LARGE DEVIATION PRINCIPLE FOR THE SAMPLE COVARIANCE FUNCTION OF A FIRST ORDER AUTOREGRESSIVE PROCESS

Haruyoshi Mita*

This paper describes a large deviation principle for the sample covariance function $\sum_{i=1}^n X_i X_{i-1} / n$ of a first order non-explosive Gaussian autoregressive process with unknown autoregressive parameter θ , the rate function is also provided. The asymptotic rate of convergence of the tail probability of the sample covariance function is also studied.

1. Introduction

We consider the first order non-explosive Gaussian autoregressive process

$$(1.1) \quad X_n = \theta X_{n-1} + U_n, \quad X_0 = 0, \quad |\theta| < 1, \quad n = 1, 2, \dots,$$

where $U_1, U_2, \dots$ are independent and normally distributed with mean 0 and unit variance. Let P_θ be the probability law governing the process (1.1), and let $S_n = \sum_{i=1}^n X_i X_{i-1} / n$. We will show in this paper the large deviation principle for the sample covariance function S_n , that is, we give the large deviation rate function I having the property that

$$\text{for any closed set } F, \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_\theta(S_n \in F) \leq - \inf_{x \in F} I(x),$$

and

$$\text{for any open set } G, \liminf_{n \rightarrow \infty} \frac{1}{n} \log P_\theta(S_n \in G) \geq - \inf_{x \in G} I(x),$$

where I is a lower semi-continuous function. A rigorous definition of the large deviation principle will be given in section 2.

Furthermore, in estimation problems for the autoregressive process (1.1) it is known that $S_n \rightarrow \theta/(1-\theta^2)$ in probability as $n \rightarrow \infty$ under P_θ . Therefore, we are interested in studying the asymptotic rate of convergence of the probabilities $P_\theta(|S_n - \theta/(1-\theta^2)| > \varepsilon)$ as $n \rightarrow \infty$, where $\varepsilon > 0$ is arbitrary.

The process $\{X_n\}$ given by (1.1) is a Markov process. For non i.i.d. processes, the Gärtner-Ellis theorem, an extension of Cramer's theorem which is a large deviation theorem for i.i.d. process (c.f., e.g., Bucklew (1990), or Dembo and Zeitouni (1993)), is useful for our problem.

For the autoregressive process (1.1), Bryc and Smolenski (1993) studied the large deviation principle for the sum of squares of the process $\{X_n\}$, that is,

Received January, 1996. Revised March, 1996. Accepted March, 1996.

* Faculty of Liberal Arts, University of The Sacred Heart, Shibuya-ku, Tokyo 150, Japan

$\sum_{i=1}^n X_{i-1}^2 / n$. In statistical inference for the model (1.1), $\sum_{i=1}^n X_{i-1}^2$ is denoted as the observed Fisher information. They also provided the explicit rate function. Although we show the large deviation principle for the sample covariance function $S_n = \sum_{i=1}^n X_i X_{i-1} / n$ in the same manner, the problem we studied is more complicated than theirs.

In section 2, we show the large deviation principle for the sample covariance function and provide the rate function. In section 3, we study the asymptotic rate of convergence of the probabilities $P_\theta(|S_n - \theta/(1 - \theta^2)| > \varepsilon)$.

2. Large deviation principle

We first give a definition of the large deviation principle (c.f., Dembo and Zeitouni (1993), and Deuschel and Stroock (1989)).

Let $\{P_n: n \geq 1\}$ be a sequence of probability measures on the Borel σ -algebra $\mathbf{B}$ of a complete separable metric space X .

DEFINITION 1. We say that a sequence $\{P_n: n \geq 1\}$ satisfies the large deviation principle with rate function I if there exists a function $I: X \rightarrow [0, \infty]$ satisfying

- (1) $0 \leq I(x) \leq \infty$ for all $x \in X$,
- (2) I is a lower semi-continuous function on X ,
- (3) for any $l < \infty$, the level set $\{x: I(x) \leq l\}$ is a compact subset of X ,
- (4) for any closed set $F \subset X$, $\limsup_{n \rightarrow \infty} \frac{1}{n} \log P_n(F) \leq -\inf_{x \in F} I(x)$,
- (5) for any open set $G \subset X$, $\liminf_{n \rightarrow \infty} \frac{1}{n} \log P_n(G) \geq -\inf_{x \in G} I(x)$.

Let $\{X_n\}$ be the autoregressive process given by (1.1), and let S_n be the sample covariance function, i.e., $S_n = \sum_{i=1}^n X_i X_{i-1} / n$. We denote P_θ as the probability law governing the process given by (1.1), and we denote $P_{n,\theta}$ as the distribution of S_n , that is, for any $A \in \mathbf{B}$, $P_{n,\theta}(A) = P_\theta(S_n \in A)$, and E_θ denotes the expectation under P_θ . Further, we denote $\varphi_{n,\theta}(s)$ as the moment generating function of the sample covariance function S_n under P_θ . For convenience, as long as there is no confusion, we often abbreviate the parameter θ for some notations throughout this paper.

LEMMA 2.1. Suppose that $\{X_n\}$ is given by (1.1). Then it follows that for any $n=2, 3, \dots$,

$$\varphi_{n,\theta}(ns) = \begin{cases} \prod_{i=1}^{n-1} (1 - 2a_i s)^{-\frac{1}{2}} & \text{for any } s \in \bigcap_{i=1}^{n-1} J_i \\ +\infty & \text{otherwise,} \end{cases}$$

where $a_1, a_2, \dots, a_n$ are given by the following recurrence:

$$(2.1) \quad a_1 = \theta + \frac{s}{2},$$

$$(2.2) \quad a_n = \frac{\theta^2 a_{n-1} + a_1}{1 - 2a_{n-1}s} \quad (n=2, 3, \dots),$$

and $J_n(n=1, 2, \dots)$ is given by $J_n = \{s: 1 - 2a_ns > 0\}$.

PROOF. By (1.1) and the assumption that $U_1, U_2, \dots$ are normally distributed, it follows that

$$\begin{aligned} (2.3) \quad \varphi_{n,\theta}(ns) &= E_\theta \left(\exp \left(s \sum_{i=1}^n X_i X_{i-1} \right) \right) \\ &= E_\theta \left(\exp \left(s \left(\sum_{i=1}^{n-1} X_i X_{i-1} + \theta X_{n-1}^2 + U_n X_{n-1} \right) \right) \right) \\ &= E_\theta \left(E_\theta \left(\exp \left(s \left(\sum_{i=1}^{n-1} X_i X_{i-1} + \theta X_{n-1}^2 + U_n X_{n-1} \right) \right) \middle| U_1, U_2, \dots, U_{n-1} \right) \right) \\ &= E_\theta \left(\exp \left(s \left(\sum_{i=1}^{n-1} X_i X_{i-1} + \theta X_{n-1}^2 \right) \right) \right. \\ &\quad \cdot E_\theta \left(\exp(s U_n X_{n-1}) \middle| U_1, U_2, \dots, U_{n-1} \right) \Big) \\ &= E_\theta \left(\exp \left(s \left(\sum_{i=1}^{n-1} X_i X_{i-1} + \theta X_{n-1}^2 \right) + \frac{1}{2} s^2 X_{n-1}^2 \right) \right) \\ &= E_\theta \left(\exp \left(s \sum_{i=1}^{n-1} X_i X_{i-1} + \left(\theta + \frac{1}{2} s \right) s X_{n-1}^2 \right) \right) \\ &= E_\theta \left(\exp \left(s \sum_{i=1}^{n-1} X_i X_{i-1} + a_1 s X_{n-1}^2 \right) \right). \end{aligned}$$

By conditioning on $U_1, U_2, \dots, U_{n-2}$, for $s \in J_1$ we have

$$\begin{aligned} (2.4) \quad \text{R.H.S. of (2.3)} &= E_\theta \left(\exp \left(s \sum_{i=1}^{n-2} X_i X_{i-1} + s(\theta X_{n-2} + U_{n-1}) X_{n-2} + a_1 s(\theta X_{n-2} + U_{n-1})^2 \right) \right) \\ &= E_\theta \left(\exp \left(s \sum_{i=1}^{n-2} X_i X_{i-1} + \theta(s + a_1 s \theta) X_{n-2}^2 + a_1 s U_{n-1}^2 \right. \right. \\ &\quad \left. \left. + (s + 2a_1 s \theta) X_{n-2} U_{n-1} \right) \right) \\ &= E_\theta \left(\exp \left(s \left(\sum_{i=1}^{n-2} X_i X_{i-1} + \theta(s + a_1 s \theta) X_{n-2}^2 \right) \right) \right. \\ &\quad \cdot E_\theta \left(\exp(a_1 s U_{n-1}^2 + (s + 2a_1 s \theta) U_{n-1} X_{n-2}) \middle| U_1, U_2, \dots, U_{n-2} \right) \Big) \\ &= (1 - 2a_1 s)^{-\frac{1}{2}} E_\theta \left(\exp \left(s \sum_{i=1}^{n-2} X_i X_{i-1} + \theta(s + a_1 s \theta) X_{n-2}^2 \right. \right. \\ &\quad \left. \left. + \frac{(s + 2a_1 s \theta)^2}{2(1 - 2a_1 s)} X_{n-2}^2 \right) \right) \\ &= (1 - 2a_1 s)^{-\frac{1}{2}} E_\theta \left(\exp \left(s \sum_{i=1}^{n-2} X_i X_{i-1} + \frac{\theta^2 a_1 s + a_1 s}{1 - 2a_1 s} X_{n-2}^2 \right) \right) \\ &= (1 - 2a_1 s)^{-\frac{1}{2}} E_\theta \left(\exp \left(s \sum_{i=1}^{n-2} X_i X_{i-1} + a_2 s X_{n-2}^2 \right) \right). \end{aligned}$$

In view of (2.3) and (2.4), and by iterating the same calculations as above, we obtain for any $s \in \bigcap_{i=1}^{n-1} J_i$,

$$\varphi_{n,\theta}(ns) = \prod_{i=1}^{n-1} (1 - 2a_i s)^{-\frac{1}{2}}.$$

It is easy to see that $\varphi_{n,\theta}(ns) = +\infty$ for any $s \notin \bigcap_{i=1}^{n-1} J_i$. This completes the proof. $\square$

LEMMA 2.2. Suppose that $\{X_n\}$ is given by (1.1). It follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \varphi_{n,\theta}(ns) = L(s),$$

where L is given by

$$L(s) = \begin{cases} -\frac{1}{2} \log \left[\frac{1}{2} (\theta^2 + 1 + \sqrt{((\theta+1)^2 + 2s)((\theta-1)^2 - 2s)}) \right] & \text{for } -\frac{1}{2}(\theta+1)^2 \leq s \leq \frac{1}{2}(\theta-1)^2, \\ +\infty & \text{otherwise.} \end{cases}$$

Moreover, it follows that

$$\lim_{s \rightarrow -\frac{(\theta+1)^2}{2} + 0} |L'(s)| = \lim_{s \rightarrow \frac{(\theta-1)^2}{2} - 0} |L'(s)| = +\infty,$$

where L' is the first derivative of L .

PROOF. We consider the following correspondence of a 2×2 matrix and a linear fractional transformation:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longleftrightarrow \frac{ax+b}{cx+d}.$$

Let $f(x) = \frac{\theta^2 x + a_1 s}{-2x + 1}$, and the corresponding matrix is denoted by A , i.e.,

$$A = \begin{pmatrix} \theta^2 & a_1 s \\ -2 & 1 \end{pmatrix}.$$

By (2.2), $a_n s$ is given by the $(n-1)$ -fold composition $f^{n-1}(a_1 s)$ and the corresponding matrix is given by A^{n-1} . Here we define $\Lambda = \{s : -(\theta+1)^2/2 \leq s \leq (\theta-1)^2/2\}$. By easy calculation, for any $s \in \Lambda^\circ$, where Λ° denotes the interior of Λ , the matrix A has the following two distinct positive eigenvalues α and β ($\beta < \alpha$):

$$\alpha = \frac{1}{2} (\theta^2 + 1 + \sqrt{(\theta^2 - 1)^2 - 8a_1 s}),$$

and

$$\beta = \frac{1}{2} (\theta^2 + 1 - \sqrt{(\theta^2 - 1)^2 - 8a_1 s}).$$

Furthermore, from the remainder theorem it is easy to see that for $n \geq 3$,

$$x^{n-1} = \Phi_A(x)Q_{n-3}(x) + \frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta}x + \frac{\alpha\beta^{n-1} - \alpha^{n-1}\beta}{\alpha - \beta},$$

where $\Phi_A(x)$ is the characteristic polynomial of the matrix A and $Q_{n-3}(x)$ is some polynomial of degree $n-3$. By the Hamilton-Cayley's theorem, we have

$$A^{n-1} = \frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta} A + \frac{\alpha\beta^{n-1} - \alpha^{n-1}\beta}{\alpha - \beta} E,$$

where E is the identity matrix. Therefore, we have

$$A^{n-1} = \begin{pmatrix} \frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta} \theta^2 + \frac{\alpha\beta^{n-1} - \alpha^{n-1}\beta}{\alpha - \beta} & \frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta} a_1 s \\ -\frac{2(\alpha^{n-1} - \beta^{n-1})}{\alpha - \beta} & \frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta} + \frac{\alpha\beta^{n-1} - \alpha^{n-1}\beta}{\alpha - \beta} \end{pmatrix}.$$

Since $a_n s = f^{n-1}(a_1 s)$ corresponds to the matrix A^{n-1} and the composition of linear fractional transformations is also a linear fractional transformation, it follows that for $n \geq 3$,

$$\begin{aligned} (2.5) \quad a_n s &= \frac{((\alpha^{n-1} - \beta^{n-1})\theta^2 + (\alpha\beta^{n-1} - \alpha^{n-1}\beta))a_1 s + (\alpha^{n-1} - \beta^{n-1})a_1 s}{-2(\alpha^{n-1} - \beta^{n-1})a_1 s + \alpha^{n-1} - \beta^{n-1} + \alpha\beta^{n-1} - \alpha^{n-1}\beta} \\ &= \frac{\left[(\theta^2 + 1 - \beta) - \left(\frac{\beta}{\alpha}\right)^{n-1}(\theta^2 + 1 - \alpha) \right] a_1 s}{1 - \beta - 2a_1 s - \left(\frac{\beta}{\alpha}\right)^{n-1}(1 - \alpha - 2a_1 s)} \\ &= \frac{[\theta^2 + 1 - \beta - \Delta^{n-1}(\theta^2 + 1 - \alpha)] a_1 s}{1 - \beta - 2a_1 s - \Delta^{n-1}(1 - \alpha - 2a_1 s)}, \end{aligned}$$

where $\Delta = \frac{\beta}{\alpha}$.

For any $s \in \Lambda^*$, from (2.5) we have

$$(2.6) \quad a_n s \rightarrow \frac{(\theta^2 + 1 - \beta)a_1 s}{-2a_1 s + 1 - \beta} = a(s) \quad (\text{say}) \quad \text{as } n \rightarrow \infty.$$

Next, we will deal with the boundary points of Λ . Let $s = (\theta - 1)^2/2$. For this case we have $\alpha = \beta = (\theta^2 + 1)/2$, and $x^{n-1} = \Phi_A(x)Q_{n-3}(x) + ax + b$ for $n \geq 3$. From the remainder theorem, we obtain $\beta^{n-1} = a\beta + b$. Moreover, by differentiating the function x^{n-1} and by the fact that the characteristic equation $\Phi_A(x) = 0$ has double root, we have $(n-1)\beta^{n-2} = a$. Therefore, we obtain $a = (n-1)\beta^{n-2}$ and $b = (2-n)\beta^{n-1}$. Again, by the Hamilton-Cayley's theorem, we have

$$\begin{aligned} A^{n-1} &= aA + bE \\ &= \begin{pmatrix} (n-1)\theta^2\beta^{n-2} - (n-2)\beta^{n-1} & (n-1)a_1 s\beta^{n-2} \\ -2(n-1)\beta^{n-2} & (n-1)\beta^{n-2} - (n-2)\beta^{n-1} \end{pmatrix}. \end{aligned}$$

Hence, we obtain

$$(2.7) \quad a_{nS} = \frac{((n-1)\theta^2 - (n-2)\beta + (n-1))a_1s}{-2(n-1)a_1s + (n-1) - (n-2)\beta}.$$

We have

$$(2.8) \quad a_{nS} \rightarrow \frac{(\theta^2 - \beta + 1)a_1s}{1 - 2a_1s - \beta} = a(s) \quad \text{as } n \rightarrow \infty.$$

Similarly, we can see that (2.8) holds for $s = -(\theta+1)^2/2$.

By combining (2.6), it has been shown that for any $s \in \Lambda$,

$$(2.9) \quad \lim_{n \rightarrow \infty} a_{nS} = a(s) \\ = -\frac{1}{4}(\theta^2 - 1 + \sqrt{((\theta+1)^2 + 2s)((\theta-1)^2 - 2s)}).$$

Next, we show the monotonicity of the sequence $\{a_{nS}\}$.

Case (1): $s \in \Lambda^+$. By (2.5), we have for $n \geq 4$,

$$(2.10) \quad a_{nS} - a_{n-1S} \\ = \frac{[\theta^2 + 1 - \beta - \Delta^{n-1}(\theta^2 + 1 - \alpha)]a_1s}{1 - \beta - 2a_1s - \Delta^{n-1}(1 - \alpha - 2a_1s)} \\ - \frac{[\theta^2 + 1 - \beta - \Delta^{n-2}(\theta^2 + 1 - \alpha)]a_1s}{1 - \beta - 2a_1s - \Delta^{n-2}(1 - \alpha - 2a_1s)} \\ = \frac{(\Delta^{n-2} - \Delta^{n-1})(\alpha - \beta)(\theta^2 + 2a_1s)a_1s}{[1 - \beta - 2a_1s - \Delta^{n-1}(1 - \alpha - 2a_1s)][1 - \beta - 2a_1s - \Delta^{n-2}(1 - \alpha - 2a_1s)]}.$$

After some calculations, we have $1 - \beta - 2a_1s > 0$ for any $s \in \Lambda^+$. It follows that if $a_1s > 0$, $-\theta^2/2 < a_1s < 0$, and $a_1s = 0$ or $-\theta^2/2$ then $1 - \alpha - 2a_1s > 0$, $1 - \alpha - 2a_1s < 0$, and $1 - \alpha - 2a_1s = 0$, respectively. Since $0 < \Delta < 1$ and $1 - \alpha - 2a_1s < 1 - \beta - 2a_1s$, from (2.10) we have for $n \geq 4$,

$$(2.11) \quad a_{nS} - a_{n-1S} \geq \frac{(\Delta^{n-2} - \Delta^{n-1})(\alpha - \beta)(\theta^2 + 2a_1s)a_1s}{[1 - \beta - 2a_1s - \Delta^{n-1}(1 - \alpha - 2a_1s)]^2} > 0 \\ \text{for any } s \in \{s: a_1s > 0\}.$$

Similarly, for $n \geq 4$, we can show that

$$(2.12) \quad a_{nS} < a_{n-1S} \quad \text{for any } s \in \left\{s: -\frac{\theta^2}{2} < a_1s < 0\right\},$$

$$(2.13) \quad a_{nS} = a_3s \quad \text{for any } s \in \left\{s: a_1s = 0 \quad \text{or} \quad -\frac{\theta^2}{2}\right\}.$$

From (2.1) and (2.2), we have

$$(2.14) \quad \begin{aligned} a_3s - a_2s &= \frac{\theta^2 a_2s + a_1s}{1 - 2a_2s} - \frac{(\theta^2 + 1)a_1s}{1 - 2a_1s} \\ &= \frac{(\theta^2 + 2a_1s)(a_2s - a_1s)}{(1 - 2a_2s)(1 - 2a_1s)}, \end{aligned}$$

and

$$(2.15) \quad \begin{aligned} a_2s - a_1s &= \frac{(\theta^2 + 1)a_1s}{1 - 2a_1s} - a_1s \\ &= \frac{(\theta^2 + 2a_1s)a_1s}{1 - 2a_1s}. \end{aligned}$$

For any $s \in \Lambda^*$, we have $-\theta^2/2 \leq a_1s < (\theta^2 - 1)^2/8 < 1/2$. Since $a_2s = (\theta^2 + 1)a_1s/(1 - 2a_1s)$, taking the hyperbola $y = (\theta^2 + 1)x/(1 - 2x)$ in the region $-\theta^2/2 \leq x < (\theta^2 - 1)^2/8$ into account, we can show that $a_2s < 1/2$. Therefore, from (2.14) and (2.15) it follows that if $a_1s > 0$, $-\theta^2/2 < a_1s < 0$, and $a_1s = 0$ or $-\theta^2/2$ then $a_1s < a_2s < a_3s$, $a_3s < a_2s < a_1s$, and $a_1s = a_2s = a_3s$, respectively.

Hence, by combining (2.11), (2.12), and (2.13), if $a_1s > 0$, $-\theta^2/2 < a_1s < 0$, and $a_1s = 0$ or $-\theta^2/2$ then $\{a_ns\}$ is monotone increasing, monotone decreasing, and $a_ns = a_1s$ for any $n \geq 1$, respectively.

Case (2) : $s \in \partial\Lambda$ (∂ denotes the boundary operator). By (2.7), for $n \geq 4$ we have

$$(2.16) \quad \begin{aligned} a_ns - a_{n-1}s &= \frac{((n-1)\theta^2 - (n-2)\beta + n-1)a_1s}{-2(n-1)a_1s + n-1 - (n-2)\beta} - \frac{((n-2)\theta^2 - (n-3)\beta + n-2)a_1s}{-2(n-2)a_1s + n-2 - (n-3)\beta} \\ &= \frac{(\theta^2 + 2a_1s)a_1s\beta}{[-2(n-1)a_1s + n-1 - (n-2)\beta][-2(n-2)a_1s + n-2 - (n-3)\beta]}. \end{aligned}$$

Straightforward computation gives us $a_1s = (\theta^2 - 1)^2/8$ for any $s \in \partial\Lambda$, and therefore we obtain $1 - 2a_1s > 0$ and $1 - 2a_1s - \beta > 0$. Then we have for any $n \geq 4$

$$\begin{aligned} -2(n-2)a_1s + n-2 - (n-3)\beta &= (n-2)(1 - 2a_1s) - (n-3)\beta \\ &> (n-3)(1 - 2a_1s - \beta) \\ &> 0. \end{aligned}$$

Hence, the denominator of R.H.S. in (2.16) is positive. Since $a_1s > 0$, by (2.16), it follows that $\{a_ns\}_{n=3}^\infty$ is a monotone increasing sequence. Furthermore, by (2.14) and (2.15) it is seen that $a_1s < a_2s < a_3s$ in the same manner as case (1). Therefore, $\{a_ns\}$ is monotone increasing.

Consequently, from cases (1) and (2) above, for any $s \in \Lambda$ it follows that if $a_1s > 0$, $-\theta^2/2 < a_1s < 0$, and $a_1s = 0$ or $-\theta^2/2$ then $\{a_ns\}$ is monotone increasing, monotone decreasing, and $a_ns = a_1s$ for any $n \geq 1$, respectively.

After some calculations, we can show that $a(s) < 1/2$ for any $s \in \Lambda$. Therefore, by the above arguments, it follows that $\{s : a_1s > 0\} \cap \Lambda \subset \bigcap_{i=1}^\infty J_i$ and $\{s : -\theta^2/2 < a_1s < 0\} \cap \Lambda \subset \bigcap_{i=1}^\infty J_i$ and $\{s : a_1s = 0 \text{ or } a_1s = -\theta^2/2\} \cap \Lambda \subset \bigcap_{i=1}^\infty J_i$. Consequently, it has been shown that $\Lambda \subset \bigcap_{i=1}^\infty J_i$. Hence, from lemma 2.1 it

follows that for any $s \in \Lambda$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \varphi_{n,\theta}(ns) &= -\frac{1}{2} \log(1-2a(s)) \\ &= -\frac{1}{2} \log \left[\frac{1}{2} (\theta^2 + 1 + \sqrt{\{(\theta+1)^2 + 2s\}\{(\theta-1)^2 - 2s\}}) \right]. \end{aligned}$$

For $s \notin \Lambda$, we shall show that $\lim_{n \rightarrow \infty} \frac{1}{n} \log \varphi_{n,\theta}(ns) = +\infty$. Recall that $f(x) = \frac{\theta^2 x + a_1 s}{1-2x}$. For any $s > (\theta-1)^2/2$ we have $a_1 s > 0$. By (2.4), if $a_1 s \geq 1/2$ then $\varphi_{n,\theta}(ns) = \infty$ for $n=2$. Therefore, we suppose that $a_1 s < 1/2$. We put $g(x) = f(x) - x$. After some calculations, we obtain

$$(2.17) \quad g'(x) = \frac{-4x^2 + 4x + (s + \theta + 1)(s + \theta - 1)}{(1-2x)^2}.$$

According to (2.17), it is easy to see that $g(x)$ attains its minimum value $s - (\theta-1)^2/2$ at $x = (1 - (\theta+s))/2$ in the region $x < 1/2$. We have $f(x) \geq x + s - (\theta-1)^2/2$ for any $s > (\theta-1)^2/2$, and it holds that $a_n s = f(a_{n-1}s) \geq a_{n-1}s + s - (\theta-1)^2/2 \geq a_1 s + (n-1)(s - (\theta-1)^2/2)$. Therefore, there exists some integer n such that $a_n s > 1/2$. Hence $\varphi_{n,\theta}(ns) = \infty$ for some n . In the same manner, we can show that for any $s < -(\theta+1)^2/2$, there exists some integer n such that $\varphi_{n,\theta}(ns) = \infty$.

We put $L(s) = -(1/2) \log(1-2a(s))$ for $s \in \Lambda$, $= +\infty$ otherwise. As mentioned above, it follows that $\lim_{n \rightarrow \infty} \frac{1}{n} \log \varphi_{n,\theta}(ns) = L(s)$.

For any $s \in \Lambda^\circ$, straightforward computation gives us

$$L'(s) = \frac{2(\theta+s)}{\{(\theta+1)^2 + 2s\}\{(\theta-1)^2 - 2s\} + (\theta^2 + 1)\sqrt{\{(\theta+1)^2 + 2s\}\{(\theta-1)^2 - 2s\}}}.$$

Hence it holds that $\lim_{s \rightarrow -(\theta+1)^2/2+0} |L'(s)| = \lim_{s \rightarrow -(\theta-1)^2/2-0} |L'(s)| = +\infty$. This completes the proof. $\square$

THEOREM 2.1. Suppose that $\{X_n\}$ is given by (1.1). Then $\{P_{n,\theta}: n \geq 1\}$ satisfies the large deviation principle with the rate function given by

$$I(x) = xs(x) - L(s(x)),$$

where $s(x)$ is a unique solution of $L'(s) = x$.

PROOF. By lemma 2.2, it is easy to see that L is both convex and a lower semi-continuous function on R^1 , and further, an essentially smooth function. Therefore, by the Gärtner-Ellis theorem (see, Ellis (1984), or Dembo and Zeitouni (1993)), the large deviation principle holds with the rate function $I(x) = \sup_s \{xs - L(s)\}$.

We identify the rate function $I(x)$ in the following. Put $\tilde{L}(s) = sx - L(s)$. We have for $s \in \Lambda$,

$$\begin{aligned}
(2.18) \quad \tilde{L}'(s) &= x - L'(s) \\
&= x - \frac{2(\theta + s)}{(\theta^2 + 1 + \sqrt{((\theta + 1)^2 + 2s)((\theta - 1)^2 - 2s)})} \\
&\quad \cdot \frac{1}{\sqrt{((\theta + 1)^2 + 2s)((\theta - 1)^2 - 2s)}} \\
&= x - \frac{2y}{(c + \sqrt{c^2 - 4y^2})\sqrt{c^2 - 4y^2}},
\end{aligned}$$

where $c = \theta^2 + 1$ and $y = \theta + s$. Here we write $\tilde{L}'(s)$ as $\tilde{L}'(y)$. For any $x \in R^1$ fixed, we solve the equation $\tilde{L}'(y) = 0$. To this end, from (2.18), we shall solve the equation

$$4y(4x^2y^3 + 4xy^2 + (1 - c^2x^2)y - c^2x) = 0.$$

Fix any $x \in R^1$ and define $h(y) = 4x^2y^3 + 4xy^2 + (1 - c^2x^2)y - c^2x$. By straightforward calculation, if $x \neq 0$ then the equation $h'(y) = 0$ has two distinct solutions $y_{\pm}(x) = (-2 \pm \sqrt{1 + 3c^2x^2})/6x$.

Case (i): $x > 0$. For this case, by (2.18), we shall solve the equation $h(y) = 0$ in the region $0 < y < c/2$. Since $h(c/2) = c/2 > 0$, $h(0) = -c^2x < 0$, and $y_-(x) < 0$, the equation $h(y) = 0$ has a unique solution $y(x)$ in the region $0 < y < c/2$. Furthermore, it is easily seen that $h(y) < 0$ for $0 < y < y(x)$ and $h(y) > 0$ for $y(x) < y < c/2$. Hence, taking (2.18) into account, it follows that $\tilde{L}(s)$ attains its maximum value at a unique point $s = s(x)$, where $s(x) = y(x) - \theta$. Of course, $L'(s(x)) = x$.

Case (ii): $x < 0$. Taking (2.18) into account, we solve the equation $h(y) = 0$ in the region $-c/2 < y < 0$. We have $h(-c/2) = -c/2 < 0$ and $h(0) = -c^2x > 0$. By similar calculations as in the case (i), it follows that the equation $h(y) = 0$ has a unique solution $y(x)$ in the region $-c/2 < y < 0$. It follows that $\tilde{L}(s)$ attains its maximum value at a unique point $s = s(x)$, where $s(x) = y(x) - \theta$, and $L'(s(x)) = x$.

Case (iii): $x = 0$. In this case, from (2.18) it follows that $y(0) = 0$, that is, $s(0) = -\theta$. It is easy to see that $L'(-\theta) = 0$ and $I(x) = \sup_s \{-L(s)\} = \log(\theta^2 + 1)/2$. This completes the proof. $\square$

3. Asymptotic rate of convergence of the tail probability of S_n

In statistical analysis of the first order autoregressive model given by (1.1), it is known that $S_n = \sum_{i=1}^n X_i X_{i-1} / n$ converges to $\theta/(1 - \theta^2)$ in probability as $n \rightarrow \infty$ under P_θ . For references concerning the statistical analysis of this model, one may refer to Anderson (1971), Akahira (1976), and Akahira and Takeuchi (1981). We will study the asymptotic rate of convergence of the tail probability of the sample covariance function $S_n = \sum_{i=1}^n X_i X_{i-1} / n$.

THEOREM 3.1. *Suppose that $\{X_n\}$ is given by (1.1). Then it follows that for any $\varepsilon > 0$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta \left(\left| S_n - \frac{\theta}{1 - \theta^2} \right| > \varepsilon \right) = \begin{cases} -I(x(\theta)), & \text{if } 0 < \varepsilon \leq \left| x(\theta) - \frac{\theta}{1 - \theta^2} \right| \\ -\min \left\{ I\left(\frac{\theta}{1 - \theta^2} - \varepsilon \right), I\left(\frac{\theta}{1 - \theta^2} + \varepsilon \right) \right\}, & \text{if } \varepsilon > \left| x(\theta) - \frac{\theta}{1 - \theta^2} \right|, \end{cases}$$

where $I(\cdot)$ is the rate function given by Theorem 2.1, and $x(\theta)$ is a solution of $I'(x) = 0$.

PROOF. In this paragraph, we let $H(x, y) = 4x^2y^3 + 4xy^2 + (1 - c^2x^2)y - c^2x$. In the proof of Theorem 2.1, it was shown that there exists a function $y = y(x)$ such that $H(x, y(x)) = 0$ for any $x \in R^1$, and furthermore, we showed that if $x > 0$ and $x < 0$ then $y_+(x) < y(x)$ and $y(x) < y_-(x)$, respectively. Hence, if $x \neq 0$ then $\partial H(x, y(x))/\partial y \neq 0$. Furthermore, it is easy to see that $\partial H(0, y(0))/\partial y = 1$. Therefore, by the fundamental theorem of implicit functions, it follows that $y = y(x)$ is differentiable and $y'(x) = -(\partial H(x, y(x))/\partial x)/(\partial H(x, y(x))/\partial y)$. For $x > 0$, it is easy to see that $\partial H(x, y)/\partial y > 0$ for $y < y_-(x)$, $\partial H(x, y)/\partial y < 0$ for $y_-(x) < y < y_+(x)$, and $\partial H(x, y)/\partial y > 0$ for $y > y_+(x)$. Hence, $\partial H(x, y(x))/\partial y > 0$ for $x > 0$. Similarly, for $x < 0$ we can show that $\partial H(x, y(x))/\partial y > 0$. Combining the fact that $\partial H(0, y(0))/\partial y = 1$, we have $\partial H(x, y(x))/\partial y > 0$ for any $x \in R^1$. We obtain

$$(3.1) \quad y'(x) = - \left(\frac{\partial H(x, y(x))}{\partial x} \right) / \left(\frac{\partial H(x, y(x))}{\partial y} \right) = - \frac{(4y(x)^2 - c^2)(2xy(x) + 1)}{\frac{\partial H(x, y(x))}{\partial y}}.$$

Since $-c/2 < y(x) < c/2$ and $xy(x) > 0$, from (3.1) we have $y'(x) > 0$. Hence $y(x)$ is a strictly increasing continuous function. Since $s(x) = y(x) - \theta$ and $I(x) = xs(x) - L(s(x))$, $I(x)$ is differentiable. It follows that

$$(3.2) \quad I'(x) = s(x) + s'(x)(x - L'(s(x))) = y(x) - \theta.$$

Suppose that $x > 0$. By the discussions in the proof of Theorem 2.1, the cubic function $h(y)$ is strictly convex on $(y_+(x), c/2)$. Since $h(y_+(x)) < 0$, there exists an intersection point on the y -axis and a straight line which passes through two points $(y_+(x), h(y_+(x)))$ and $(c/2, c/2)$. We denote the intersection point as $(\tilde{y}(x), 0)$. We obtain $\tilde{y}(x) = (c/2)((y_+(x) - h(y_+(x)))/(c/2 - h(y_+(x))))$. It follows that $y_+(x) \rightarrow c/(2\sqrt{3})$ as $x \rightarrow +\infty$. Since, by straightforward calculation, we can show that $h(y_+(x)) \rightarrow -\infty$ as $x \rightarrow +\infty$, it follows that $\tilde{y}(x) \rightarrow c/2$ as $x \rightarrow +\infty$. Since $\tilde{y}(x) < y(x) < c/2$, we have $y(x) \rightarrow c/2$ as $x \rightarrow +\infty$. Therefore, by (3.2) it follows that $I'(x) \rightarrow c/2 - \theta > 0$ as $x \rightarrow +\infty$. Similarly, we can show that $I'(x) \rightarrow -c/2 - \theta < 0$ as $x \rightarrow -\infty$. Therefore, there exists a unique point

$x = x(\theta)$ such that $I'(x) = y(x) - \theta = 0$. It follows that $I(x)$ is strictly decreasing on $(-\infty, x(\theta))$ and strictly increasing on $(x(\theta), +\infty)$, attaining its minimum for $x = x(\theta)$.

Now it follows that

$$(3.3) \quad P_\theta\left(\left|S_n - \frac{\theta}{1-\theta^2}\right| > \varepsilon\right) = P_\theta\left(S_n > \frac{\theta}{1-\theta^2} + \varepsilon\right) + P_\theta\left(S_n < \frac{\theta}{1-\theta^2} - \varepsilon\right).$$

We consider the following three cases: (a) $\theta/(1-\theta^2) + \varepsilon \leq x(\theta)$, (b) $\theta/(1-\theta^2) - \varepsilon < x(\theta) < \theta/(1-\theta^2) + \varepsilon$, (c) $x(\theta) \leq \theta/(1-\theta^2) - \varepsilon$. For the case (a), by Theorem 2.1 and the property of I , which was investigated above, it follows that

$$(3.4) \quad \begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n > \frac{\theta}{1-\theta^2} + \varepsilon\right) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n \geq \frac{\theta}{1-\theta^2} + \varepsilon\right) \\ &\leq - \inf_{x \geq \frac{\theta}{1-\theta^2} + \varepsilon} I(x) \\ &= -I(x(\theta)), \end{aligned}$$

and

$$(3.5) \quad \begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n > \frac{\theta}{1-\theta^2} + \varepsilon\right) &\geq - \inf_{x > \frac{\theta}{1-\theta^2} + \varepsilon} I(x) \\ &= -I(x(\theta)). \end{aligned}$$

By (3.4) and (3.5), we have

$$(3.6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n > \frac{\theta}{1-\theta^2} + \varepsilon\right) = -I(x(\theta)).$$

Similarly, by Theorem 2.1 and the property of I , we have

$$(3.7) \quad \begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n < \frac{\theta}{1-\theta^2} - \varepsilon\right) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n \leq \frac{\theta}{1-\theta^2} - \varepsilon\right) \\ &\leq - \inf_{x \leq \frac{\theta}{1-\theta^2} - \varepsilon} I(x) \\ &= -I\left(\frac{\theta}{1-\theta^2} - \varepsilon\right), \end{aligned}$$

and

$$(3.8) \quad \begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n < \frac{\theta}{1-\theta^2} - \varepsilon\right) &\geq - \inf_{x < \frac{\theta}{1-\theta^2} - \varepsilon} I(x) \\ &= -I\left(\frac{\theta}{1-\theta^2} - \varepsilon\right). \end{aligned}$$

By (3.7) and (3.8), we have

$$(3.9) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta\left(S_n < \frac{\theta}{1-\theta^2} - \varepsilon\right) = -I\left(\frac{\theta}{1-\theta^2} - \varepsilon\right).$$

Since $I(x(\theta)) < I(\theta/(1-\theta^2) - \epsilon)$, by (3.3), (3.6), and (3.9) we have

$$(3.10) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta \left(\left| S_n - \frac{\theta}{1-\theta^2} \right| > \epsilon \right) = -I(x(\theta)).$$

In a similar fashion, we can see that (3.10) holds for the case (c), and for the case (b) it holds that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta \left(\left| S_n - \frac{\theta}{1-\theta^2} \right| > \epsilon \right) = -\min \left\{ I \left(\frac{\theta}{1-\theta^2} - \epsilon \right), I \left(\frac{\theta}{1-\theta^2} + \epsilon \right) \right\}.$$

This completes the proof. $\square$

From Theorem 3.1, if $x(\theta) \neq \theta/(1-\theta^2)$ then for all sufficiently small $\epsilon > 0$, we have $\lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta(|S_n - \theta/(1-\theta^2)| > \epsilon) = -I(x(\theta))$. However, for any θ satisfying $x(\theta) = \theta/(1-\theta^2)$, we have $\lim_{n \rightarrow \infty} \frac{1}{n} \log P_\theta(|S_n - \theta/(1-\theta^2)| > \epsilon) = -\min\{I(\theta/(1-\theta^2) - \epsilon), I(\theta/(1-\theta^2) + \epsilon)\}$, that is, this limit does not attain the minimum of the rate function $I(x)$, even if $\epsilon > 0$ is sufficiently small.

Acknowledgments

The author would like to thank the anonymous referees for their valuable comments.

REFERENCES

- [1] Akahira, M. (1976). On the asymptotic efficiency of estimators in an autoregressive process, *Ann. Inst. Statist. Math.*, 28, 35-48.
- [2] Akahira, M. and Takeuchi, K. (1981). *Asymptotic Efficiency of Statistical Estimators*, Lecture Notes in Statistics, 7, Springer.
- [3] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*, Wiley.
- [4] Bryc, W. and Smolenski, W. (1993). On the large deviation principle for a quadratic functional of the autoregressive process, *Statist and Prob Letters*, 17, 281-285.
- [5] Bucklew, J. A. (1990). *Large Deviations Techniques in Decision, Simulation, and Estimation*, Wiley.
- [6] Dembo, A. and Zeitouni, O. (1993). *Large Deviations Techniques and Applications*, Jones and Bartlett.
- [7] Deuschel, J.-D. and Stroock, D. W. (1989). *Large Deviations*, Academic Press.
- [8] Ellis, R. S. (1984). Large deviations for a general class of random vectors, *Ann. Probab.*, 12, 1-12.