsymbol{S}]. \qquad (4.6)
\end{aligned}
$$

On the other hand, the mean squared prediction error due to the random innovations is from (3.1) and (3.2)

$$E(x_{n,1}-x_{n+1})^2=\sigma^2. \qquad (4.7)$$

Thus, from (4.6) and (4.7), an explicit formula of the A.M.S.E. of the one-step ahead prediction for the $BL(p, 0, p, 1)$ model is given by

$$\text{A.M.S.E.}(\hat{x}_{n,1})\sim\sigma^2+\mathrm{tr}[\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{\Gamma}]+\mathrm{tr}[\boldsymbol{V}(\hat{\boldsymbol{b}})\boldsymbol{W}]+2\mathrm{tr}[Cov(\hat{\boldsymbol{a}},\hat{\boldsymbol{b}})\boldsymbol{S}]. \qquad (4.8)$$

We will proceed to deduce the formula for $h\geq 2$. The formula can be obtained in the same way as the above case $h=1$, though somewhat complicatedly. What corresponds to (4.6) in case $h\geq 2$ is as follows:

$$
\begin{aligned}
\eta_1 &\equiv E(\hat{x}_{n,h}-x_{n,h})^2\\
&\simeq E\Big\{(\hat{\boldsymbol{a}}-\boldsymbol{a})'\boldsymbol{M}(h)\boldsymbol{x}_n+(\hat{\boldsymbol{a}}-\boldsymbol{a})'\boldsymbol{M}(h-1)\boldsymbol{B}\boldsymbol{x}_n e_n+\boldsymbol{H}'\boldsymbol{A}^{h-1}\boldsymbol{H}(\hat{\boldsymbol{b}}-\boldsymbol{b})'\boldsymbol{x}_n e_n\\
&\qquad +\sigma^2(\hat{\boldsymbol{a}}-\boldsymbol{a})'\sum_{j=0}^{h-2}\boldsymbol{M}(j)\boldsymbol{B}\boldsymbol{C}+\sigma^2\boldsymbol{H}'\sum_{j=0}^{h-2}\boldsymbol{A}^j\boldsymbol{H}(\hat{\boldsymbol{b}}-\boldsymbol{b})'\boldsymbol{C}\Big\}^2\\
&=\mathrm{tr}[\boldsymbol{M}'(h)\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{M}(h)\boldsymbol{\Gamma}]+\mathrm{tr}[\boldsymbol{B}'\boldsymbol{M}'(h-1)\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{M}(h-1)\boldsymbol{B}\boldsymbol{W}]\\
&\qquad +(a^*_{h-1})^2\mathrm{tr}[\boldsymbol{V}(\hat{\boldsymbol{b}})\boldsymbol{W}]+\sigma^4\boldsymbol{C}'\boldsymbol{B}'\Big\{\sum_{i=0}^{h-2}\sum_{j=0}^{h-2}\boldsymbol{M}'(i)\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{M}(j)\Big\}\boldsymbol{B}\boldsymbol{C}\\
&\qquad +\sigma^4\Big(\sum_{j=0}^{h-2}a^*_j\Big)^2\boldsymbol{C}'\boldsymbol{V}(\hat{\boldsymbol{b}})\boldsymbol{C}+2\mathrm{tr}[\boldsymbol{M}'(h)\boldsymbol{V}(\hat{\boldsymbol{b}})\boldsymbol{M}(h-1)\boldsymbol{B}\boldsymbol{S}]\\
&\qquad +2\sigma^2\boldsymbol{m}'\boldsymbol{M}'(h)\Big\{\boldsymbol{V}(\hat{\boldsymbol{a}})\sum_{j=0}^{h-2}\boldsymbol{M}(j)\boldsymbol{B}\boldsymbol{C}+Cov(\hat{\boldsymbol{a}},\hat{\boldsymbol{b}})\boldsymbol{C}\sum_{j=0}^{h-2}a^*_j\Big\}\\
&\qquad +2a^*_{h-1}\Big\{\mathrm{tr}[\boldsymbol{M}'(h)Cov(\hat{\boldsymbol{a}},\hat{\boldsymbol{b}})\boldsymbol{S}]+\mathrm{tr}[\boldsymbol{B}'\boldsymbol{M}'(h-1)Cov(\hat{\boldsymbol{a}},\hat{\boldsymbol{b}})\boldsymbol{W}]\\
&\qquad +\sigma^4\boldsymbol{C}'\boldsymbol{B}'\sum_{j=0}^{h-2}\boldsymbol{M}'(j)Cov(\hat{\boldsymbol{a}},\hat{\boldsymbol{b}})\boldsymbol{C}+\sigma^4\boldsymbol{C}'\boldsymbol{V}(\hat{\boldsymbol{b}})\boldsymbol{C}\sum_{j=0}^{h-2}a^*_j\Big\}
\end{aligned}
$$

$$+2\sigma^4 \boldsymbol{C}' \boldsymbol{B}' \left\{ \boldsymbol{M}'(h-1) Cov(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}) \boldsymbol{C} \sum_{j=0}^{h-2} a_j^* + \boldsymbol{M}'(h-1) \boldsymbol{V}(\hat{\boldsymbol{a}}) \sum_{j=0}^{h-2} \boldsymbol{M}(j) \boldsymbol{BC} + \sum_{j=0}^{h-2} \boldsymbol{M}'(j) Cov(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}) \boldsymbol{C} \sum_{i=0}^{h-2} a_i^* \right\}, \quad (4.9)$$

where $a_j^* \equiv \boldsymbol{H}' \boldsymbol{A}^j \boldsymbol{H}$ in the last relation.

Secondly, let us derive a relation corresponding to (4.7). From (3.1) and (3.2), we have the following relation first:

$$\eta_2 \equiv E(x_{n,h} - x_{n+h})^2$$
$$= E\left\{ \boldsymbol{H}' \left(\boldsymbol{A}^h + \boldsymbol{A}^{h-1} \boldsymbol{B} e_n - \prod_{j=1}^{h} (\boldsymbol{A} + \boldsymbol{B} e_{n+h-j}) \right) \boldsymbol{x}_n + \sigma^2 \boldsymbol{H}' \sum_{j=0}^{h-2} \boldsymbol{A}^j \boldsymbol{BC} - \boldsymbol{H}' \boldsymbol{C} e_{n+h} - \boldsymbol{H}' \sum_{i=1}^{h-1} \left(\prod_{j=1}^{i} (\boldsymbol{A} + \boldsymbol{B} e_{n+h-j}) \right) \boldsymbol{C} e_{n+h-i} \right\}^2, \quad (4.10)$$

Developing the right hand side of this relation, and taking expectation on each term, each term is as follows:

1). $$z_1 \equiv E\left\{ \boldsymbol{H}' \left(\boldsymbol{A}^h + \boldsymbol{A}^{h-1} \boldsymbol{B} e_n - \prod_{j=1}^{h} (\boldsymbol{A} + \boldsymbol{B} e_{n+h-j}) \right) \boldsymbol{x}_n \right\}^2$$

$$= \mathrm{tr}\left[\boldsymbol{A}' \sum_{j=1}^{h-1} \sigma^{2j} \boldsymbol{P}' \begin{bmatrix} \boldsymbol{A}=h-1-j \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{P} \begin{bmatrix} \boldsymbol{A}=h-1-j \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{A\Gamma} \right]$$
$$+ 2\mathrm{tr}\left[\boldsymbol{A}' \sum_{j=1}^{h-1} \sigma^{2j} \boldsymbol{P}' \begin{bmatrix} \boldsymbol{A}=h-1-j \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{P} \begin{bmatrix} \boldsymbol{A}=h-1-j \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{BS} \right]$$
$$+ \mathrm{tr}\left[\boldsymbol{B}' \sum_{j=1}^{h-1} \sigma^{2j} \boldsymbol{P}' \begin{bmatrix} \boldsymbol{A}=h-1-j \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{P} \begin{bmatrix} \boldsymbol{A}=h-1-j \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{BW} \right], \quad (4.11)$$

where $\boldsymbol{P}\begin{bmatrix} \boldsymbol{A}=i \\ \boldsymbol{B}=j \end{bmatrix} \equiv \begin{pmatrix} \boldsymbol{H}' \boldsymbol{A}^i \boldsymbol{B}^j \\ \boldsymbol{H}' \boldsymbol{A}^{i-1} \boldsymbol{BAB}^{j-1} \\ \boldsymbol{H}' \boldsymbol{A}^{i-1} \boldsymbol{B}^2 \boldsymbol{AB}^{j-2} \\ \vdots \\ \boldsymbol{H}' \boldsymbol{B}^j \boldsymbol{A}^i \end{pmatrix} : \dfrac{(i+j)!}{i!j!} \times p$ matrix. For example

$$\boldsymbol{P}\begin{bmatrix} \boldsymbol{A}=2 \\ \boldsymbol{B}=2 \end{bmatrix} = (\boldsymbol{B}^{2\prime} \boldsymbol{A}^{2\prime} \boldsymbol{H}, \boldsymbol{B}' \boldsymbol{A}' \boldsymbol{B}' \boldsymbol{A}' \boldsymbol{H}, \boldsymbol{A}' \boldsymbol{B}^{2\prime} \boldsymbol{A}' \boldsymbol{H}, \boldsymbol{B}' \boldsymbol{A}^{2\prime} \boldsymbol{B}' \boldsymbol{H}, \boldsymbol{A}' \boldsymbol{B}' \boldsymbol{A}' \boldsymbol{B}' \boldsymbol{H}, \boldsymbol{A}^{2\prime} \boldsymbol{B}^{2\prime} \boldsymbol{H})'.$$

2). $$z_2 \equiv E\left\{ \boldsymbol{H}' \sum_{i=1}^{h-1} \left(\prod_{j=1}^{i} (\boldsymbol{A} + \boldsymbol{B} e_{n+h-j}) \right) \boldsymbol{C} e_{n+h-i} \right\}^2$$

$$= \sum_{j=1}^{h-1} \left\{ \boldsymbol{C}' \sum_{i=0}^{j} \sigma^{2(i+1)} \boldsymbol{P}^{*\prime} \begin{bmatrix} \boldsymbol{A}=j-i \\ \boldsymbol{B}=i \end{bmatrix} \boldsymbol{P} \begin{bmatrix} \boldsymbol{A}=j-i \\ \boldsymbol{B}=j \end{bmatrix} \boldsymbol{C} \right\}$$
$$+ \delta(h:2) \left\{ 2\boldsymbol{C}' \left(\sum_{j=1}^{h-2} \sum_{i=0}^{h-2-j} \boldsymbol{Q}_j' (\boldsymbol{I}_{2^j} \otimes \boldsymbol{HH}') \tilde{\boldsymbol{Q}}_j \boldsymbol{A}^i \boldsymbol{B} \right) \boldsymbol{C} \right\}. \quad (4.12)$$

Here $\delta(h:i)$ is 0 for $h \leq i$, 1 for $h = i+1, i+2, \ldots$. Further, $\boldsymbol{Q}_j$ and $\tilde{\boldsymbol{Q}}_j$ are, respectively,

$$\boldsymbol{Q}_j \equiv \begin{cases} \sigma^4 \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix} & \text{for } j=1, \\ \begin{bmatrix} \boldsymbol{A} \odot \boldsymbol{Q}_{j-1} \\ \sigma^2 \boldsymbol{B} \odot \boldsymbol{Q}_{j-1} \end{bmatrix} & \text{for } j=2, 3, \dots, h-2, \end{cases}$$

$$\tilde{\boldsymbol{Q}}_j \equiv \begin{cases} \begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{A} \end{bmatrix} & \text{for } j=1, \\ \begin{bmatrix} \boldsymbol{A} \odot \tilde{\boldsymbol{Q}}_{j-1} \\ \boldsymbol{B} \odot \tilde{\boldsymbol{Q}}_{j-1} \end{bmatrix} & \text{for } j=2, 3, \dots, h-2, \end{cases}$$

where, let $\boldsymbol{R}_i$ $(i=1,2,\dots,s)$ be $p \times p$ matrices and let $\boldsymbol{R}=(\boldsymbol{R}_1', \boldsymbol{R}_2', \dots, \boldsymbol{R}_s')'$, the operator is then defined by $\boldsymbol{A} \odot \boldsymbol{R} \equiv \begin{pmatrix} \boldsymbol{A}\boldsymbol{R}_1 \\ \vdots \\ \boldsymbol{A}\boldsymbol{R}_s \end{pmatrix}$. Also, let $\tau=(i+j)!/i!j!$, $\boldsymbol{P}\begin{bmatrix} \boldsymbol{A}=i \\ \boldsymbol{B}=j \end{bmatrix} = \begin{pmatrix} \boldsymbol{p}_1 \\ \vdots \\ \boldsymbol{p}_\tau \end{pmatrix}$, $\boldsymbol{P}^*\begin{bmatrix} \boldsymbol{A}=i \\ \boldsymbol{B}=j \end{bmatrix} = \begin{pmatrix} \boldsymbol{p}_1^* \\ \vdots \\ \boldsymbol{p}_\tau^* \end{pmatrix}$, and let $\boldsymbol{p}_r, \boldsymbol{p}_r^*$ $(r=1, 2, \dots, \tau)$ be $1 \times p$ row vectors. The new matrix $\boldsymbol{P}^*\begin{bmatrix} \boldsymbol{A}=i \\ \boldsymbol{B}=j \end{bmatrix}$ is defined by

$$\boldsymbol{p}_r^* = \begin{cases} 3\boldsymbol{p}_r & \text{if the last matrix of matrices defining the row vector } \boldsymbol{p}_r \text{ in } \boldsymbol{p}(\cdot) \text{ is matrix } \boldsymbol{B}, \\ \boldsymbol{p}_r & \text{otherwise,} \end{cases}$$

for $r=1, 2, \dots, \tau$. For instance, $\boldsymbol{P}^*\begin{bmatrix} \boldsymbol{A}=1 \\ \boldsymbol{B}=2 \end{bmatrix} = \begin{pmatrix} 3\boldsymbol{H}'\boldsymbol{A}\boldsymbol{B}^2 \\ 3\boldsymbol{H}'\boldsymbol{B}\boldsymbol{A}\boldsymbol{B} \\ \boldsymbol{H}'\boldsymbol{B}^2\boldsymbol{A} \end{pmatrix}$. It should be noted that, in case $h=2$, the second term of (4.12) does not appear.

3). $$z_3 \equiv -2E\Big\{\boldsymbol{H}'\Big(\boldsymbol{A}^h + \boldsymbol{A}^{h-1}\boldsymbol{B}e_n - \prod_{j=1}^{h}(\boldsymbol{A}+\boldsymbol{B}e_{n+h-j})\Big)\boldsymbol{x}_n\boldsymbol{H}' \sum_{i=1}^{h-1}\Big(\prod_{j=1}^{i}(\boldsymbol{A}+\boldsymbol{B}e_{n+h-j})\Big)\boldsymbol{C}e_{n+h-i}\Big\}$$

$$\begin{aligned} &= 2\sigma^2(\boldsymbol{m}'\boldsymbol{A}' + \sigma^2\boldsymbol{C}'\boldsymbol{B}')\boldsymbol{A}^{h-2'}\boldsymbol{B}'\boldsymbol{H}\boldsymbol{H}'\boldsymbol{A}\boldsymbol{C} \\ &\quad + \delta(h:2)\Big[-2\sigma^2(\boldsymbol{m}'\boldsymbol{A}' + \sigma^2\boldsymbol{C}'\boldsymbol{B}')\boldsymbol{A}^{h-1'}\boldsymbol{H}\boldsymbol{H}'\sum_{j=1}^{h-2}\boldsymbol{A}^j\boldsymbol{B}\boldsymbol{C} \\ &\qquad + 2\sigma^2(\boldsymbol{m}'\boldsymbol{A}' + \sigma^2\boldsymbol{C}'\boldsymbol{B}') \\ &\qquad \Big\{\delta(h:3)\sum_{i=2}^{h-2}\Big(\boldsymbol{A}^{h-i'}\sum_{r=0}^{i-1}\sigma^{2r}\boldsymbol{P}'\begin{bmatrix} \boldsymbol{A}=i-1-r \\ \boldsymbol{B}=r \end{bmatrix}\boldsymbol{P}\begin{bmatrix} \boldsymbol{A}=i-1-r \\ \boldsymbol{B}=r \end{bmatrix}\boldsymbol{B}\boldsymbol{C} \\ &\qquad + \boldsymbol{A}^{h-1-i'}\boldsymbol{B}'\sum_{r=0}^{i-1}\sigma^{2r}\boldsymbol{P}'\begin{bmatrix} \boldsymbol{A}=i-1-r \\ \boldsymbol{B}=r \end{bmatrix}\boldsymbol{P}\begin{bmatrix} \boldsymbol{A}=i-1-r \\ \boldsymbol{B}=r \end{bmatrix}\boldsymbol{A}\boldsymbol{C}\Big) \\ &\qquad + \boldsymbol{B}'\sum_{r=0}^{h-2}\sigma^{2r}\boldsymbol{P}'\begin{bmatrix} \boldsymbol{A}=h-2-r \\ \boldsymbol{B}=r \end{bmatrix}\boldsymbol{P}\begin{bmatrix} \boldsymbol{A}=h-2-r \\ \boldsymbol{B}=r \end{bmatrix}\boldsymbol{A}\boldsymbol{C} \end{aligned}$$

$$+\boldsymbol{A}'\sum_{r=0}^{h-2}\sigma^{2r}\boldsymbol{P}'\begin{bmatrix}A=h-2-r\\ B=r\end{bmatrix}\boldsymbol{P}\left[\begin{matrix}A=h-2-r\\ B=r\end{matrix}\right]\boldsymbol{BC}\Big\}\Bigg]. \tag{4.13}$$

Here also, it is to be noted that, in case $h=2$, only the first term of (4.13) appears.

4). $$z_4 \equiv -2\sigma^2 E\left\{\left(\boldsymbol{H}'\sum_{j=0}^{h-2}\boldsymbol{A}^j\boldsymbol{BC}\right)\boldsymbol{H}'\sum_{i=1}^{h-1}\left[\prod_{j=1}^{i}(\boldsymbol{A}+\boldsymbol{B}e_{n+h-j})\right]\boldsymbol{C}e_{n+h-i}\right\}$$

$$=-2\sigma^4\left\{\boldsymbol{H}'\sum_{j=0}^{h-2}\boldsymbol{A}^j\boldsymbol{BC}\right\}^2. \tag{4.14}$$

5). $$z_5 \equiv E\left\{\sigma^2\boldsymbol{H}'\sum_{j=0}^{h-2}\boldsymbol{A}^j\boldsymbol{BC}\right\}^2+E\{\boldsymbol{H}'\boldsymbol{C}e_{n+h}\}^2=\sigma^4\left\{\boldsymbol{H}'\sum_{j=0}^{h-2}\boldsymbol{A}^j\boldsymbol{BC}\right\}^2+\sigma^2. \tag{4.15}$$

Noting that the other terms disappear by taking expectations, by means of (4.11), (4.12), (4.13), (4.14) and (4.15), η_2 is given by

$$\eta_2=z_1+z_2+z_3+z_4+z_5. \tag{4.16}$$

Consequently from (4.9) and (4.16), the A.M.S.E. of the multi-step ahead prediction for the $BL(p, 0, p, 1)$ model is finally expressed as

$$\text{A. M. S. E. }(\hat{x}_{n,h})\sim\eta_1+\eta_2 \qquad \text{for } h\geq 2. \tag{4.17}$$

As it has been mentioned in the introduction, the $BL(p, 0, p, 1)$ model includes the $AR(p)$ process as a special case. Setting the matrix $\boldsymbol{B}$ to $\boldsymbol{O}$ in (2.3), from (4.8) we have

$$\text{A.M.S.E.}(\hat{x}_{n,1})\sim\sigma^2+\mathrm{tr}[\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{\Gamma}].$$

Since $(\hat{\boldsymbol{a}}-\boldsymbol{a})\sim N(0,\ n^{-1}\sigma^2\boldsymbol{\Gamma}^{-1})$ in the $AR(p)$ process (see Box and Jenkins[3]), it follows that

$$\text{A. M. S. E. }(\hat{x}_{n,1})\sim\sigma^2(1+p/n), \tag{4.18}$$

which is the same result as Bloomfield[2]. Furthermore, in this case, $\eta_1=\mathrm{tr}[\boldsymbol{M}'(h)\ \boldsymbol{V}(\hat{\boldsymbol{a}})\ \boldsymbol{M}(h)\ \boldsymbol{\Gamma}]$, $z_1=0$, $z_2=\sigma^2\sum_{j=1}^{h-1}(\boldsymbol{C}'\boldsymbol{A}^j\boldsymbol{H})^2$, $z_3=0$, $z_4=0$, $z_5=\sigma^2$, we thus find that

$$\text{A. M. S. E. }(\hat{x}_{n,h})\sim\mathrm{tr}[\boldsymbol{M}'(h)\boldsymbol{V}(\hat{\boldsymbol{a}})\boldsymbol{M}(h)\boldsymbol{\Gamma}]+\sigma^2\sum_{j=1}^{h-1}(\boldsymbol{C}'\boldsymbol{A}^j\boldsymbol{H})^2+\sigma^2$$

$$=n^{-1}\sigma^2\mathrm{tr}[\boldsymbol{M}'(h)\boldsymbol{\Gamma}^{-1}\boldsymbol{M}(h)\boldsymbol{\Gamma}]+\sigma^2\sum_{j=1}^{h-1}(\boldsymbol{C}'\boldsymbol{A}^j\boldsymbol{H})^2+\sigma^2, \tag{4.19}$$

which is also equal to the A.M.S.E. of the multi-step ashead prediction for the $AR(p)$ process given by Yamamoto[11].

5. Concluding remarks

In the cause of non-linearity, an explicit formula (4.17) has in some degree become a tedious form, but it is found that the approach applying to the linear time series $AR(p)$ process is also useful to the non-linear time series model $BL(p, 0, p, 1)$ process. Moreover, we infer that it may be possible to apply this approach to a more general bilinear model.

There are some points which need to be pointed out. Firstly, we didn't refer to the structures of $\boldsymbol{V}(\hat{\boldsymbol{a}})$, $\boldsymbol{V}(\hat{\boldsymbol{b}})$ and $Cov(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}})$, however the properties of their structures need to be understood in detail. Secondly, for simplicity, the variance of e_t, σ^2, is assumed as a known parameter. Even if σ^2 is an unknown parameter, the formula (4.17) will be slightly modified and our approach will be still available. Finally, in the case of relatively small sample size, the accuracy of approximate naturally become dull, thus we must be careful of the use of the formula (4.17).

References

[1] Baillie, R. T. (1979), "Asymptotic Prediction Mean Squared Error for Vector Autoregressive Models", *Biometrika*, Vol. 66, No. 3, 675-678.

[2] Bloomfield, P. (1972), "On the Error of Prediction of a Time Series", *Biometrika*, Vol. 59, No. 3, 501-507.

[3] Box, G.E.P. and G. M. Jenkins (1970), "*Time Series Analysis Forecasting and Control*", Holden-Day, San Francisco.

[4] Cramér, H. (1946), "*Mathematical Methods of Statistics*", Princeton University Press, Princeton, N. J.

[5] Fotopoulos, S. B. and W. D. Ray (1983), "Components of Prediction Errors for a Stationary Process with Estimated Parameters", *Journal of Time Series*

Analysis, Vol. 4. No. 1, 1-8.

[6] Granger, C.W.J. and A.P. Andersen (1978), "*An Intoroduction to Bilinear Time Series Models*", Vandenhoek and Ruprecht, Göttingen.

[7] Neudecker, H. (1969), "Some Theorems on Matrix Differentiation with Special Reference to Kronecker Matrix Products", *Journal of the American Statistical Association*, Vol. 64, 953-963.

[8] Reinsel, G. (1980), "Asymptotic Properties of Prediction Errors for the Multivariate Autoregressive Model Using Estimated Parameters", *J.R.S.S.*, Vol. 42, No. 3, 328-333.

[9] Subba Rao, T. (1981), "On the Theory of Bilinear Time Series Models", *J.R.S.S.*, Vol. 43, No. 2, 244-255.

[10] Subba Rao, T. and M. M. Gabr (1984), "*An Introduction to Bispectral and Bilinear Time Series Models*", Springer-Verlag.

[11] Yamamoto, T. (1976), "Asymptotic Mean Square Prediction Error for an Autoregressive Model with Estimated Coefficients", *Applied Statistics*, Vol. 25, No. 2, 123-127.

[12] Yamomoto, T. (1981), "Predictions of Multivariate Autoregressive-Moving Average Models", *Biometrika*, Vol. 68, No. 2, 485-492.