Ann. Inst. Statist. Math.
36 (1984), Part A, 419-430

ASYMPTOTIC BIAS OF THE LEAST SQUARES ESTIMATOR FOR MULTIVARIATE AUTOREGRESSIVE MODELS*

TAKU YAMAMOTO AND NAOTO KUNITOMO

(Received May 18, 1983; revised Nov. 7, 1983)

Summary

The asymptotic bias of the least squares estimator for the multivariate autoregressive models is derived. The formulas for the low order univariate autoregressive models are given in terms of the simple functions of parameters. Our results are useful to the bias correction method of the least squares estimation.

1. Introduction and model

To estimate the multivariate autoregressive models, the least squares estimation method has been commonly used. Kendall [5] obtained the asymptotic bias of the least squares estimator to order n^{-1} for the first-order autoregressive model with a constant term, while White [8] to orders higher than n^{-1} for the model without constant term, where n is the sample size. Sawa [6] has recently shown that these approximations are quite accurate even for small n, and thus recommended their use for the bias correction. His study is based on the exact moments of the least squares estimator. Shenton and Johnson [7] reached a similar conclusion by examining the accuracy of these approximations based on the Monte Carlo experiments.

In the present note, we derive the asymptotic bias of the least squares estimator to order n^{-1} for the multivariate autoregressive models with a constant term. Since the resulting bias formula is expressed as a simple matrix function of parameters, it can be readily used for the bias correction. It includes various simpler models as special cases, and the bias formulas for low order univariate autoregressive models

* This work was supported by National Science Foundation Grant SES79-13976 at the Institute for Mathematical Studies in the Social Sciences, Stanford University. This paper is a revision of Discussion Paper No. 504, The Center for Mathematical Studies in Economics and Management Science, Northwestern University, October 1981.

Some key words: Asymptotic bias, least squares, autoregressive models, multivariate time series, bias correction.

are turned out to be relatively simple. Incidentally, we will point out an error in the asymptotic bias formula given by Groenwald and de Waal [4].

Let us consider the m-variable qth order autoregressive process

$$(1) \qquad \boldsymbol{x}_t=\boldsymbol{B}_1\boldsymbol{x}_{t-1}+\cdots+\boldsymbol{B}_q\boldsymbol{x}_{t-q}+\boldsymbol{b}_0+\boldsymbol{w}_t ,$$

where $\boldsymbol{x}_t=(x_{t1}, x_{t2},\cdots, x_{tm})'$, $\boldsymbol{w}_t=(w_{t1}, w_{t2},\cdots, w_{tm})'$, $\boldsymbol{b}_0=(b_{01}, b_{02},\cdots, b_{0m})'$, and $\boldsymbol{B}_k$ $(k=1,\cdots, q)$ are the $m\times m$ coefficient matrices. We make the following assumptions:

(A1) The $\boldsymbol{w}_t$'s are independently, identically distributed with mean zero and covariance matrix $\boldsymbol{\Omega}_w$ and that all characteristic roots of $|\lambda^q\boldsymbol{I}_m-\lambda^{q-1}\boldsymbol{B}_1-\cdots\boldsymbol{B}_q|=0$ are less than unity in absolute value.

(A2) For some s_0, $\mathrm{E}\{|w_{ti}|^{s_0}\}<+\infty$, $i=1,\cdots, m$.

Let $\boldsymbol{y}_t$, $\boldsymbol{v}_t$, and $\boldsymbol{b}$ be $p\times 1$ $(p=m\times q)$ vectors such that

$$\boldsymbol{y}_t=\begin{bmatrix}\boldsymbol{x}_t\\ \boldsymbol{x}_{t-1}\\ \vdots\\ \boldsymbol{x}_{t-q+1}\end{bmatrix}, \qquad \boldsymbol{v}_t=\begin{bmatrix}\boldsymbol{w}_t\\ \boldsymbol{0}\\ \vdots\\ \boldsymbol{0}\end{bmatrix}, \qquad \boldsymbol{b}=\begin{bmatrix}\boldsymbol{b}_0\\ \boldsymbol{0}\\ \vdots\\ \boldsymbol{0}\end{bmatrix},$$

and $\boldsymbol{B}$ and $\mathrm{E}(\boldsymbol{v}_t\boldsymbol{v}_t')=\boldsymbol{\Omega}_v$ be the $p\times p$ matrices such that

$$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1\ \boldsymbol{B}_2\cdots\boldsymbol{B}_q\\ \boldsymbol{I}_{m(q-1)} \quad \boldsymbol{0}\end{bmatrix}, \qquad \boldsymbol{\Omega}_v=\begin{bmatrix}\boldsymbol{\Omega}_w & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix},$$

where $\boldsymbol{I}_{m(q-1)}$ is the identity matrix of order $m(q-1)$. Following Anderson [1], (1) can be rewritten as a first order vector autoregressive representation:

$$(2) \qquad \boldsymbol{y}_t=\boldsymbol{B}\boldsymbol{y}_{t-1}+\boldsymbol{b}+\boldsymbol{v}_t ,$$

where $\boldsymbol{y}_t=[y_{t1}, y_{t2},\cdots, y_{tp}]'$, $\boldsymbol{v}_t=[v_{t1}, v_{t2},\cdots, v_{tp}]'$, and $\boldsymbol{b}=(\beta_{i0})$ are $p\times 1$ vectors, and all characteristic roots of $\boldsymbol{B}$ are less than unity in absolute value. Alternatively, the process is expressed as

$$(3) \qquad \boldsymbol{z}_t=\boldsymbol{A}\boldsymbol{z}_{t-1}+\boldsymbol{u}_t ,$$

where $\boldsymbol{z}_t=[\boldsymbol{y}_t', 1]'$, $\boldsymbol{u}_t=[\boldsymbol{v}_t', 0]'$ with $\boldsymbol{\Omega}_u=\mathrm{E}(\boldsymbol{u}_t\boldsymbol{u}_t')$, and $\boldsymbol{A}$ and $\boldsymbol{\Omega}_u$ are the $(p+1)\times(p+1)$ matrices

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{B} & \boldsymbol{b}\\ \boldsymbol{0} & 1\end{bmatrix}, \qquad \text{and} \qquad \boldsymbol{\Omega}_u=\begin{bmatrix}\boldsymbol{\Omega}_v & \boldsymbol{0}\\ \boldsymbol{0} & 0\end{bmatrix}.$$

From the successive substitution, we can express $\boldsymbol{z}_t$ and $\mathrm{E}(\boldsymbol{z}_t\boldsymbol{z}_t')$ as

$$\boldsymbol{z}_t=\boldsymbol{d}+\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{u}_{t-i} , \qquad \boldsymbol{\Gamma}=\mathrm{E}(\boldsymbol{z}_t\boldsymbol{z}_t')=\boldsymbol{d}\boldsymbol{d}'+\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i ,$$

and

$$\boldsymbol{\Gamma}_1 = \mathrm{E}(\boldsymbol{z}_t\boldsymbol{z}'_{t-1}) = \boldsymbol{d}\boldsymbol{d}' + \boldsymbol{A}\sum_{i=0}^{\infty} \boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i ,$$

where $\boldsymbol{A}^0=\boldsymbol{I}_{p+1}$, $\boldsymbol{I}_k$ is the identity matrix of order k, and $\boldsymbol{d}$ is the $(p+1)$th column of $\boldsymbol{A}^\infty$, i.e., $\boldsymbol{d}=[((\boldsymbol{I}_p-\boldsymbol{B})^{-1}\boldsymbol{b})', 1]'$.

2. Asymptotic bias of the least squares estimator

For given observations $\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_n$, the least squares estimator of $\boldsymbol{A}$ is given by

$$\hat{\boldsymbol{A}} = \hat{\boldsymbol{\Gamma}}_1\hat{\boldsymbol{\Gamma}}^{-1} , \tag{4}$$

where

$$\hat{\boldsymbol{\Gamma}} = \frac{1}{n}\sum_{t=1}^{n} \boldsymbol{z}_{t-1}\boldsymbol{z}'_{t-1} \quad \text{and} \quad \hat{\boldsymbol{\Gamma}}_1 = \frac{1}{n}\sum_{t=1}^{n} \boldsymbol{z}_t\boldsymbol{z}'_{t-1} .$$

To derive the asymptotic bias we assume that the initial observation $\boldsymbol{y}_0$ obeys the same multivariate distribution as $\boldsymbol{y}_t$ for $t>1$. From the proof in Appendix it is easily seen that all the results hold, even if we replace the assumption on $\boldsymbol{y}_0$ by $\boldsymbol{y}_0=0$. The following lemma is needed for determining the order of magnitude in the expansion of expectations, which has been given as Lemma 3.3 in Bhansali [2].

LEMMA 1. *Assume* (A1) *and* (A2) *with* $s_0=4k$, *where* $k\geqq 1$ *is a prefixed integer. Then, as* $n\to+\infty$,

$$\mathrm{E}\{\|\hat{\boldsymbol{\Gamma}}-\boldsymbol{\Gamma}\|^{2k}\} = O(n^{-k}) , \tag{5}$$

and

$$\mathrm{E}\{\|\hat{\boldsymbol{\Gamma}}_1-\boldsymbol{\Gamma}_1\|^{2k}\} = O(n^{-k}) , \tag{6}$$

where $\|\boldsymbol{C}\|=\sup(\boldsymbol{\beta}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{\beta})^{1/2}$ $(\boldsymbol{\beta}'\boldsymbol{\beta}\leqq 1)$ *for any matrix* $\boldsymbol{C}$ *and vector* $\boldsymbol{\beta}$ *is the matrix norm.*

We further assume:

(A3) $\mathrm{E}\{\|\hat{\boldsymbol{\Gamma}}^{-1}\|^2\}$ be bounded for some $n>N_k$.

It may be noted that the assumptions (A2) and (A3) are satisfied, if the distribution of $\boldsymbol{w}_t$ is normal (see Fuller and Hasza [2]).

Now, we define the asymptotic bias of $\hat{\boldsymbol{A}}$, denoted by ABIAS$(\hat{\boldsymbol{A}})$, $\mathrm{E}(\hat{\boldsymbol{A}}-\boldsymbol{A})$, *ignoring the term of order* $O(n^{-3/2})$ for notational convenience.

THEOREM. *Assume* (A1), (A2), *and* (A3) for $s_0=16$. *Then, as* $n\to+\infty$, *the asymptotic bias of the least squares estimator* (4) *for the multi-*

variate first-order autoregressive model with a constant term (3) *is*

$$(7) \qquad \mathrm{ABIAS}(\hat{A}) = -n^{-1}\Omega_u \sum_{k=0}^{\infty} \{A'^k \operatorname{tr}(A^{k+1}) + A'^{2k+1}\}\Gamma^{-1} .$$

PROOF. See Appendix.

It is sometimes convenient to express the above result in terms of the original notations B, b and Ω_v associated with the model representation (2). By (A.12) and (A.14) in Appendix, it is easy to show that

$$\Omega_u \sum_{k=0}^{\infty} A'^{mk}\Gamma^{-1} = \begin{bmatrix} \Omega_u \sum_{k=0}^{\infty} B'^{mk} \\ 0' \end{bmatrix} D^{-1}(I_p, -d_1) , \qquad m=1, 2 ,$$

where $D=\sum_{i=0}^{\infty} B^i \Omega_v B'^i$ and $d_1=(I_p-B)^{-1}b$. Then, combining (A.15), (A.16), (A.17) and (A.18) in Appendix, we get

$$(8) \qquad \mathrm{ABIAS}(\hat{B}, \hat{b}) = -n^{-1}\Omega_v \sum_{k=0}^{\infty} \{B'^k + B'^k \operatorname{tr}(B^{k+1}) + B'^{2k+1}\} \times D^{-1}(I_p, -d_1) .$$

It is interesting to note that, while $\hat{B}$ and $\hat{b}$ are biased of order n^{-1}, the unconditional mean of the process calculated by them is not biased in the following sense:

$$\{I_p - A\,\mathrm{E}(\hat{B})\}^{-1} A\,\mathrm{E}(\hat{b}) = [I_p - \{B + \mathrm{ABIAS}(\hat{B})\}]^{-1}[b + \mathrm{ABIAS}(\hat{b})] = (I_p - B)^{-1} b .$$

In the above, the use was made of the fact that $\mathrm{ABIAS}(\hat{b}) = -\{\mathrm{ABIAS}(\hat{B})\}d_1$ and $d_1=(I_p-B)^{-1}b$. For the univariate pth order autoregressive model with a constant term, we may just pick the first row of (7) or (8). However, the formula (8) can be simplified as follows:

$$(9) \qquad \mathrm{ABIAS}(\hat{\beta}', \hat{\beta}_0) = -n^{-1}e' \sum_{k=0}^{\infty} \{B'^k + B'^k \operatorname{tr}(B^{k+1}) + B'^{2k+1}\} \times F^{-1}(I_p, -c\mathbf{1})$$

where

$$B = \begin{bmatrix} \beta' & \\ I_{p-1} & 0 \end{bmatrix} ,$$

$\beta=[\beta_1, \beta_2, \cdots, \beta_p]'$, $F=\sum_{i=0}^{\infty} B^i M B'^i$, $M=ee'$, $e=[1, 0, \cdots, 0]'$, $\mathbf{1}=[1, 1, \cdots, 1]'$, $c=\beta_0 \Big/ \left(1-\sum_{k=1}^{p} \beta_k\right)$. The above expression (9) depends solely upon the coefficient parameters, but not the variance parameter. Thus, it is more convenient to use in the bias correction scheme. When p is not quite large, the formula (9) can be further simplified.

COROLLARY 1. *For the pth order univariate autoregressive model with a constant term* ($p=1, 2, 3$), *the asymptotic bias of the least squares estimator is given by*

$$\text{(10)} \qquad \text{ABIAS}(\hat{\beta}_1, \hat{\beta}_0) = -n^{-1}(1+3\beta_1)\left[1, -\frac{\beta_0}{1-\beta_1}\right] \qquad \textit{for } p=1,$$

$$\text{(11)} \qquad \text{ABIAS}(\hat{\beta}_2, \hat{\beta}_1, \hat{\beta}_0) = -n^{-1}\left[2+4\beta_2, 1+\beta_1+\beta_2, -\frac{\beta_0(3+\beta_1+5\beta_2)}{1-\beta_1-\beta_2}\right] \qquad \textit{for } p=2,$$

$$\text{(12)} \qquad \text{ABIAS}(\hat{\beta}_3, \hat{\beta}_2, \hat{\beta}_1, \hat{\beta}_0) = -n^{-1}\left[5\beta_3, 2+4\beta_2, \beta_1+2\beta_3, -\frac{\beta_0(2+\beta_1+4\beta_2+7\beta_3)}{1-\beta_1-\beta_2-\beta_3}\right] \qquad \textit{for } p=3.$$

PROOF. See Appendix.

The result of $\text{ABIAS}(\hat{\beta}_1)$ for $p=1$ is consistent with Kendall [5], while, as far as we know, other formulas have not been previously obtained. The completely parallel result to (7) is also derived for the case of no constant term.

COROLLARY 2. *Let* $\boldsymbol{B}$ *be the least squares estimator for the multivariate first order autoregressive model in* (2) *with a priori knowledge of* $\boldsymbol{b}_0=\boldsymbol{0}$. *Then the asymptotic bias of* $\hat{\boldsymbol{B}}$ *is*

$$\text{(13)} \qquad \text{ABIAS}(\hat{\boldsymbol{B}}) = -n^{-1}\boldsymbol{\Omega}_v \sum_{k=0}^{\infty} \{\boldsymbol{B}'^k \operatorname{tr}(\boldsymbol{B}^{k+1}) + \boldsymbol{B}'^{2k+1}\}\boldsymbol{D}^{-1},$$

where

$$\hat{\boldsymbol{B}} = \left(\sum_{t=1}^{n} \boldsymbol{y}_t \boldsymbol{y}'_{t-1}\right)\left(\sum_{t=1}^{n} \boldsymbol{y}_{t-1} \boldsymbol{y}'_{t-1}\right)^{-1}.$$

PROOF. Since $\boldsymbol{d}=\boldsymbol{0}$ or $\boldsymbol{d}_1=\boldsymbol{0}$ in this case, the first two terms of (A.1) in Appendix are dropped. Correspondingly, the first term of (8) is dropped. Then, it becomes equivalent to (13). Q.E.D.

We note that, while the first term of (13) is the same as the second term of equation (3.13) in Groenwald and de Waal [4], the second term in (13) corrects an error in their first term. For the univariate pth order autoregressive model without constant term, the result in (13) is reduced to

$$\text{ABIAS}(\hat{\boldsymbol{\beta}}') = -n^{-1}\boldsymbol{e}' \sum_{k=0}^{\infty} \{\boldsymbol{B}'^{k} \operatorname{tr}(\boldsymbol{B}^{k+1}) + \boldsymbol{B}'^{2k+1}\}\boldsymbol{F}^{-1}. \tag{14}$$

Further, for the univariate low order autoregressive models without constant term, the above formula can be further simplified.

COROLLARY 3. *For the pth order univariate autoregressive model without constant term* ($p=1, 2, 3$), *the asymptotic bias of the least squares estimator is given by*

$$\text{ABIAS}(\hat{\beta}_1) = -n^{-1}[2\beta_1] \qquad \text{for } p=1, \tag{15}$$

$$\text{ABIAS}(\hat{\beta}_2, \hat{\beta}_1) = -n^{-1}[1+3\beta_2, \beta_1] \qquad \text{for } p=2, \tag{16}$$

$$\text{ABIAS}(\hat{\beta}_3, \hat{\beta}_2, \hat{\beta}_1) = -n^{-1}[4\beta_3, 1+\beta_2, \beta_1+\beta_3] \qquad \text{for } p=3. \tag{17}$$

PROOF. See Appendix.

The result of ABIAS $(\hat{\beta}_1)$ for $p=1$ is consistent with White [8] to order n^{-1} while other formulas have not been previously obtained. For the higher order univariate autoregressive models with $p>3$, the bias formula of the least squares estimator becomes more complicated.

Finally, although the bias formulas in (7)–(9) and (13) are expressed as the infinite sums of matrix arguments, we can truncate them by taking the summation from 0 to n in practice. This truncation is mathematically valid since the remaining terms are in the order $o(n^{-1})$. Therefore, our results have some applicability to the bias correction of the least squares estimation method in the multivariate autoregressive models.

Appendix: Proofs

We first establish a lemma which gives ABIAS $(\hat{\boldsymbol{A}})$ in terms of the matrices $\boldsymbol{A}$, $\boldsymbol{\Omega}_u$, $\boldsymbol{\Gamma}^{-1}$ and the vector $\boldsymbol{d}$.

LEMMA 2. *The asymptotic bias of the least squares estimator* (4) *for the multivariate first order autoregressive model with a constant term* (3) *is*

$$\begin{aligned}\text{ABIAS}(\hat{\boldsymbol{A}}) = -n^{-1}\boldsymbol{\Omega}_u\Bigg[\sum_{k=0}^{\infty} \boldsymbol{A}'^{k}\boldsymbol{\Gamma}^{-1}\Bigg\{\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\boldsymbol{d} + \boldsymbol{d}\boldsymbol{d}'\boldsymbol{\Gamma}^{-1} + \Bigg(\sum_{i=0}^{\infty} \boldsymbol{A}^{i}\boldsymbol{\Omega}_u\boldsymbol{A}'^{i}\Bigg) \\ \times \boldsymbol{A}'^{k+1}\boldsymbol{\Gamma}^{-1} + \operatorname{tr}\Bigg(\boldsymbol{A}^{k+1}\Bigg(\sum_{i=0}^{\infty} \boldsymbol{A}^{i}\boldsymbol{\Omega}_u\boldsymbol{A}'^{i}\Bigg)\boldsymbol{\Gamma}^{-1}\Bigg)\Bigg\}\Bigg] + o(n^{-1}).\end{aligned} \tag{A.1}$$

PROOF. First, we expand $\hat{\boldsymbol{A}} - \boldsymbol{A}$ as

$$\sqrt{n}(\hat{\boldsymbol{A}} - \boldsymbol{A}) = \Bigg(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{u}_t \boldsymbol{z}'_{t-1}\Bigg)\hat{\boldsymbol{\Gamma}}^{-1} = \boldsymbol{W}\boldsymbol{\Gamma}^{-1} \sum_{i=0}^{k-1} (-\boldsymbol{V})^{i} + \boldsymbol{W}\hat{\boldsymbol{\Gamma}}^{-1}(-\boldsymbol{V})^{k}, \tag{A.2}$$

where

$$W=(1/\sqrt{n})\sum_{t=1}^{n}u_t z'_{t-1} \qquad \text{and} \qquad V=(\hat{\Gamma}-\Gamma)\Gamma^{-1}.$$

Then we take $k=2$ and

$$\text{(A.3)} \qquad \mathrm{E}\,\|\sqrt{n}(\hat{A}-A)-W\Gamma^{-1}(I-V)\|\leqq \mathrm{E}\,\|W\hat{\Gamma}^{-1}(-V)^2\|\,.$$

By the use of Lemma 1 and (A3) the right-hand side of (A.3) is in the order of n^{-1}. Ignoring the term of $O(n^{-3/2})$, ABIAS$(\hat{A})$ is reduced to

$$\text{(A.4)} \qquad \mathrm{ABIAS}\,(\hat{A})=-n^{-1}\,\mathrm{E}\left\{n^{-1}\sum_{t=1}^{n}u_t z'_{t-1}\Gamma^{-1}\left(\sum_{t=0}^{n-1}z_t z'_t\right)\Gamma^{-1}\right\}.$$

Using the expression of z_t, we can express $\sum_{t=1}^{n}u_t z'_{t-1}$ and $\sum_{t=0}^{n-1}z_t z'_t$ as

$$\text{(A.5)} \qquad \begin{aligned}&\sum_{t=1}^{n}u_t z'_{t-1}=\sum_{t=1}^{n}(g_t+h_t)\,,\\ &\sum_{t=0}^{n-1}z_t z'_t=\sum_{t=0}^{n-1}(dd'+F_t+G_t+G'_t+H_t+H'_t)\,,\end{aligned}$$

where

$$g_t=u_t d'\,,$$

$$h_t=\sum_{i=1}^{\infty}h_{t,i}=\sum_{i=1}^{\infty}u_t u'_{t-i}A'^{i-1}\,,$$

$$F_t=\sum_{i=0}^{\infty}A^i u_{t-i}u'_{t-i}A'^i\,,$$

$$G_t=\sum_{i=0}^{\infty}G_{t,i}=\sum_{i=0}^{\infty}du'_{t-i}A'^i\,,$$

$$H_t=\sum_{i=0}^{\infty}H_{t,i}=\sum_{i=0}^{\infty}A^i u_{t-i}\left(\sum_{j=1}^{\infty}u'_{t-i-j}A'^{i+j}\right).$$

Noting that $\mathrm{E}\,\{(u_s z'_{s-1})(z_t z'_t)\}=0$ for $s>t$, the expectation for given s is

$$\text{(A.6)} \qquad \begin{aligned}&\mathrm{E}\left\{u_s z'_{s-1}\Gamma^{-1}\left(\sum_{t=0}^{n-1}z_t z'_t\right)\Gamma^{-1}\right\}\\ &=\mathrm{E}\left\{u_s z'_{s-1}\Gamma^{-1}\left(\sum_{t=s}^{\infty}z_t z'_t\right)\Gamma^{-1}\right\}-\mathrm{E}\left\{u_s z'_{s-1}\Gamma^{-1}\left(\sum_{t=n}^{\infty}z_t z'_t\right)\Gamma^{-1}\right\},\\ &\qquad\qquad s=1,2,\cdots,n\,.\end{aligned}$$

We first evaluate the first term of the above. Since the expectation exists only when the time indices of u_t's are equal or pairwise equal, it can be expressed by (A.5) as

$$\text{(A.7)}\qquad \begin{aligned} &\mathrm{E}\left\{\boldsymbol{u}_s\boldsymbol{z}'_{s-1}\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{z}_t\boldsymbol{z}'_t\Big)\boldsymbol{\Gamma}^{-1}\right\} \\ &\quad=\mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{G}_t\Big)\boldsymbol{\Gamma}^{-1}+\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{G}'_t\Big)\boldsymbol{\Gamma}^{-1}\right. \\ &\qquad\left.+\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{H}_t\Big)\boldsymbol{\Gamma}^{-1}+\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{H}'_t\Big)\boldsymbol{\Gamma}^{-1}\right\} . \end{aligned}$$

The first term of the above can be reduced to

$$\begin{aligned} \mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{G}_t\Big)\boldsymbol{\Gamma}^{-1}\right\} &=\mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\sum_{i=0}^{\infty}\boldsymbol{G}_{t,i}\Big)\boldsymbol{\Gamma}^{-1}\right\} \\ &=\mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{k=0}^{\infty}\boldsymbol{G}_{s+k,k}\Big)\boldsymbol{\Gamma}^{-1}\right\} \\ &=\mathrm{E}\left\{\boldsymbol{u}_s\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\sum_{k=0}^{\infty}\boldsymbol{d}\boldsymbol{u}'_s\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\right\} . \end{aligned}$$

Since $\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\boldsymbol{d}$ is a scalar, we get

$$\text{(A.8)}\qquad \mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{G}_t\Big)\boldsymbol{\Gamma}^{-1}\right\}=\boldsymbol{\Omega}_u\Big(\sum_{k=0}^{\infty}\boldsymbol{A}'^k\Big)\boldsymbol{\Gamma}^{-1}(\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\boldsymbol{d}) .$$

Similarly, the second term of (A.6) is reduced to

$$\begin{aligned} \mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{G}'_t\Big)\boldsymbol{\Gamma}^{-1}\right\} &=\mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{k=0}^{\infty}\boldsymbol{G}'_{s+k,k}\Big)\boldsymbol{\Gamma}^{-1}\right\} \\ &=\mathrm{E}\left\{\boldsymbol{u}_s\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{u}_s\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\right\} . \end{aligned}$$

Since $\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{u}_s$ is a scalar, first transposing it and then taking the expectation, we get

$$\mathrm{E}\left\{\boldsymbol{g}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{G}'_t\Big)\boldsymbol{\Gamma}^{-1}\right\}=\boldsymbol{\Omega}_u\Big(\sum_{k=0}^{\infty}\boldsymbol{A}'^k\Big)\boldsymbol{\Gamma}^{-1}\boldsymbol{d}\boldsymbol{d}'\boldsymbol{\Gamma}^{-1} .$$

The third term of (A.7) can be written as

$$\begin{aligned} \mathrm{E}\left\{\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{H}_t\Big)\boldsymbol{\Gamma}^{-1}\right\} &=\mathrm{E}\left\{\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\sum_{i=0}^{\infty}\boldsymbol{H}_{t,i}\Big)\boldsymbol{\Gamma}^{-1}\right\} \\ &=\mathrm{E}\left\{\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\boldsymbol{h}_{s,i}\boldsymbol{\Gamma}^{-1}\boldsymbol{H}_{s+k,k}\boldsymbol{\Gamma}^{-1}\right\} \\ &=\mathrm{E}\left\{\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\boldsymbol{u}_s\boldsymbol{u}'_{s-i}\boldsymbol{A}'^{i-1}\boldsymbol{\Gamma}^{-1}\boldsymbol{A}^k\boldsymbol{u}_s\boldsymbol{u}'_{s-i}\boldsymbol{A}'^{k+i}\boldsymbol{\Gamma}^{-1}\right\} . \end{aligned}$$

Since $\boldsymbol{u}'_{s-i}\boldsymbol{A}'^{i-1}\boldsymbol{\Gamma}^{-1}\boldsymbol{A}^k\boldsymbol{u}_s$ is a scalar, first transposing it and then taking the expectation, we get

$$\mathrm{E}\left\{\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\Big(\sum_{t=s}^{\infty}\boldsymbol{H}_t\Big)\boldsymbol{\Gamma}^{-1}\right\}=\boldsymbol{\Omega}_u\left\{\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\Big(\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i\Big)\boldsymbol{A}'^{k+1}\right\}\boldsymbol{\Gamma}^{-1} .$$

The fourth term of (A.7) can be similarly written as

$$\begin{aligned}\mathrm{E}&\left\{\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\left(\sum_{t=s}^{\infty}\boldsymbol{H}_t'\right)\boldsymbol{\Gamma}^{-1}\right\}\\ &=\mathrm{E}\left\{\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\boldsymbol{h}_{s,i}\boldsymbol{\Gamma}^{-1}\boldsymbol{H}_{s+k,k}'\boldsymbol{\Gamma}^{-1}\right\}\\ &=\mathrm{E}\left\{\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\boldsymbol{u}_s\boldsymbol{u}_{s-i}'\boldsymbol{A}'^{i-1}\boldsymbol{\Gamma}^{-1}\boldsymbol{A}^{k+i}\boldsymbol{u}_{s-i}\boldsymbol{u}_s'\boldsymbol{A}'^{k}\boldsymbol{\Gamma}^{-1}\right\}\ .\end{aligned}$$

Since $\boldsymbol{u}_{s-i}'\boldsymbol{A}'^{i-1}\boldsymbol{\Gamma}^{-1}\boldsymbol{A}^{k+1}\boldsymbol{u}_{s-i}$ is a scalar, we use the trace operation and take the expectation, we get

$$\mathrm{E}\left\{\boldsymbol{h}_s\boldsymbol{\Gamma}^{-1}\left(\sum_{t=s}^{\infty}\boldsymbol{H}_t'\right)\boldsymbol{\Gamma}^{-1}\right\}=\boldsymbol{\Omega}_u\left[\sum_{k=0}^{\infty}\boldsymbol{A}'^{k}\boldsymbol{\Gamma}^{-1}\operatorname{tr}\left\{\boldsymbol{A}^{k+1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^{i}\boldsymbol{\Omega}_u\boldsymbol{A}'^{i}\right)\boldsymbol{\Gamma}^{-1}\right\}\right]\ . \tag{A.9}$$

Thus, the first term of (A.7) is

$$\begin{aligned}\mathrm{E}&\left\{\boldsymbol{u}_s\boldsymbol{z}_{s-1}'\boldsymbol{\Gamma}^{-1}\left(\sum_{t=0}^{\infty}\boldsymbol{z}_t\boldsymbol{z}_t'\right)\boldsymbol{\Gamma}^{-1}\right\}\\ &=\boldsymbol{\Omega}_u\left[\sum_{k=0}^{\infty}\boldsymbol{A}'^{k}\boldsymbol{\Gamma}^{-1}\left\{\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\boldsymbol{d}+\boldsymbol{d}\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}+\left(\sum_{i=0}^{\infty}\boldsymbol{A}^{i}\boldsymbol{\Omega}_u\boldsymbol{A}'^{i}\right)\boldsymbol{A}'^{k+1}\boldsymbol{\Gamma}^{-1}\right.\right.\\ &\qquad\left.\left.+\operatorname{tr}\left(\boldsymbol{A}^{k+1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^{i}\boldsymbol{\Omega}_u\boldsymbol{A}'^{i}\right)\boldsymbol{\Gamma}^{-1}\right)\right\}\right]\ ,\qquad s=1,2,\cdots,n\ .\end{aligned} \tag{A.10}$$

It is easily seen that the second term of (A.6) is $O(\boldsymbol{B}^{n-s})$ for $s=1,2,\cdots,n$, because $\boldsymbol{A}^{i}\boldsymbol{u}_t=[(\boldsymbol{B}^{i}\boldsymbol{v}_t)',0]'$. Since $O(\boldsymbol{B}^{n-s})=O(\rho^{n-s})$ where $0<\rho<1$ by the assumption on the characteristic roots, we have $\sum_{s=1}^{n}O(\boldsymbol{B}^{n-s})=O(1)$. Thus,

$$\mathrm{E}\left\{n^{-1}\sum_{t=1}^{n}\boldsymbol{u}_t\boldsymbol{z}_{t-1}'\boldsymbol{\Gamma}^{-1}\left(\sum_{t=0}^{n-1}\boldsymbol{z}_t\boldsymbol{z}_t'\right)\boldsymbol{\Gamma}^{-1}\right\}=\mathrm{E}\left\{\boldsymbol{u}_s\boldsymbol{z}_{s-1}'\boldsymbol{\Gamma}^{-1}\left(\sum_{t=0}^{\infty}\boldsymbol{z}_t\boldsymbol{z}_t'\right)\boldsymbol{\Gamma}^{-1}\right\}+O(n^{-1})\ .$$

From (A.2) and (A.10), we obtain (A.1) in Lemma 2. Q.E.D.

PROOF OF THEOREM. We now reduce (A.1) into a simpler formula. We note three basic relations. First, we have

$$\boldsymbol{A}^{i}=\begin{bmatrix}\boldsymbol{B}^{i} & \left(\sum_{k=0}^{i-1}\boldsymbol{B}^{k}\right)\boldsymbol{b}_0\\ \boldsymbol{0} & 1\end{bmatrix}\ ,\qquad i=1,2,\cdots\ . \tag{A.11}$$

Second, using the above relation, we have

$$\sum_{i=0}^{\infty}\boldsymbol{A}^{i}\boldsymbol{\Omega}_u\boldsymbol{A}'^{i}=\begin{bmatrix}\boldsymbol{D} & \boldsymbol{0}\\ \boldsymbol{0}' & 0\end{bmatrix}\ , \tag{A.12}$$

where $\boldsymbol{D}=\sum_{i=0}^{\infty}\boldsymbol{B}^{i}\boldsymbol{\Omega}_v\boldsymbol{B}'^{i}$. Finally, $\boldsymbol{\Gamma}^{-1}$ is expressed as

(A.13)
$$\boldsymbol{\Gamma}^{-1}=\begin{bmatrix} \boldsymbol{D}^{-1} & -\boldsymbol{D}^{-1}\boldsymbol{d}_1 \\ -\boldsymbol{d}_1'\boldsymbol{D}^{-1} & 1+\boldsymbol{d}_1'\boldsymbol{D}^{-1}\boldsymbol{d}_1 \end{bmatrix},$$

since $\boldsymbol{\Gamma}$ is given by

$$\boldsymbol{\Gamma}=\begin{bmatrix} \boldsymbol{D}+\boldsymbol{d}_1\boldsymbol{d}_1' & \boldsymbol{d}_1 \\ \boldsymbol{d}_1' & 1 \end{bmatrix},$$

where $\boldsymbol{d}_1=(\boldsymbol{I}_p-\boldsymbol{B})^{-1}\boldsymbol{b}$.

Noting that $\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\boldsymbol{d}=1$ by (A.13), the first term of (A.1) is given by

(A.14)
$$-n^{-1}\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\boldsymbol{d}=-n^{-1}\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}.$$

For the second term of (A.1), we have, by (A.11) and the fact that $\boldsymbol{\Gamma}^{-1}\boldsymbol{d}=[0, 0,\cdots, 0, 1]'$,

(A.15)
$$-n^{-1}\boldsymbol{\Omega}_u\left(\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\boldsymbol{d}\boldsymbol{d}'\boldsymbol{\Gamma}^{-1}\right)=-n^{-1}\begin{bmatrix} \boldsymbol{\Omega}_v\sum_{k=0}^{\infty}\boldsymbol{B}'^k & 0 \\ \boldsymbol{0} & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}=\boldsymbol{0}.$$

Thus, the second term of (A.1) vanishes. By (A.11), (A.12) and (A.13), it is easily verified that

$$\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i\right)\boldsymbol{A}'^{k+1}=\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^{2k+1}.$$

Thus, the third term of (A.1) is reduced to

(A.16)
$$-n^{-1}\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i\right)\boldsymbol{A}'^{k+1}\boldsymbol{\Gamma}^{-1}=-n^{-1}\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^{2k+1}\boldsymbol{\Gamma}^{-1}.$$

By (A.11), (A.12) and (A.13), we have

$$\boldsymbol{A}^{k+1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i\right)\boldsymbol{\Gamma}^{-1}=\begin{bmatrix} \boldsymbol{B}^{k+1} \\ \boldsymbol{0} \end{bmatrix}[\boldsymbol{I}_p, -\boldsymbol{d}_1].$$

Then

$$\mathrm{tr}\left\{\boldsymbol{A}^{k+1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i\right)\boldsymbol{\Gamma}^{-1}\right\}=\mathrm{tr}\,(\boldsymbol{B}^{k+1}).$$

Thus, the fourth term of (A.1) is given by

(A.17)
$$-n^{-1}\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\,\mathrm{tr}\left\{\boldsymbol{A}^{k+1}\left(\sum_{i=0}^{\infty}\boldsymbol{A}^i\boldsymbol{\Omega}_u\boldsymbol{A}'^i\right)\boldsymbol{\Gamma}^{-1}\right\}$$
$$=-n^{-1}\boldsymbol{\Omega}_u\sum_{k=0}^{\infty}\boldsymbol{A}'^k\boldsymbol{\Gamma}^{-1}\,\mathrm{tr}\,\{\boldsymbol{B}^{k+1}\}.$$

Combining (A.14), (A.15), (A.16) and (A.17), and noting that $1+\mathrm{tr}\,(\boldsymbol{B}^{k+1})=\mathrm{tr}\,(\boldsymbol{A}^{k+1})$, we get the desired result. Q.E.D.

Sketch of the Proof of Corollaries 1 *and* 3. We illustrate the derivations of bias formulas in Corollaries 1 and 3 by (16) for $p=2$. Let λ_i $(i=1, 2)$ be the distinct characteristic roots of $\boldsymbol{B}$. Then by the use of decomposition $\boldsymbol{B}=\boldsymbol{H\Lambda H}^{-1}$ where $\Lambda=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$ and $\boldsymbol{H}=\dfrac{1}{\lambda_1-\lambda_2}\begin{pmatrix}\lambda_1 & \lambda_2\\ 1 & 1\end{pmatrix}$, $\boldsymbol{F}$ in (9) is reduced to

$$\text{(A.18)}\qquad \boldsymbol{F}=\boldsymbol{H}\left[\sum_{k=0}^{\infty}\Lambda^k\begin{pmatrix}1\\-1\end{pmatrix}(1, -1)\boldsymbol{\Lambda}^k\right]\boldsymbol{H}'=\boldsymbol{H}\begin{pmatrix}\dfrac{1}{1-\lambda_1^2} & \dfrac{-1}{1-\lambda_1\lambda_2}\\ \dfrac{-1}{1-\lambda_1\lambda_2} & \dfrac{1}{1-\lambda_2^2}\end{pmatrix}\boldsymbol{H}',$$

and hence

$$\text{(A.19)}\qquad \boldsymbol{H}'\boldsymbol{F}^{-1}\boldsymbol{H}=\frac{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)^2}{(\lambda_1-\lambda_2)^2}\begin{pmatrix}\dfrac{1}{1-\lambda_2^2} & \dfrac{1}{1-\lambda_1\lambda_2}\\ \dfrac{1}{1-\lambda_1\lambda_2} & \dfrac{1}{1-\lambda_1^2}\end{pmatrix},$$

where we used the relation $\boldsymbol{H}^{-1}\boldsymbol{e}=(1, -1)'$ in (A.18). Then the first component of (14) becomes

$$\text{(A.20)}\qquad \boldsymbol{e}'\boldsymbol{H}^{-1\prime}\sum_{k=0}^{\infty}\left[\boldsymbol{\Lambda}^k\left(\sum_{i=1}^{2}\lambda_i^{k+1}\right)+\boldsymbol{\Lambda}^{2k+1}\right]\boldsymbol{H}'\boldsymbol{F}^{-1}\boldsymbol{e}$$

$$=\frac{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}{(\lambda_1-\lambda_2)^2}(1, -1)$$

$$\times\begin{pmatrix}\dfrac{2\lambda_1}{1-\lambda_1^2}+\dfrac{\lambda_2}{1-\lambda_1\lambda_2} & 0\\ 0 & \dfrac{\lambda_1}{1-\lambda_1\lambda_2}+\dfrac{2\lambda_2}{1-\lambda_2^2}\end{pmatrix}$$

$$\times\begin{pmatrix}\dfrac{1}{1-\lambda_2^2} & \dfrac{1}{1-\lambda_1\lambda_2}\\ \dfrac{1}{1-\lambda_1\lambda_2} & \dfrac{1}{1-\lambda_1^2}\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}=1-3\lambda_1\lambda_2\,.$$

Since $\lambda_1\lambda_2=-\beta_2$ and $\lambda_1+\lambda_2=\beta_1$, we obtain ABIAS$(\hat{\beta}_2)=-n^{-1}(1+3\beta_2)$. For ABIAS$(\hat{\beta}_1)$ in (16), we use $(0, 1)'$ instead of $\boldsymbol{e}$ in the last term of (A.20) and the remaining calculations are exactly the same. If two roots λ_1 and λ_2 are equal, the Jordan canonical form should be used in the decomposition of $\boldsymbol{B}=\boldsymbol{H\Lambda H}^{-1}$ where $\boldsymbol{H}=\begin{pmatrix}\lambda & \lambda+1\\ 1 & 1\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}$ and $\lambda=\lambda_1=\lambda_2$. The resulting formulas for ABIAS$(\hat{\beta}_1)$ and ABIAS$(\hat{\beta}_2)$ are the same in the above. Similarly, this method can be used to derive (10)–(12) and (15)–(17) with minor modifications. Q.E.D.

Acknowledgement

We are grateful to Professor T. W. Anderson and an anonymous referee of this journal for their comments on the earlier version of the paper.

YOKOHAMA NATIONAL UNIVERSITY
UNIVERSITY OF TOKYO

REFERENCES

[1] Anderson, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations, *Ann. Math. Statist.*, **30**, 676-687.
[2] Bhansali, R. J. (1981). Effects of not knowing the order of an autoregressive process on the mean squared error of prediction—I, *J. Amer. Statist. Ass.*, **76**, 588-597.
[3] Fuller, W. A. and Hasza, D. P. (1981). Properties of predictors for autoregressive time series, *J. Amer. Statist. Ass.*, **76**, 155-161.
[4] Groenewald, P. C. N. and de Waal, D. J. (1979). An asymptotic distribution for the coefficient in a multiple time series, *South African Statistics Journal*, **13**, 15-28.
[5] Kendall, M. G. (1954). Note on bias in the estimation of autocorrelation, *Biometrika*, **41**, 403-404.
[6] Sawa, T. (1978). The exact moments of the least squares estimator for the autoregressive model, *Journal of Econometrics*, **8**, 159-172.
[7] Shenton, L. R. and Johnson, W. L. (1965). Moments of a serial correlation coefficient, *J. R. Statist. Soc.*, B, **27**, No. 2, 308-320.
[8] White, J. S. (1961). Asymptotic expansion for the mean and variance of the serial correlation coefficient, *Biometrika*, **48**, 85-95.