Ann. Inst. Statist. Math.
Vol. 53, No. 1, 142–158 (2001)

SEQUENTIAL ESTIMATION FOR A FUNCTIONAL OF THE SPECTRAL DENSITY OF A GAUSSIAN STATIONARY PROCESS

TAKAYUKI SHIOHAMA AND MASANOBU TANIGUCHI

Department of Mathematical Science, Graduate School of Engineering Science, Osaka University, Machikaneyama-cho 1-3, Toyonaka, Osaka 560-8531, Japan,
e-mail: shiohama@sigmath.es.osaka-u.ac.jp; taniguti@sigmath.es.osaka-u.ac.jp

(Received April 10, 2000; revised August 12, 2000)

Abstract. Integral functional of the spectral density of stationary process is an important index in time series analysis. In this paper we consider the problem of sequential point and fixed-width confidence interval estimation of an integral functional of the spectral density for Gaussian stationary process. The proposed sequential point estimator is based on the integral functional replaced by the periodogram in place of the spectral density. Then it is shown to be asymptotically risk efficient as the cost per observation tends to zero. Next we provide a sequential interval estimator, which is asymptotically efficient as the width of the interval tends to zero. Finally some numerical studies will be given.

Key words and phrases: Stationary process, spectral density, sequential point estimation, sequential interval estimation, periodogram, integral functional.

1. Introduction

There have been a large number of articles and significant developments on both sequential point and interval estimation for i.i.d. random variables e.g., Robbins (1959) and Chow and Robbins (1965). Under the influence of these developments in the i.i.d. case, the literature on sequential estimation in time series emerged in the last decade. Sriram (1987) developed the sequential point and interval estimation for the mean of a stationary autoregressive model of order 1 (AR(1)) with unknown autoregressive coefficient. Then he showed the proposed sequential point and interval estimators are asymptotically efficient. Fakhre-Zakeri and Lee (1992) extends the results by Sriram (1987) to the case when the process concerned is a linear process. Further, Fakhre-Zakeri and Lee (1993) discussed the case of multivariate linear process. As for the other parameters (except for mean), Sriram (1988) considered the problem of sequential point estimation of the autoregressive coefficient of a stationary AR(1) model, and showed the asymptotic efficiency of the proposed sequential procedure. Lee (1994) generalized Sriram's (1988) results to the case of the autoregressive coefficients of AR(p) model.

In this paper we argue the sequential procedures for an integral functional of spectral density, which represents many important indices in time series analysis. Let $\boldsymbol{f}(\lambda)$ be the spectral density matrix of an m-dimensional Gaussian stationary process. We consider the problem of sequential point and interval estimation for an integral functional of the spectral density matrix of the form

$$\boldsymbol{\theta} = \begin{bmatrix} \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_1(\lambda)\boldsymbol{f}(\lambda)\}\mathrm{d}\lambda \\ \vdots \\ \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_q(\lambda)\boldsymbol{f}(\lambda)\}\mathrm{d}\lambda \end{bmatrix}$$

where $\boldsymbol{\psi}_j(\lambda)$'s are given $m \times m$-matrix-valued functions. The proposed sequential point estimator of $\boldsymbol{\theta}$ is based on

$$\hat{\boldsymbol{\theta}}_n = \begin{bmatrix} \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_1(\lambda)\boldsymbol{I}_n(\lambda)\}\mathrm{d}\lambda \\ \vdots \\ \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_q(\lambda)\boldsymbol{I}_n(\lambda)\}\mathrm{d}\lambda \end{bmatrix}$$

where $\boldsymbol{I}_n(\lambda)$ is the periodogram matrix of n-consecutive observations from the process concerned. Then we show that it is asymptotically risk efficient as the cost per observation tends to zero. Also a sequential interval estimator is given, and is shown to be asymptotically efficient as the width of the interval tends to zero. Some numerical studies will be given in Section 4. They agree with the theoretical results.

2. Sequential procedures for multivariate Gaussian processes

Let $\{\boldsymbol{X}_t; t = 0, \pm 1, \ldots\}$ be an m-dimensional Gaussian stationary process with mean $E(\boldsymbol{X}_t) = \boldsymbol{0}$, spectral density matrix $\boldsymbol{f}(\lambda)$, and covariance function $\boldsymbol{\Gamma}(l) = E(\boldsymbol{X}_t\boldsymbol{X}'_{t+l})$. The covariance matrices satisfy

$$\sum_{l=-\infty}^{\infty} |l| \|\boldsymbol{\Gamma}(l)\| < \infty, \tag{2.1}$$

where $\|\boldsymbol{\Gamma}(l)\|$ is the square root of maximum eigenvalue of $\boldsymbol{\Gamma}(l)\boldsymbol{\Gamma}(l)'$. Consider an integral functional of $\boldsymbol{f}(\lambda)$ of the form

$$\boldsymbol{\theta} = \begin{bmatrix} \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_1(\lambda)\boldsymbol{f}(\lambda)\}\mathrm{d}\lambda \\ \vdots \\ \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_q(\lambda)\boldsymbol{f}(\lambda)\}\mathrm{d}\lambda \end{bmatrix} \tag{2.2}$$

where $\boldsymbol{\psi}_j(\lambda)$, $j = 1, \ldots, q$ be $m \times m$ matrix valued continuous functions on $[-\pi, \pi]$ such that $\boldsymbol{\psi}_j(\lambda) = \boldsymbol{\psi}_j(\lambda)^*$. To estimate $\boldsymbol{\theta}$ we use

$$\hat{\boldsymbol{\theta}}_n = \begin{bmatrix} \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_1(\lambda)\boldsymbol{I}_n(\lambda)\}\mathrm{d}\lambda \\ \vdots \\ \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_q(\lambda)\boldsymbol{I}_n(\lambda)\}\mathrm{d}\lambda \end{bmatrix} \tag{2.3}$$

where $\boldsymbol{I}_n(\lambda)$ is the periodogram matrix defined by

$$\boldsymbol{I}_n(\lambda) = \frac{1}{2\pi n}\left\{\sum_{t=1}^{n} \boldsymbol{X}_t e^{it\lambda}\right\}\left\{\sum_{t=1}^{n} \boldsymbol{X}_t e^{it\lambda}\right\}^* . \tag{2.4}$$

Denote by $f_{ab}(\lambda)$, $I_n^{(a,b)}(\lambda)$ and $\Gamma_{ab}(l)$ the (a,b)-th element of $\boldsymbol{f}(\lambda)$, $\boldsymbol{I}_n(\lambda)$ and $\boldsymbol{\Gamma}(l)$, respectively.

Suppose that one wishes to estimate $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}_n$, given n consecutive observations, subject to the loss function

$$L_n = (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})' \boldsymbol{Q} (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + cn \tag{2.5}$$

where $\boldsymbol{Q}$ is a given $q \times q$ positive definite matrix and $c > 0$ is the cost per observation. It is known that

$$nE(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})' \to \boldsymbol{V} \quad \text{as} \quad n \to \infty \tag{2.6}$$

where the (i,j) element of $\boldsymbol{V}$ is

$$4\pi \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_i(\lambda)\boldsymbol{f}(\lambda)\boldsymbol{\psi}_j(\lambda)\boldsymbol{f}(\lambda)\}\mathrm{d}\lambda \tag{2.7}$$

(see, for example, Hosoya and Taniguchi (1982)). The associated risk is

$$R_n = EL_n = n^{-1}\,\mathrm{tr}\,\boldsymbol{QV} + cn + o(n^{-1}). \tag{2.8}$$

If $\boldsymbol{V}$ is known, (2.8) is approximately minimized by the best fixed sample size

$$n_0 \simeq [\mathrm{tr}\,\boldsymbol{QV}/c]^{1/2} \tag{2.9}$$

with corresponding risk

$$R_{n_0} \simeq 2cn_0. \tag{2.10}$$

However, when $\boldsymbol{V}$ is unknown, as is typically the case, n_0 cannot be used and there is no fixed sample size procedure that will achieve the risk (2.10).

Motivated by (2.9), we shall consider the following stopping rule,

$$N = \inf\{n \geq m_0 : n \geq c^{-1/2}[(\mathrm{tr}\,\boldsymbol{Q}\hat{\boldsymbol{V}}_n)^{1/2} + n^{-h}]\}, \tag{2.11}$$

where m_0 is a predetermined and fixed initial sample size, $h > 0$ is an arbitrary positive constant, and $\hat{\boldsymbol{V}}_n$ is an appropriate consistent estimator of $\boldsymbol{V}$. To develop the asymptotic theory we need the following assumption.

ASSUMPTION 1. (i) There exists $\rho \in (0,1)$ such that the covariance function of $\boldsymbol{X}_t$ satisfies

$$\Gamma_{ab}(j) = O(\rho^{|j|}), \qquad a,b = 1,\ldots,m. \tag{2.12}$$

(ii) The weight function $\boldsymbol{\psi}_j(\lambda)$ is expressed as

$$\boldsymbol{\psi}_j(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \boldsymbol{\eta}_j(k) e^{-ik\lambda}, \tag{2.13}$$

where $\boldsymbol{\eta}_j(k)$'s are known $m \times m$ matrices and the (a,b) elements satisfy

$$\eta_j^{(a,b)}(k) = O(\beta^{|k|}), \qquad a,b = 1,\ldots,m \quad \text{for some } \beta \in (0,1). \tag{2.14}$$

Let $\{h_n\} = [(\log n)^{1+\xi}]$, $\xi \in (0,1)$, where $[x]$ is the greatest integer less than or equal to x. Then

$$\sqrt{n}\rho^{h_n} \to 0 \quad \text{and} \quad h_n/(n^\delta) \to 0 \quad \text{for all } \delta > 0. \tag{2.15}$$

We can check $\sqrt{n}\rho^{h_n} \to 0$ by setting $\rho = e^{-c}$, $(c > 0)$, and writing $\sqrt{n}\rho^{h_n} = n^{1/2}e^{-c(\log n)(\log n)^\xi} = n^{1/2}\{n^{-c}\}^{(\log n)^\xi} = n^{1/2-c(\log n)^\xi}$. From Assumption 1, it is easily seen that

$$\begin{aligned} V_{ij} &= 4\pi\int_{-\pi}^{\pi} \operatorname{tr}\{\boldsymbol{\psi}_i(\lambda)\boldsymbol{f}(\lambda)\boldsymbol{\psi}_j(\lambda)\boldsymbol{f}(\lambda)\}\mathrm{d}\lambda \\ &= 4\pi\int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\right)^4 \operatorname{tr}\left\{\sum_{k_1,k_2,k_3,k_4=-\infty}^{\infty} \boldsymbol{\Gamma}(k_1)\boldsymbol{\eta}_i(k_2)\boldsymbol{\Gamma}(k_3)\boldsymbol{\eta}_j(k_4) \times e^{-i(k_1+k_2+k_3+k_4)\lambda}\right\}\mathrm{d}\lambda \\ &= \frac{1}{2\pi^2}\operatorname{tr}\left\{\sum_{k_1,k_2,k_3=-\infty}^{\infty} \boldsymbol{\Gamma}(k_1)\boldsymbol{\eta}_i(k_2)\boldsymbol{\Gamma}(k_3)\boldsymbol{\eta}_j(-k_1-k_2-k_3)\right\}. \end{aligned} \tag{2.16}$$

Then it is natural to estimate V_{ij} by

$$\hat{V}_{n,ij} = \frac{1}{2\pi^2}\operatorname{tr}\left\{\sum_{k_1,k_2,k_3=-h_n}^{h_n} \hat{\boldsymbol{\Gamma}}_n(k_1)\boldsymbol{\eta}_i(k_2)\hat{\boldsymbol{\Gamma}}_n(k_3)\boldsymbol{\eta}_j(-k_1-k_2-k_3)\right\} \tag{2.17}$$

where

$$\hat{\boldsymbol{\Gamma}}_n(l) = \frac{1}{n}\sum_{t=1}^{n-l}\boldsymbol{X}_t\boldsymbol{X}'_{t+l}. \tag{2.18}$$

The proposed sequential estimator of $\boldsymbol{\theta}$ is $\hat{\boldsymbol{\theta}}_N$ and its risk is

$$R_N = \operatorname{tr}\boldsymbol{Q}E(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})' + cEN. \tag{2.19}$$

It will be proved in Section 5 that this sequential procedure is "asymptotically risk efficient" in Robbins' (1959) sense, i.e., $R_N/R_{n_0} \to 1$ as $c \to 0$. The general nature of a sequential estimation procedure is to consider an optimum or a reasonable fixed-sample-size for known variance structure, and then to make appropriate modification to motivate a sequential procedure. Exact properties of such a sequential procedure are often difficult to study, though asymptotic properties as $c \to 0$ can be established by using sophisticated statistical tools.

Confidence region

Consider the problem of finding a confidence set for $\boldsymbol{\theta}$ with prescribed size $2d$ and coverage probability $1-\alpha$, $0 < \alpha < 1$. First note that,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow[\mathcal{L}]{} \mathcal{N}(\boldsymbol{0}, \boldsymbol{V}) \quad \text{as} \quad n \to \infty, \tag{2.20}$$

where the (i,j)-th entry of $\boldsymbol{V}$ is as in (2.7). Based on (2.20), if the covariance matrix $\boldsymbol{V}$ is known, the following ellipsoid is an asymptotic $(1-\alpha)$ confidence set for $\boldsymbol{\theta}$,

$$\{\boldsymbol{\theta} \in \boldsymbol{R}^q : n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})'\boldsymbol{V}^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \le \chi^2_{1-\alpha}(q)\}, \tag{2.21}$$

where $\chi^2_{1-\alpha}(q)$ is the upper $1-\alpha$ point of a chi-square distribution with q degrees of freedom. However, when $\boldsymbol{V}$ is unknown, there is no fixed sample size procedure to construct confidence set of fixed size with guaranteed coverage probability.

Let $\lambda_{\max}(\boldsymbol{V})$ and $\lambda_{\max}(\hat{\boldsymbol{V}}_n)$ denote the largest eigenvalue of $\boldsymbol{V}$ and $\hat{\boldsymbol{V}}_n$, respectively. To construct a fixed accuracy confidence set with prescribed coverage probability we adopt the sequential procedure (T, R_T) in which the stopping rule T is defined by

$$T = \inf\{n \ge m_0 : (\lambda_{\max}(\hat{\boldsymbol{V}}_n) + n^{-h}) \le d^2 n/\chi^2_{1-\alpha}(q)\} \tag{2.22}$$

and the terminal decision rule is

$$R_T = \{\boldsymbol{\theta} \in \boldsymbol{R}^q : (\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \le d^2\}. \tag{2.23}$$

We state the main results in Section 3, and give the proofs in Section 5.

3. Asymptotic results for multivariate Gaussian processes

We can provide the following theorems which describe asymptotic performances of the sequential procedures with stopping rules N and T defined in (2.11) and (2.22), respectively.

THEOREM 1. *Suppose that Assumption* 1 *holds. Then as* $c \to 0$,

$$N/n_0 \to 1 \quad a.s., \tag{3.1}$$

$$E|N/n_0 - 1| \to 0 \tag{3.2}$$

and

$$R_N/R_{n_0} \to 1. \tag{3.3}$$

Theorem 1 implies that the sequential point estimator is asymptotically risk efficient as $c \to 0$. The following theorem shows the asymptotic normality of $\hat{\boldsymbol{\theta}}_N$.

THEOREM 2. *Under Assumption* 1,

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}) \xrightarrow[\mathcal{L}]{} N(\boldsymbol{0}, \boldsymbol{V}), \quad as \quad c \to 0. \tag{3.4}$$

The following theorem states that the fixed accuracy confidence set procedure is asymptotically efficient with prescribed coverage probability.

THEOREM 3. *Suppose that Assumption* 2 *holds. Then,*

$$\lim_{d\to 0} T/k_0 = 1 \quad a.s., \tag{3.5}$$

$$\lim_{d\to 0} P(\boldsymbol{\theta} \in R_T) \ge 1-\alpha, \tag{3.6}$$

and

$$\lim_{d\to 0} E(T/k_0) = 1, \tag{3.7}$$

where $k_0 = [\lambda_{\max}(\boldsymbol{V}) \cdot \chi^2_{1-\alpha}(q) d^{-2}] + 1$.

Table 1. Sequential point estimation for $\theta = \Gamma(0)$.

$\Gamma(0) = 1.961$

c	n_0	N	N/n_0	$\hat{\theta}_N$
1	48	42	0.875	1.755
0.1	150	135	0.900	1.769
0.01	474	415	0.876	1.807
0.005	671	650	0.969	1.958
0.003	866	864	0.998	2.024

Table 2. Sequential point estimation for $\theta = \Gamma(1)$.

$\Gamma(1) = 1.373$

c	n_0	N	N/n_0	$\hat{\theta}_N$
1	46	40	0.870	1.086
0.1	146	131	0.897	1.231
0.01	460	390	0.848	1.221
0.005	650	517	0.795	1.141
0.003	839	817	0.974	1.359

4. Numerical studies

In this section we give some numerical results to illustrate the sequential procedures described in the previous sections. Example 1 deals with a scalar process, and Example 2 treats a multivariate process. For the scalar case, we consider the following autoregressive process (AR(1))

$$X_t = 0.7X_{t-1} + \varepsilon_t \tag{4.1}$$

where $\varepsilon_t \sim$ i.i.d. $\mathcal{N}(0,1)$. The spectral density is given by $f(\lambda) = (2\pi)^{-1}|1 - 0.7e^{-i\lambda}|^{-2}$. We generate $X_1, \ldots, X_{1000}$ from (4.1), (i.e., $n = 1000$), and choose $h_n = 18$.

Example 1. (Sequential estimation for the autocovariance function) The autocovariance function $\Gamma(l)$ of $\{X_t\}$ can be expressed as $\int_{-\pi}^{\pi} \psi(\lambda) f(\lambda) \mathrm{d}\lambda$ with $\psi(\lambda) = e^{il\lambda}$. The autocovariances at lag 0 and 1 of (4.1) are 1.961 and 1.373, respectively. Table 1 summarizes the simulation results of the sequential point estimation for $\theta = \Gamma(0)$, and Table 2 shows those of $\Gamma(1)$. It is clear from these tables that the stopping rule N approaches the best fixed-sample-size n_0 as the cost per observation tends to zero.

We also investigated the problem of finding a 90 % confidence interval for $\Gamma(0)$ and $\Gamma(1)$ with prescribed width $2d$. Tables 3 and 4 summarize the results of estimating $\Gamma(0)$ and $\Gamma(1)$ with various width $d = 0.8,\ 0.6,\ 0.4, 0.3$ and 0.25. It is seen that T/k_0 tends to 1 as width d approaches zero.

As an illustration of the case of multivariate process, we consider the following vector autoregressive process (VAR(1))

$$\begin{aligned} X_{1t} &= 0.7X_{1,t-1} + \varepsilon_{1t} \\ X_{2t} &= 0.3X_{1,t-1} + 0.6X_{2,t-1} + \varepsilon_{2t} \end{aligned} \tag{4.2}$$

Table 3. 90% confidence interval of sequential estimation for $\theta = \Gamma(0)$.

$\Gamma(0) = 1.961$

d	k_0	T	T/k_0	$[\hat{\theta}_T - d, \hat{\theta}_T + d]$
0.8	95	74	0.779	[1.019, 2.619]
0.6	169	135	0.799	[1.169, 2.369]
0.4	380	240	0.632	[1.240, 2.040]
0.3	676	490	0.725	[1.527, 2.127]
0.25	975	942	0.966	[1.738, 2.238]

Table 4. 90% confidence interval of sequential estimation for $\theta = \Gamma(1)$.

$\Gamma(1) = 1.373$

d	k_0	T	T/k_0	$[\hat{\theta}_T - d, \hat{\theta}_T + d]$
0.8	90	70	0.778	[0.422, 2.022]
0.6	159	128	0.805	[0.640, 1.840]
0.4	357	193	0.541	[0.616, 1.416]
0.3	635	439	0.691	[0.892, 1.492]
0.25	914	880	0.963	[1.133, 1.633]

Table 5. Sequential point estimation for $\text{vec}\{\mathbf{\Gamma}(0)\}$.

$\boldsymbol{\theta} \equiv \text{vec}\{\mathbf{\Gamma}(0)\} = (1.961, 1.572, 1.572, 1.941)'$

c	n_0	N	N/n_0	$\hat{\boldsymbol{\theta}}_N$
1	112	105	0.938	$(2.037, 1.587, 1.587, 1.745)'$
0.1	353	372	1.054	$(2.105, 1.729, 1.729, 2.068)'$
0.05	499	531	1.064	$(2.089, 1.718, 1.450, 2.051)'$
0.01	1116	1064	0.953	$(1.927, 1.499, 1.499, 1.814)'$
0.005	1578	1558	0.987	$(1.928, 1.508, 1.508, 1.862)'$

where $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a sequence of i.i.d. $N(\mathbf{0}, \mathbf{\Sigma})$ random variables with

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}.$$

We generate 2000 observations from (4.2) and choose $h_n = 10$.

Example 2. (Sequential estimation for the autocovariance functions) Since $\boldsymbol{\theta} = \text{vec}\{\mathbf{\Gamma}(l)\}$ is expressed as an integral functional of the spectral density matrix $\boldsymbol{f}(\lambda)$, we study the sequential point and interval estimation for $\text{vec}\{\mathbf{\Gamma}(0)\}$ and $\text{vec}\{\mathbf{\Gamma}(1)\}$, numerically. The results obtained are summarized in Tables 5–8, and they agree with the theoretical results for the asymptotic efficiency of Theorems 1 and 3 in Section 3.

As an application of the results in Section 2, we can consider the following problem of misspecified prediction. Suppose $\{\boldsymbol{X}_t\}$ is a p-vector autoregressive process of order 1

Table 6. 90% confidence interval of sequential estimation for vec$\{\mathbf{\Gamma}(0)\}$.

$\boldsymbol{\theta} \equiv \text{vec}\{\mathbf{\Gamma}(0)\} = (1.961, 1.572, 1.572, 1.941)'$				
d	k_0	T	T/k_0	$\hat{\boldsymbol{\theta}}_T$
1.1	648	593	0.915	$(1.907, 1.512, 1.512, 1.783)'$
1.0	784	731	0.932	$(1.925, 1.543, 1.543, 1.804)'$
0.9	967	987	1.021	$(1.956, 1.557, 1.557, 1.870)'$
0.8	1224	1277	1.043	$(1.963, 1.551, 1.551, 1.855)'$
0.7	1599	1572	0.983	$(1.920, 1.521, 1.521, 1.849)'$

Table 7. Sequential point estimation for vec$\{\mathbf{\Gamma}(1)\}$.

$\boldsymbol{\theta} \equiv \text{vec}\{\mathbf{\Gamma}(1)\} = (1.373, 1.100, 1.531, 1.636)'$				
c	n_0	N	N/n_0	$\hat{\boldsymbol{\theta}}_N$
1	106	98	0.925	$(1.597, 1.189, 1.611, 1.526)'$
0.1	334	326	0.976	$(1.519, 1.154, 1.661, 1.679)'$
0.05	472	502	1.064	$(1.600, 1.249, 1.750, 1.785)'$
0.01	1055	947	0.898	$(1.339, 1.022, 1.430, 1.459)'$
0.005	1493	1419	0.950	$(1.360, 1.037, 1.459, 1.537)'$

Table 8. 90% confidence interval of sequential estimation for vec$\{\mathbf{\Gamma}(1)\}$.

$\boldsymbol{\theta} \equiv \text{vec}\{\mathbf{\Gamma}(0)\} = (1.373, 1.100, 1.531, 1.636)'$				
d	k_0	T	T/k_0	$\hat{\boldsymbol{\theta}}_T$
1.1	583	537	0.921	$(1.270, 1.029, 1.405, 1.504)'$
1.0	706	619	0.877	$(1.351, 1.065, 1.436, 1.479)'$
0.9	871	773	0.887	$(1.360, 1.089, 1.489, 1.530)'$
0.8	1102	1076	0.976	$(1.345, 1.089, 1.481, 1.566)'$
0.7	1438	1445	1.005	$(1.324, 1.072, 1.454, 1.558)'$

(VAR(1)). Then the best linear predictor of $\boldsymbol{X}_t$ based on $\boldsymbol{X}_{t-1}, \boldsymbol{X}_{t-2}, \ldots$, is given by

$$\mathbf{\Gamma}(1)\mathbf{\Gamma}(0)^{-1}\boldsymbol{X}_{t-1}. \tag{4.3}$$

Although $\boldsymbol{X}_t$ is actually VAR(1), we consider the case when the fitted model is incorrectly specified by the following VMA(1) process

$$\boldsymbol{X}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{A}\boldsymbol{\varepsilon}_{t-1} \tag{4.4}$$

where $\boldsymbol{A}$ is a $p \times p$-matrix, and $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $N(\boldsymbol{0}, \boldsymbol{I})$, ($\boldsymbol{I}$; the $p \times p$-identity matrix). For the model (4.4), it is easily seen that

$$\mathbf{\Gamma}(0) = \boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}', \qquad \mathbf{\Gamma}(1) = \boldsymbol{A}$$

then the misspecified prediction error for the predictor (4.3) is

$$\text{tr}\, E[\{\boldsymbol{X}_t - \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{X}_{t-1}\}\{\boldsymbol{X}_t - \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{X}_{t-1}\}'] \tag{4.5}$$

$$= \operatorname{tr}\left[\int_{-\pi}^{\pi} \left\{I - A(I + AA')^{-1}e^{i\lambda}\right\}\left\{I - A(I + AA')^{-1}e^{i\lambda}\right\}^* f(\lambda)\mathrm{d}\lambda\right].$$

If we set $\psi(\lambda) = \left\{I - A(I + AA')^{-1}e^{i\lambda}\right\}\left\{I - A(I + AA')^{-1}e^{i\lambda}\right\}^*$, then we can discuss the sequential estimation procedure for (4.5).

5. Proofs

In this section we will give the proofs of Theorems 1–3. The proofs of theorems are based on the following lemmas. Write $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)'$ and $\hat{\boldsymbol{\theta}}_n = (\hat{\theta}_{n,1}, \ldots, \hat{\theta}_{n,q})'$.

LEMMA 1. *For any integer $r \geq 2$, we have*

$$E(\hat{\theta}_{n,j} - \theta_j)^{2r} = O(n^{-r}), \qquad j = 1, \ldots, q. \tag{5.1}$$

PROOF. From the inequality $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ we have

$$\begin{aligned} E(\hat{\theta}_{n,j} - \theta_j)^{2r} &= E(\hat{\theta}_{n,j} - E\hat{\theta}_{n,j} + E\hat{\theta}_{n,j} - \theta_j)^{2r} \\ &\leq 2^{2r-1}\{E(\hat{\theta}_{n,j} - E\hat{\theta}_{n,j})^{2r} + E(E\hat{\theta}_{n,j} - \theta_j)^{2r}\} \\ &= 2^{2r-1}\left[L_1 + L_2\right] \qquad \text{(say)}. \end{aligned} \tag{5.2}$$

Note the following identity

$$E(X_1 X_2 \cdots X_{2r}) = \sum_{\nu = \{\nu_1, \ldots, \nu_p\}} \operatorname{cum}\{X_{j_1}; j_1 \in \nu_1\} \times \cdots \times \operatorname{cum}\{X_{j_p}; j_p \in \nu_p\} \tag{5.3}$$

where the summation is over all partitions $(\nu_1, \nu_2, \ldots, \nu_p)$ of $\{1, 2, \ldots, 2r\}$ (see (8.11) of Brillinger (1969)). Using this identity, we have

$$E(\hat{\theta}_{n,j} - E\hat{\theta}_{n,j})^{2r} = \sum \operatorname{cum}^{(j_1)}\{\hat{\theta}_{n,j} - E\hat{\theta}_{n,j}\} \times \cdots \times \operatorname{cum}^{(j_p)}\{\hat{\theta}_{n,j} - E\hat{\theta}_{n,j}\}, \tag{5.4}$$

where the summation is over $(j_1, \ldots, j_p)$ satisfying $j_1 + \cdots + j_p = 2r$. Using the argument of the proof for *Theorem* 7.6.1 of Brillinger (1981) and *Lemma* A.3.3 of Hosoya and Taniguchi (1982) we have

$$E\hat{\theta}_{n,j} = \theta_j + O(n^{-1}) \tag{5.5}$$

and

$$\operatorname{cum}^{(J)}\{\hat{\theta}_{n,j} - E\hat{\theta}_{n,j}, \ldots, \hat{\theta}_{n,j} - E\hat{\theta}_{n,j}\} = O(n^{-J+1}). \tag{5.6}$$

From (5.5) it follows that

$$L_2 = O(n^{-2r}). \tag{5.7}$$

In view of (5.6) we can see that the main order term in (5.4) is $\operatorname{cum}^{(2)}(\hat{\theta}_{n,j}) \ldots \operatorname{cum}^{(2)}(\hat{\theta}_{n,j})$. Since

$$\operatorname{cum}^{(2)}(\hat{\theta}_{n,j}) = \operatorname{var}(\hat{\theta}_{n,j}) = n^{-1}V_{jj} + o(1) = O(n^{-1}), \tag{5.8}$$

we observe

$$L_1 = E(\hat{\theta}_{n,j} - E\hat{\theta}_{n,j})^{2r} = O(n^{-r}),$$

which implies, together with (5.2) and (5.7), the desired result. □

LEMMA 2. *For any integer $r \geq 2$, we have*

$$E \max_{1\leq i,j\leq q} |\hat{V}_{n,ij} - V_{ij}|^{2r} = O(n^{-r} \cdot h_n^{6r}). \tag{5.9}$$

PROOF. As in Lemma 1, we have

$$\begin{aligned} E|\hat{V}_{n,ij} - V_{ij}|^{2r} &= E|\hat{V}_{n,ij} - E\hat{V}_{n,ij} + E\hat{V}_{n,ij} - V_{ij}|^{2r} \\ &\leq 2^{2r-1}\{E|\hat{V}_{n,ij} - E\hat{V}_{n,ij}|^{2r} + E|E\hat{V}_{n,ij} - V_{ij}|^{2r}\} \\ &= 2^{2r-1}\,[M_1 + M_2], \qquad \text{(say)}. \end{aligned} \tag{5.10}$$

Using the identity (5.3), we have

$$\begin{aligned} E|\hat{V}_{n,ij} - E\hat{V}_{n,ij}|^{2r} = &\sum_{j_1+\cdots+j_p=2r} \mathrm{cum}^{(j_1)}\{\hat{V}_{n,ij} - E\hat{V}_{n,ij}\} \times \cdots \\ &\times \mathrm{cum}^{(j_p)}\{\hat{V}_{n,ij} - E\hat{V}_{n,ij}\}. \end{aligned} \tag{5.11}$$

From (2.17) and (5.5) it is seen that

$$\begin{aligned} E\hat{V}_{n,ij} = &\frac{1}{2\pi^2}\,\mathrm{tr}\left\{\sum_{k_1,k_2,k_3=-h_n}^{h_n} \mathbf{\Gamma}(k_1)\boldsymbol{\eta}_i(k_2)\mathbf{\Gamma}(k_3)\boldsymbol{\eta}_j(-k_1-k_2-k_3)\right\} \\ &+O(n^{-1}h_n^3). \end{aligned} \tag{5.12}$$

Noting (2.16), (5.12), Assumption 1 and $h_n = [(\log n)]^{1+\xi}$ we can see that

$$E\hat{V}_{n,ij} = V_{ij} + O(n^{-1}h_n^3),$$

which implies

$$M_2 = O(n^{-2r}h_n^{6r}). \tag{5.13}$$

Next we turn to evaluate the J-th order joint cumulant of $\hat{V}_{n,ij} - E\hat{V}_{n,ij}$

$$\begin{aligned} &\mathrm{cum}^{(J)}\{\hat{V}_{n,ij} - E\hat{V}_{n,ij}, \ldots, \hat{V}_{n,ij} - E\hat{V}_{n,ij}\} \\ &= \left(\frac{1}{2\pi^2}\right)^J \sum_{|k_1^{(1)}|\leq h_n} \sum_{|k_2^{(1)}|\leq h_n} \sum_{|k_3^{(1)}|\leq h_n} \cdots \sum_{|k_1^{(J)}|\leq h_n} \sum_{|k_2^{(J)}|\leq h_n} \sum_{|k_3^{(J)}|\leq h_n} \\ &\times \sum\nolimits^{*} \mathrm{cum}\{\hat{\Gamma}_{a^{(1)}b^{(1)}}(k_1^{(1)})\hat{\Gamma}_{c^{(1)}d^{(1)}}(k_3^{(1)}), \ldots, \hat{\Gamma}_{a^{(J)}b^{(J)}}(k_1^{(J)})\hat{\Gamma}_{c^{(J)}d^{(J)}}(k_3^{(J)})\} \\ &\times \eta_i^{(b^{(1)},c^{(1)})}(k_2^{(1)}) \cdots \eta_i^{(b^{(J)},c^{(J)})}(k_2^{(J)}) \\ &\times \eta_j^{(d^{(1)},a^{(1)})}(-k_1^{(1)} - k_2^{(1)} - k_3^{(1)}) \cdots \eta_j^{(d^{(J)},a^{(J)})}(-k_1^{(J)} - k_2^{(J)} - k_3^{(J)}), \end{aligned} \tag{5.14}$$

where $\sum^*$ is the summation for $a^{(1)}, b^{(1)}, c^{(1)}, d^{(1)}, \ldots, a^{(J)}, b^{(J)}, c^{(J)}, d^{(J)} = 1, \ldots, q$. Now

$$\begin{aligned} &\mathrm{cum}\{\hat{\Gamma}_{a^{(1)}b^{(1)}}(k_1^{(1)})\hat{\Gamma}_{c^{(1)}d^{(1)}}(k_3^{(1)}), \ldots, \hat{\Gamma}_{a^{(J)}b^{(J)}}(k_1^{(J)})\hat{\Gamma}_{c^{(J)}d^{(J)}}(k_3^{(J)})\} \\ &= \mathrm{cum}[Y\{1^{(1)}\}Y\{3^{(1)}\}, \ldots, Y\{1^{(J)}\}Y\{3^{(J)}\}] \qquad \text{(say)} \\ &= \sum_{\nu} \mathrm{cum}\{Y\{i_1\} : i_1 \in \nu_1\} \times \cdots \times \mathrm{cum}\{Y\{i_p\} : i_p \in \nu_p\} \end{aligned} \tag{5.15}$$

where the summation extends over all indecomposable partition $(\nu_1, \ldots, \nu_p)$ of the table

$$\begin{matrix}(1^{(1)}, 3^{(1)})\\ \vdots \\ (1^{(J)}, 3^{(J)})\end{matrix}$$

(see Brillinger (1981)). From the *Theorem* 5.1 of Brillinger (1969), we have

$$\text{cum}\{\hat{\Gamma}_{ab}(k_1^{(1)}), \ldots, \hat{\Gamma}_{ab}(k_1^{(J)})\} = O(n^{-J+1})$$

and

$$\begin{aligned}\text{cov}\{\hat{\Gamma}_{a_1b_1}(k_1^{(u_1)}), \hat{\Gamma}_{a_2b_2}(k_1^{(u_2)})\} = \frac{2\pi}{n}\Bigg[\int_{-\pi}^{\pi}\{e^{i\lambda(k_1^{(u_1)}-k_1^{(u_2)})}f_{a_1a_2}(\lambda)f_{b_1b_2}(-\lambda)\mathrm{d}\lambda \\ + \int_{-\pi}^{\pi} e^{i\lambda(k_1^{(u_1)}+k_1^{(u_2)})}\}f_{a_1b_2}(\lambda)f_{b_1a_2}(-\lambda)\mathrm{d}\lambda\Bigg] \\ + o(n^{-1}).\end{aligned}$$

Then it is not difficult to show that the main order of (5.15) is that of

$$\begin{aligned}E\{\hat{\Gamma}_{a^{(u_1)}b^{(u_1)}}(k_1^{(u_1)})\}E\{\hat{\Gamma}_{a^{(u_2)}b^{(u_2)}}(k_3^{(u_2)})\}\text{cum}\{\hat{\Gamma}_{a^{(l_1)}b^{(l_1)}}(k_1^{(l_1)}), \hat{\Gamma}_{a^{(l'_1)}b^{(l'_1)}}(k_1^{(l'_1)})\} \\ \times \cdots \times \text{cum}\{\hat{\Gamma}_{a^{(l_{J-1})}b^{(l_{J-1})}}(k_1^{(l_{J-1})}), \hat{\Gamma}_{a^{(l'_{J-1})}b^{(l'_{J-1})}}(k_1^{(l'_{J-1})})\},\end{aligned}$$

which is of order $O(n^{-J+1})$. Hence $\text{cum}^{(J)}\{\hat{V}_{n,ij}, \ldots, \hat{V}_{n,ij}\} = O(n^{-J+1} \cdot h_n^{3J})$. Therefore main order term in M_1 is $\text{cum}^{(2)}\{\hat{V}_{n,ij}\} \times \cdots \times \text{cum}^{(2)}\{\hat{V}_{n,ij}\}$. Then, from (5.11) we have

$$E|\hat{V}_{n,ij} - V_{ij}|^{2r} = O(n^{-1} \cdot h_n^6)^r = O(n^{-r} \cdot h_n^{6r}).$$

Using this relation,

$$\begin{aligned}E \max_{1\le i,j\le q} |\hat{V}_{n,ij} - V_{ij}|^{2r} &\le \sum_{i,j=1}^{m} E|\hat{V}_{n,ij} - V_{ij}|^{2r} \\ &= \sum_{i,j=1}^{m} O(n^{-r} \cdot h_n^{6r}) \\ &= O(n^{-r} \cdot h_n^{6r}),\end{aligned}$$

which completes the proof. □

LEMMA 3. *For any integer $r \ge 2$, we have*

(5.16) $$\|\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n - \operatorname{tr} \boldsymbol{Q}\boldsymbol{V}\|_{2r} = O(n^{-1/2} \cdot h_n^3)$$

and

(5.17) $$\|\lambda_{\max}(\hat{\boldsymbol{V}}_n) - \lambda_{\max}(\boldsymbol{V})\|_{2r} = O(n^{-1/2} \cdot h_n^3).$$

In particular, for any $\varepsilon > 0$,

(5.18) $$P(|\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n - \operatorname{tr} \boldsymbol{Q}\boldsymbol{V}| > \varepsilon) = O(n^{-r/2} \cdot h_n^{3r})$$

and

(5.19) $$P(|\lambda_{\max}(\hat{\boldsymbol{V}}_n) - \lambda_{\max}(\boldsymbol{V})| > \varepsilon) = O(n^{-r/2} \cdot h_n^{3r}).$$

PROOF. We can prove (5.16) and (5.18) in the same way as in Lemma 2. To prove (5.17) we use the following result (see Anderson (1984), p. 354): For symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, $\lambda_{\max}(\boldsymbol{A}+\boldsymbol{B}) \leq \lambda_{\max}(\boldsymbol{A}) + \lambda_{\max}(\boldsymbol{B})$. From the above inequality we obtain $\lambda_{\max}(\hat{\boldsymbol{V}}_n) - \lambda_{\max}(\boldsymbol{V}) \leq \lambda_{\max}(\hat{\boldsymbol{V}}_n - \boldsymbol{V})$ and $-\lambda_{\max}(\boldsymbol{V} - \hat{\boldsymbol{V}}_n) \leq \lambda_{\max}(\hat{\boldsymbol{V}}_n) - \lambda_{\max}(\boldsymbol{V})$. Then, note that

$$\begin{aligned} |\lambda_{\max}(\hat{\boldsymbol{V}}_n - \boldsymbol{V})| &= \left|\max_{\boldsymbol{a}'\boldsymbol{a}=1} \boldsymbol{a}'(\hat{\boldsymbol{V}}_n - \boldsymbol{V})\boldsymbol{a}\right| \\ &\leq O\left\{\max_{1\leq i,j\leq q} |\hat{V}_{n,ij} - V_{ij}|\right\} \end{aligned}$$

and, similarly

$$|-\lambda_{\max}(\boldsymbol{V} - \hat{\boldsymbol{V}}_n)| \leq O\left\{\max_{1\leq i,j\leq q} |\hat{V}_{n,ij} - V_{ij}|\right\}.$$

In view of Lemma 2 we get (5.17). (5.19) follows from the Markov inequality and (5.17). □

Let

$$n_1 = [c^{-1/2(1+h)}], \quad n_2 = [n_0(1-\varepsilon)] \quad \text{and} \quad n_3 = [n_0(1+\varepsilon)], \quad \text{for } 0<\varepsilon<1. \tag{5.20}$$

Henceforth, all unidentified limits are taken as $c \to 0$, and $r(\geq 2)$ is an arbitrary integer given in Lemma 2.

LEMMA 4. *For every* $\zeta \in (0, r/2-1)$ *and* $\varepsilon \in (0,1)$

$$P[N \leq n_2] = O(c^{(r/2-\zeta-1)/2(1+h)}) \tag{5.21}$$

and

$$\sum_{n\geq n_3} P[N > n] = O(c^{(r/2-\zeta-1)/2}). \tag{5.22}$$

PROOF. In the definition of stopping rule N in (2.11), we can see that $n \geq c^{-1/2}n^{-h}$, which implies $n \geq c^{-1/2(1+h)} = n_1$, hence $N \geq n_1$. Observe that

$$\begin{aligned} P[N \leq n_2] &\leq P[(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n)^{1/2} \leq nc^{1/2}, \text{ for some } n_1 \leq n \leq n_2] \\ &\leq P[(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n)^{1/2} \leq (1-\varepsilon)n_0 c^{1/2}, \text{ for some } n \in [n_1, n_2]] \\ &\leq P[\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n \leq \operatorname{tr} \boldsymbol{Q}\boldsymbol{V}(1-\varepsilon)^2, \text{ for some } n \in [n_1, n_2]] \\ &\leq P[|\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n - \operatorname{tr} \boldsymbol{Q}\boldsymbol{V}| \geq \varepsilon(2-\varepsilon)\operatorname{tr} \boldsymbol{Q}\boldsymbol{V}, \text{ for some } n \in [n_1, n_2]] \\ &\leq \sum_{n=n_1}^{\infty} O(h_n^{3r} \cdot n^{-r/2}), \end{aligned}$$

where the last inequality follows from Lemma 3. Using (2.15) we obtain

$$P[N \leq n_2] = O(n_1^{-r/2+\zeta+1}), \qquad \forall \zeta \in (0, r/2-1),$$

which, together with (5.20) implies (5.21). Next we prove (5.22). From (2.11) it follows that, for $n \geq n_3$,

$$\begin{aligned} P[N > n] &\leq P[(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n)^{1/2} > c^{1/2}n - n_3^{-h}] \\ &\leq P[(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n)^{1/2} - (\operatorname{tr} \boldsymbol{Q}\boldsymbol{V})^{1/2} > c^{1/2}(n_3 - n_0) - n_3^{-h}]. \end{aligned}$$

Choose c small enough so that

$$\varepsilon(\operatorname{tr} \boldsymbol{QV})^{1/2} - \{c^{1/2}/[(\operatorname{tr} \boldsymbol{QV})^{1/2}(1+\varepsilon)]\}^h > \varepsilon(\operatorname{tr} \boldsymbol{QV})^{1/2}/2.$$

Then

$$\begin{aligned} P[N > n] &\leq P[(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n)^{1/2} - (\operatorname{tr} \boldsymbol{QV})^{1/2} > \varepsilon(\operatorname{tr} \boldsymbol{QV})^{1/2}/2] \\ &\leq P[|\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n - \operatorname{tr} \boldsymbol{QV}| > \varepsilon^2 \operatorname{tr} \boldsymbol{QV}/4]. \end{aligned} \tag{5.23}$$

In the same way as (5.21) we conclude (5.22). □

Let $A = [n_2 < N < n_3]$, $B = [N \leq n_2]$ and $D = [N \geq n_3]$. We denote by I_F and $\bar{F}$ the indicator and the complement of a set F, respectively.

PROOF OF THEOREM 1. From (5.18) and the Borel-Cantelli lemma we can see

$$\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_n \to \operatorname{tr} \boldsymbol{QV} \quad \text{a.s.} \quad \text{as } n \to \infty. \tag{5.24}$$

Therefore, for fixed c we have $N < \infty$ a.s. (recall (2.11)). While, since $N \geq n_1$, if $c \to 0$, then $N \to \infty$. Hence, from this and (5.24)

$$\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_N \to \operatorname{tr} \boldsymbol{QV} \quad \text{a.s.} \quad \text{as } c \to 0. \tag{5.25}$$

But

$$c^{-1/2}(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_N)^{1/2} \leq N \tag{5.26}$$

and

$$N \leq c^{-1/2}[(\operatorname{tr} \boldsymbol{Q}\hat{\boldsymbol{V}}_{N-1})^{1/2} + (N-1)^{-h}] + n_1. \tag{5.27}$$

Therefore, from (5.24), (5.25), (5.26) and $n_0 \simeq c^{-1/2}(\operatorname{tr} \boldsymbol{QV})^{1/2}$

$$N/n_0 \to 1 \quad \text{a.s.} \quad \text{as } c \to 0.$$

We next prove the L_1-convergence (3.2). Write

$$N/n_0 - 1 = (N/n_0)I_B + (N/n_0 - 1)I_A + (N/n_0)I_D - I_{B\cup D}.$$

Then

$$E|N/n_0 - 1| \leq (2-\varepsilon)P[N \leq n_2] + \varepsilon + n_0^{-1} \sum_{n \geq n_3} P[N > n] + P[N > n_3].$$

Since ε is arbitrary, the required result follows from Lemma 4.

Now, it only remains to deal with the asymptotic risk efficiency. Since

$$\frac{R_N}{R_{n_0}} \simeq \frac{\operatorname{tr} \boldsymbol{Q}E(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})' + cEN}{2cn_0} \quad \text{and} \quad c \simeq \operatorname{tr} \boldsymbol{QV}/n_0^2,$$

it is sufficient to prove that

$$\frac{\operatorname{tr} \boldsymbol{Q}E(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})'}{cn_0} \to 1. \tag{5.28}$$

Recalling (2.6) and $c \simeq \operatorname{tr} \boldsymbol{QV}/n_0^2$, we observe that $\operatorname{tr} \boldsymbol{Q} E(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta})'/cn_0 \to 1$. Next we evaluate $\boldsymbol{M} = \{M_{ij}\} \equiv E\{(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta})'(I_A - 1)\}/cn_0$. By Schwarz inequality, it is seen that M_{ij}'s are of order $O[\{E(I_A - 1)^2\}^{1/2}]$. Since $I_A \to 1$ a.s., by the dominated convergence theorem we can see $\boldsymbol{M} \to \mathbf{0}$, hence

$$\operatorname{tr} \boldsymbol{Q} E(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta})' I_A/cn_0 \to 1.$$

Therefore, to prove (5.28) we show that

$$\operatorname{tr} \boldsymbol{Q} E(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta})' I_{\bar{A}}/cn_0 \to 0 \tag{5.29}$$

and

$$\operatorname{tr} \boldsymbol{Q} E(\hat{\boldsymbol{\theta}}_N - \hat{\boldsymbol{\theta}}_{n_0})(\hat{\boldsymbol{\theta}}_N - \hat{\boldsymbol{\theta}}_{n_0})' I_A/cn_0 \to 0. \tag{5.30}$$

We prove (5.29). Using Schwarz inequality, Lemma 1 and Lemma 4 yields for each i and j

$$\begin{aligned} &|E(\hat{\theta}_{N,i} - \theta_i)(\hat{\theta}_{N,j} - \theta_j) I_B| \\ &\quad \le E \max_{n_1 \le n \le n_2} |(\hat{\theta}_{n,i} - \theta_i)(\hat{\theta}_{n,j} - \theta_j)| I_B \\ &\quad \le \left\{ E \max_{n_1 \le n \le n_2} (\hat{\theta}_{n,i} - \theta_i)^2 (\hat{\theta}_{n,j} - \theta_j)^2 \right\}^{1/2} \{P(B)\}^{1/2} \\ &\quad \le \left\{ \sum_{n=n_1}^{n_2} E(\hat{\theta}_{n,i} - \theta_i)^2 (\hat{\theta}_{n,j} - \theta_j)^2 \right\}^{1/2} \{P(B)\}^{1/2} \\ &\quad \le \left\{ \sum_{n=n_1}^{n_2} E^{1/2}(\hat{\theta}_{n,i} - \theta_i)^4 E^{1/2}(\hat{\theta}_{n,j} - \theta_j)^4 \right\}^{1/2} \{P(B)\}^{1/2} \\ &\quad \le O(n_2^{1/2} \cdot n_1^{-1}) O(c^{(r/2 - \zeta - h)/4(1+h)}). \end{aligned} \tag{5.31}$$

From the definition of n_1 and n_2, (5.31) becomes $O(c^{(r/2-\zeta-h)/4(1+h)})$. Since we can choose r sufficiently large, therefore

$$E(\hat{\theta}_{N,i} - \theta_i)(\hat{\theta}_{N,j} - \theta_j) I_B/cn_0 \to 0, \quad \text{as} \quad c \to 0.$$

Also we have

$$|E(\hat{\theta}_{N,i} - \theta_i)(\hat{\theta}_{N,j} - \theta_j) I_D| \le \sum_{n \ge n_3} E\{|(\hat{\theta}_{N,i} - \theta_i)(\hat{\theta}_{N,j} - \theta_j) I_{D_n}|\} \tag{5.32}$$

where $D_n = \{N = n\}$, $n \ge n_3$. By Schwarz inequality and Lemma 1, the right hand side of (5.32) is dominated by

$$\begin{aligned} &\sum_{n \ge n_3} E^{1/2}((\hat{\theta}_{N,i} - \theta_i) I_{D_n})^2 E^{1/2}((\hat{\theta}_{N,j} - \theta_j) I_{D_n})^2 \\ &\quad \le \sum_{n \ge n_3} E^{1/4}(\hat{\theta}_{n,i} - \theta_i)^4 P(N = n)^{1/4} E^{1/4}(\hat{\theta}_{n,j} - \theta_j)^4 P(N = n)^{1/4} \\ &\quad \le O(n_3^{-1}) \sum_{n \ge n_3} P(N = n)^{1/2}. \end{aligned} \tag{5.33}$$

Recalling (5.22) we can see that $\sum_{n \geq n_3} P(N = n)^{1/2} = o(1)$. Hence,

$$E(\hat{\theta}_{N,i} - \theta_i)(\hat{\theta}_{N,j} - \theta_j)I_D/cn_0 \to 0, \qquad \text{as} \quad c \to 0,$$

which together with (5.32) proves (5.29). As for (5.30), we have for each i and j

$$\begin{aligned}(5.34) \qquad & |E(\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i})(\hat{\theta}_{N,j} - \hat{\theta}_{n_0,j})I_A| \\ &= \left| E\left[\sum_{n=n_2}^{n_3} (\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i})(\hat{\theta}_{N,j} - \hat{\theta}_{n_0,j})I_{A_n} \right] \right| \\ &\leq \sum_{n_2 \leq n \leq n_3} |E\{(\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i})(\hat{\theta}_{N,j} - \hat{\theta}_{n_0,j})I_{A_n}\}|\end{aligned}$$

where $A_n = \{N = n\}$, $n_2 \leq n \leq n_3$. In order to evaluate $E(\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i})(\hat{\theta}_{N,j} - \hat{\theta}_{n_0,j})$, we need the following preparation. Let

$$h_{n_3}^{(n)}(s) = \begin{cases} 1 \text{ if } & 0 \leq s \leq n/n_3 \\ 0 \text{ otherwise.} \end{cases}$$

Define the tapered periodogram by

$$(5.35) \qquad \tilde{I}_{n_3,ab}^{(n)}(\lambda) = \{2\pi H_2^{(n)}(0)\}^{-1} d_{n_3,a}^{(n)}(\lambda) d_{n_3,b}^{(n)}(-\lambda),$$

where

$$d_{n_3,a}^{(n)}(\lambda) = \sum_{t=1}^{n_3} h_{n_3}^{(n)}(t/n_3) X_{t,a}\, e^{-it\lambda}$$

and

$$H_2^{(n_3)}(0) = \sum_{t=1}^{n_3} h_{n_3}^{(n)}(t/n_3)^2.$$

For $n = n_2, \ldots, n_3$, we can write $\hat{\theta}_{n,i}$ as

$$\hat{\theta}_{n,i} = \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_i(\lambda) \tilde{\boldsymbol{I}}_{n_3}^{(n)}(\lambda)\} \mathrm{d}\lambda,$$

where $\tilde{\boldsymbol{I}}_{n_3}^{(n)}(\lambda)$ is the periodogram matrix whose (a,b)-th entry is as in (5.35). A slight modification of *Theorem* 7.7.1 of Brillinger (1981) and (2.1) of Dahlhaus (1988) leads to

$$\begin{aligned} & n_3 \operatorname{cov}(\hat{\theta}_{n,i}, \hat{\theta}_{n_0,i}) \\ &= \frac{\int h_{n_3}^{(n)}(t)^2 h_{n_3}^{(n_0)}(t)^2 \mathrm{d}t}{\int h_{n_3}^{(n)}(t)^2 \mathrm{d}t \int h_{n_3}^{(n_0)}(t)^2 \mathrm{d}t} \left[4\pi \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\psi}_i(\lambda) \boldsymbol{f}(\lambda) \boldsymbol{\psi}_j(\lambda) \boldsymbol{f}(\lambda)\} \mathrm{d}\lambda \right] + o(1).\end{aligned}$$

Using above relation, we can show that

$$\begin{aligned} E(\hat{\theta}_{n,i} - \hat{\theta}_{n_0,i})^2 &= E[\{\hat{\theta}_{n,i} - \theta_i - (\hat{\theta}_{n_0,i} - \theta_i)\}^2] \\ &= O\left(\frac{1}{n} - \frac{1}{n_0}\right) \simeq \varepsilon \times O\left(\frac{1}{n_0}\right).\end{aligned}$$

Noting (5.4) it is seen that

$$E(\hat{\theta}_{n,i} - \hat{\theta}_{n_0,i})^4 = \varepsilon^2 \times O\left(\frac{1}{n_0^2}\right). \tag{5.36}$$

As in (5.34), by Schwarz inequality and (5.36) we evaluate (5.35) as

$$\begin{aligned}
&\sum_{n=n_2}^{n_3} |E(\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i})(\hat{\theta}_{N,j} - \hat{\theta}_{n_0,j}) I_{A_n}| \\
&\quad \le \sum_{n=n_2}^{n_3} E^{1/2}((\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i}) I_{A_n})^2 E^{1/2}((\hat{\theta}_{N,j} - \hat{\theta}_{n_0,j}) I_{A_n})^2 \\
&\quad \le \sum_{n=n_2}^{n_3} E^{1/4}(\hat{\theta}_{n,i} - \hat{\theta}_{n_0,i})^4 P(N=n)^{1/4} E^{1/4}(\hat{\theta}_{n,j} - \hat{\theta}_{n_0,j})^4 P(N=n)^{1/4} \\
&\quad \le \varepsilon \times O\left(\frac{1}{n_0}\right).
\end{aligned}$$

Since ε is arbitrary,

$$\operatorname{tr} \boldsymbol{Q} E(\hat{\boldsymbol{\theta}}_N - \hat{\boldsymbol{\theta}}_{n_0})(\hat{\boldsymbol{\theta}}_N - \hat{\boldsymbol{\theta}}_{n_0})' I_A / c n_0 \to 0, \tag{5.37}$$

which completes the proof. □

Proof of Theorem 2. Write

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}) = \sqrt{N/n_0}\sqrt{n_0}(\hat{\boldsymbol{\theta}}_N - \hat{\boldsymbol{\theta}}_{n_0}) + \sqrt{N/n_0}\sqrt{n_0}(\hat{\boldsymbol{\theta}}_{n_0} - \boldsymbol{\theta}).$$

By (5.29), (5.30) and since $cn_0^2 = \operatorname{tr} \boldsymbol{QV}$

$$n_0 E(\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i})^2 \to 0, \tag{5.38}$$

which implies

$$\sqrt{n_0}|\hat{\theta}_{N,i} - \hat{\theta}_{n_0,i}| \underset{P}{\to} 0. \tag{5.39}$$

Thus the proof follows from (2.11), (3.1) and the Slutsky theorem. □

Proof of Theorem 3. The proofs of (3.5) and (3.7) are very much similar to the proofs of (3.1) and (3.2), respectively. Thus we prove (3.6) only. We have

$$\begin{aligned}
&P(\boldsymbol{\theta} \in R_T) \\
&\quad = P\{(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \le d^2\} \\
&\quad = P\{\boldsymbol{a}'(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'\boldsymbol{a} \le d^2 \text{ for all } \boldsymbol{a} \text{ such that } \boldsymbol{a}'\boldsymbol{a} = 1\} \\
&\quad = P\left\{\frac{T\boldsymbol{a}'(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'\boldsymbol{a}}{\lambda_{\max}(\boldsymbol{V})} \le \frac{d^2 T}{\lambda_{\max}(\boldsymbol{V})} \text{ for all } \boldsymbol{a} : \boldsymbol{a}'\boldsymbol{a} = 1\right\} \\
&\quad \ge P\left\{\frac{T\boldsymbol{a}'(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}} \le \frac{d^2 T}{\lambda_{\max}(\boldsymbol{V})} \text{ for all } \boldsymbol{a} : \boldsymbol{a}'\boldsymbol{a} = 1\right\} \\
&\quad = P\left\{\sup_{\boldsymbol{a}'\boldsymbol{a}=1} \frac{T\boldsymbol{a}'(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{V}\boldsymbol{a}} \le \frac{d^2 T}{\lambda_{\max}(\boldsymbol{V})}\right\} \\
&\quad = P\{T(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta})'\boldsymbol{V}^{-1}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \le (d^2 T/\lambda_{\max}(\boldsymbol{V}))\}.
\end{aligned}$$

By (3.5) and $d^2T/\lambda_{\max}(\boldsymbol{V}) \to \chi^2_{1-\alpha}(q)$, it follows from Anscombe's Theorem that as $d \to 0 \sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) \to_{\mathcal{L}} \mathcal{N}(\boldsymbol{0}, \boldsymbol{V})$. Hence $\lim_{d \to 0} P(\boldsymbol{\theta} \in R_T) \geq 1 - \alpha$. Thus the proof is completed. □

Acknowledgements

The authors thank two anonymous referees for careful reading and instructive comments. Thanks are also due to Professor S. Shirahata for his support and encouragement.

References

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, New York.

Brillinger, D. R. (1969). Asymptotic properties of spectral estimates of second order. *Biometrika* **60** 217–226.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory: Expanded ed.*, Holt, Rinehart and Winston, New York.

Chow, Y. S. and Robbins, H. E. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean, *Ann. Math. Statist.*, **36**, 457–462.

Dahlhaus, R. (1988). Empirical spectral processes and their applications to time series analysis, *Stochastic Process. Appl.*, **30**, 69–83.

Fakhre-Zakeri, I. and Lee, S. (1992). Sequential estimation of the mean of a linear process, *Sequential Anal.*, **11**, 181–197.

Fakhre-Zakeri, I. and Lee, S. (1993). Sequential estimation of the mean vector of a multivariate linear process, *J. Multivariate Anal.*, **47**, 196–209.

Hosoya, Y. and Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear process, *Ann. Statist.*, **10**, 132–153 (Correction: *ibid* **21**, 1115–1117).

Lee, S. (1994). Sequential estimation for the parameter of a stationary autoregressive model, *Sequential Anal.*, **13**, 301–317.

Robbins, H. (1959). Sequential estimation of the mean of a normal population, *Probability and Statistics* (ed. U. Grenander) 235–245, Wiley, New York.

Sriram, T. N. (1987). Sequential estimation of the mean of a first order stationary process, *Ann. Statist.*, **15**, 1079–1090.

Sriram, T. N. (1988). Sequential estimation of the autoregressive parameter in a first order autoregressive process, *Sequential Anal.*, **7**, 53–74.