Statistics & Probability Letters 17 (1993) 281-285
North-Holland

13 July 1993

On the large deviation principle for a quadratic functional of the autoregressive process

Włodzimierz Bryc

Department of Mathematics, University of Cincinnati, OH, USA

Włodzimierz Smolenski

Department of Mathematics, The University of Kansas, Lawrence, KS, USA

Received October 1992
Revised November 1992

Abstract: For the sum of squares of an autoregressive system, the large deviation principle with the explicit rate function is established.

AMS 1991 Subject Classification: Primary 60F10.

Keywords: Large deviations; discrete stochastic control; quadratic cost.

1. Introduction

Let $\{\gamma_k\}_{k\in\mathbb{N}}$ be i.i.d. N(0, 1) r.v. and consider X_k defined by the recurrence: $X_0 = 0$,

$$X_{k+1} = aX_k + \gamma_{k+1}. \tag{1.1}$$

In this note we prove the large deviation principle for the quadratic functional $S_n = (1/n)\Sigma_{k=1}^n X_k^2$ and we give the explicit rate function. Benitz and Bucklew (1990) present the large deviation principle that covers quadratic functionals of a large class of stationary Gaussian sequences, provided the spectral density satisfies a certain technical condition (which is satisfied for a stationary solution of (1.1)). However, they point out that in general the rate function equation is transcendental. Their proof is based on the Grenander–Szegö theory, and it might be of some benefit to have an elementary proof for the situation as simple as (1.1).

Our original motivation for studying S_n came from a discrete optimal control problem, where S_n represents the optimal empirical quadratic cost in a non-adaptive setup, see e.g. Hall and Heyde (1980). Quadratic functional S_n occurs also as auxiliary objects in estimation problems, see Duflo, Senoussi and Touati (1991) and Weiss (1990), in electrical engineering, see Bucklew (1990, pp. 103–105) and in some problems of statistical physics, see Ellis (1985, Section III.4).

Correspondence to: Włodzimierz Bryc, Department of Mathematics, University of Cincinnati, Old Chemistry Bldg. (ML 25), Cincinnati, OH 45221-0025, USA.

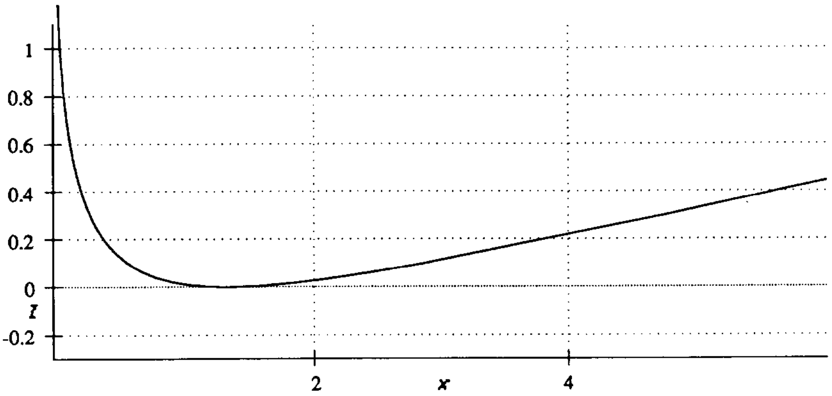

Fig. 1. Graph of $\mathbb{I}(x)$ for $a = \frac{1}{2}$.

Recall that S_n, $n \geqslant 1$, *satisfies the large deviation principle*, if there is a convex lower semicontinuous rate function $\mathbb{I} : \mathbb{R} \to [0, \infty]$ with compact level sets $\mathbb{I}^{-1}[0, a]$, $a \geqslant 0$, such that

$$\limsup_{n\to\infty} \frac{1}{n} \log P(S_n \in A) \leqslant -\inf_{x\in A} \mathbb{I}(x) \quad \text{for each closed set } A \subset \mathbb{R},$$

$$\liminf_{n\to\infty} \frac{1}{n} \log P(S_n \in A) \geqslant -\inf_{x\in A} \mathbb{I}(x) \quad \text{for each open set } A \subset \mathbb{R}.$$

The main result of this note is the following large deviation principle with the explicit rate function. (See Figure 1.)

Theorem 1. *If* $-1 < a < 1$ *then* $\{S_n\}$ *satisfies the large deviation principle with the rate function given by*

$$\mathbb{I}(x) = \begin{cases} -\frac{1}{2} \ln\left[\dfrac{2x}{1+\sqrt{4a^2x^2+1}}\right] + \frac{1}{2}[a^2+1]x - \frac{1}{2}\sqrt{4a^2x^2+1} & \textit{for } x < 0, \\ \infty & \textit{for } x \leqslant 0. \end{cases} \tag{1.2}$$

2. Proof

The following direct and rather elementary proof is based on a suitable variant of the well known Laplace transform criterion for the large deviation principle, see e.g. Dembo and Zeitouni (1993, Theorem 2.3.6).

Lemma A. *Suppose*

$$\mathbb{L}(y) = \lim_{n\to\infty} \frac{1}{n} \log E\{\exp(nyS_n)\} \tag{2.1}$$

exists as an extended number. Assume there is $\delta > 0$ *such that* $\mathbb{L}(y)$ *is finite and differentiable for* $-\infty < y < \delta$, *and*

$$\frac{\mathrm{d}}{\mathrm{d}y}\mathbb{L}(y) \to \infty \quad \textit{as } y \to \delta^-. \tag{2.2}$$

Then S_n satisfies the large deviation principle with the rate function

$$\mathbb{I}(x)=\sup_y\{xy-\mathbb{L}(y)\}. \qquad \square \tag{2.3}$$

Proof of Theorem 1. We shall verify that Lemma A can be applied with $\delta=\frac{1}{2}(1-|a|)^2$. Clearly, for $-1<a<1$ we have $\delta>0$.

Elementary integration of the normal density shows that for $\lambda<\frac{1}{2}$,

$$E\{\exp(\lambda X_{k+1}^2\mid \gamma_0,\dots,\gamma_k\}=(1-2\lambda)^{-1/2}\exp\left(\lambda^2a^2\frac{1+2\lambda}{1-2\lambda}X_k^2\right).$$

Therefore

$$E\{\exp(nyS_n)\}=\prod_{k=1}^{n}(1-2\lambda_k)^{-1/2}, \tag{2.4}$$

where the sequence $\{\lambda_n\}_{n\geqslant 0}$ solves the recurrence

$$\lambda_{n+1}=\lambda_n\frac{a^2}{1-2\lambda_n}+y \tag{2.5}$$

with the initial condition $\lambda_0=y<\frac{1}{2}$, provided that all the resulting numbers λ_n are strictly less than $\frac{1}{2}$; otherwise $E\{\exp(nyS_n)\}=\infty$.

We shall consider separately three cases: $y<\delta$, $y>\delta$ and $y=\delta$.

Case 1: $y<\delta$. In this case we shall show that the limit (2.1) exsits, is finite, differentiable and (2.2) holds. To this end, we use linear algebra to write the explicit expression for λ_n. Notice that λ_n is given by the n-fold composition $f^n(\lambda_0)$ of the Möbius function $f(x)$ associated with the matrix

$$\boldsymbol{M}=\begin{bmatrix} a^2-2y & y\\ -2 & 1\end{bmatrix}$$

in the (well known) correspondence

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

A calculation shows that for $-\infty<y<\delta:=\frac{1}{2}(1-|a|)^2$, $\boldsymbol{M}$ has two distinct positive eigenvalues, $\lambda_{\pm}=\frac{1}{2}(a^2+1-2y\pm((a^2+1-2y)^2-4a^2)^{1/2})$.

The remainder from the division of the polynomial x^n by the characteristic polynomial of $\boldsymbol{M}$ is

$$\frac{\lambda_+\lambda_-^n-\lambda_-\lambda_+^n}{\lambda_+-\lambda_-}+\frac{\lambda_+^n-\lambda_-^n}{\lambda_+-\lambda_-}x;$$

this is easily checked by evaluating the polynomials at $x=\lambda_{\pm}$. Therefore by the Cayley–Hamilton Theorem

$$\boldsymbol{M}^n=\frac{\lambda_+\lambda_-^n-\lambda_-\lambda_+^n}{\lambda_+-\lambda_-}\boldsymbol{I}+\frac{\lambda_+^n-\lambda_-^n}{\lambda_+-\lambda_-}\boldsymbol{M}.$$

Since the composition of Möbius functions is a Möbius function associated with the product of the corresponding matrices, therefore after some arithmetic,

$$\lambda_n=\frac{(1-q^n)q\lambda_-+a^2(1-q^n)+q^n-2q^ny+2y-1}{(1-q^n)q\lambda_-+q^n-2q^ny+2y-1}y, \tag{2.6}$$

where $0<q=\lambda_-/\lambda_+<1$. This shows that λ_n converges as $n\to\infty$ to the limit

$$g(y)=\frac{q\lambda_-+a^2+2y-1}{q\lambda_-+2y-1}y.$$

This simplifies to

$$g(y)=\tfrac{1}{4}\left[1-a^2+2y-\sqrt{a^4-2a^2+4y^2-4a^2y-4y+1}\right]. \tag{2.7}$$

Hence, by (2.4), the limit (2.1) exists for $-\infty<y<\delta$ and is given by $\mathbb{L}(y)=-\frac{1}{2}\log(1-2g(y))$; in particular, $\mathbb{L}(y)$ is finite and differentiable in the interval $-\infty<y<\delta$. Furthermore, the expression $a^4-2a^2+4y^2-4a^2y-4y+1$ vanishes at $y=\frac{1}{2}(1\pm a)^2$; hence $(\mathrm{d}/\mathrm{d}y)\mathbb{L}(y)=g'/(1-2g)$ satisfies (2.2) with $\delta=\frac{1}{2}(1-|a|)^2$.

Case 2: $y>\delta$. In this case we shall show that the expression under the limit in (2.1) becomes infinite starting from some $n\geqslant 1$. If $y\geqslant\frac{1}{2}$, then already the first term $n=1$ in (2.4) is infinite. If $y<\frac{1}{2}$, write again the recurrence (2.5) in the form $\lambda_{n+1}=f(\lambda_n)$. Clearly, for $x<\frac{1}{2}$ we have $f(x)\geqslant x$ and by calculus $f(x)-x$ attains the minimum value $y-\delta$ at $x=\frac{1}{2}(1-|a|)$. Therefore for $y>\delta$ we have $\lambda_n\geqslant y+n(y-\delta)$ exceeds $\frac{1}{2}$ eventually.

Case 3: $y=\delta$. Direct computation shows that $\lambda_1=2^{-1}|a|(1-|a|)^2(2-|a|)^{-1}>\frac{1}{2}(1-|a|)$. Therefore for $\lambda_1\leqslant x<\frac{1}{2}$ we have $f'(x)=a^2(1-2x)^{-2}\geqslant q>1$ and hence $f(x)>y+qx$. Thus from (2.5) we get $\lambda_{n+1}\geqslant y(1+q+\cdots+q^n)$ exceeds $\frac{1}{2}$ eventually.

Rate function identification. Formula (1.2) now follows by calculus from (2.3) and (2.7). The details are as follows. For $x<0$, $\mathbb{I}(x)=\lim_{y\to-\infty}\{xy-\mathbb{L}(y)\}$. Since $\lim_{y\to-\infty}\mathbb{L}(y)=0$, $\mathbb{I}(x)=\infty$. Let now $x>0$ be fixed. Clearly, $\mathbb{I}(x)=xy-\mathbb{L}(y)$, where y solves

$$\frac{\partial}{\partial y}\mathbb{L}(y)=x. \tag{2.8}$$

After some calculation, (2.8) gives

$$2y-z-a^2-1=xz(2y-z-a^2-1), \tag{2.9}$$

where

$$z=\sqrt{4y^2-4y+a^4-4ya^2-2a^2+1}. \tag{2.10}$$

Since $2y-z-a^2-1=-z-(z^2+a^2)^{1/2}<0$, (2.9) implies $z=1/x$. Since $y<\frac{1}{2}(1-|a|)^2$, from (2.10) $y=\frac{1}{2}(a^2+1-(z^2+4a^2)^{1/2})$. Using $z=1/x$, we now get $y=\frac{1}{2}(a^2+1-\sqrt{4a^2x^2+1}/x)$. After some calculation (1.2) now follows. □

Note added in proof

Theorem 1 has been extended by the first named author and A. Dembo in "Large deviations for quadratic functionals of Gaussian processes" (in preparation).

Acknowledgement

We would like to thank T.E. Duncan for suggesting the problem. We would also like to thank A. Dembo and the referee for pointing out the relevant references.

References

Benitz, G.R. and J.A. Bucklew (1990), Large deviation rate calculations for nonlinear detectors in Gaussian noise, *IEEE Trans. Inform. Theory* **36**, 358–371.

Bucklew, J.A. (1990) *Large Deviations Techniques in Decision, Simulation, and Estimation* (Wiley, New York).

Dembo, A. and O. Zeitouni (1993), *Large deviations and Applications* (Jones and Bartlett, Boston, MA).

Duflo, M., R. Senoussi and A. Touati (1991), Almost sure asymptotic properties of the least square estimators in a vectorial autoregressive model, *Ann. H. Poincaré Probab. Statist.* **27**(1), 1–25.

Ellis, R.S. (1985) *Entropy, Large Deviations and Statistical Mechanics* (Springer, New York).

Hall, P. and C.C. Heyde (1980), *Martingale Limit Theory and its Application* (Academic Press, New York).

Weiss, L. (1990), Maximum probability estimation for an autoregresive process, in: N.U. Prabhu and I.V. Basawa, eds., *Statistical Inference in Stochastic Processes* (Dekker, New York).