Jiří Anděl

The efficiency of estimates in stationary autoregressive series

Aplikace matematiky, Vol. 15 (1970), No. 1, 18–30

Persistent URL: http://dml.cz/dmlcz/103264

THE EFFICIENCY OF ESTIMATES IN STATIONARY AUTOREGRESSIVE SERIES

JIŘÍ ANDĚL

(Received May 30, 1968)

Let $\{X_t\}_{-\infty}^{\infty}$ be a stationary random sequence whose correlation function is known. Let its expectation be $EX_t = \alpha\varphi_t$ where φ_t is a given function and α an unknown parameter. Denote by $\hat{\alpha}_N$ the least squares estimate for α based on the random variables $X_1, \ldots, X_N$. This paper is devoted to the evaluation of the efficiency of $\hat{\alpha}_N$ for some frequently occurring functions φ_t. A theorem on asymptotic efficiency is formulated in the last section.

1. INTRODUCTION

Let $\boldsymbol{X} = (X_1, \ldots, X_N)'$ be a random vector with a regular covariance matrix $\boldsymbol{G}$ such that

$$EX_t = \alpha\varphi_t\,, \quad 1 \leqq t \leqq N\,,$$

where $\varphi_t\,(1 \leqq t \leqq N)$ is a given function and α is an unknown parameter. The matrix $\boldsymbol{G}$ is supposed to be known.

The estimate for α, say $\hat{\alpha}$, is called the least squares estimate, if $\hat{\alpha}$ minimizes the expression $\sum_{t=1}^{N}(X_t - \alpha\varphi_t)^2$.

Lemma 1. *Put* $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_N)'$. *If* $\boldsymbol{\varphi} \neq 0$ *then the least squares estimate* $\hat{\alpha}$ *is*

$$\hat{\alpha} = (\boldsymbol{\varphi}'\boldsymbol{\varphi})^{-1}\,\boldsymbol{\varphi}'\boldsymbol{X} \tag{1}$$

and has the variance

$$\operatorname{var}\hat{\alpha} = (\boldsymbol{\varphi}'\boldsymbol{\varphi})^{-2}\,\boldsymbol{\varphi}'\boldsymbol{G}\boldsymbol{\varphi}\,. \tag{2}$$

Proof is obvious. Compare with [3], p. 130.

Lemma 2. *Denote by $\tilde{\alpha}$ the best linear unbiased estimate for the parameter α (i.e., $\tilde{\alpha}$ is the linear unbiased estimate with minimal variance). Then for $\boldsymbol{\varphi} \neq \mathbf{0}$*

$$\tilde{\alpha} = (\varphi' \mathbf{G}^{-1} \varphi)^{-1} \varphi' \mathbf{G}^{-1} \mathbf{X} \tag{3}$$

and

$$\operatorname{var} \tilde{\alpha} = (\varphi' \mathbf{G}^{-1} \varphi)^{-1} . \tag{4}$$

Proof. See [1]. Compare with [3], p. 131.

Let us recall some advantages of the least squares estimate $\hat{\alpha}$. It may be simply evaluated and does not depend on the covariance matrix $\mathbf{G}$. The variance of $\hat{\alpha}$ may be easily determined by (2).

On the other hand the variance of the best linear unbiased estimate $\tilde{\alpha}$ may be substantially smaller than that of the least squares estimate $\hat{\alpha}$. But the evaluation of $\tilde{\alpha}$ is more complicated as $\tilde{\alpha}$ depends on $\mathbf{G}$, even through $\mathbf{G}^{-1}$. If N is not very small, the inversion of the matrix $\mathbf{G}$ is generally very difficult.

The efficiency e_N is introduced by

$$e_N = \operatorname{var} \tilde{\alpha} / \operatorname{var} \hat{\alpha}$$

as a measure of the quality of the estimate $\hat{\alpha}$. In our case we get with respect to (2) and (4)

$$e_N = (\varphi' \varphi)^2 \left[(\varphi' \mathbf{G} \varphi) (\varphi' \mathbf{G}^{-1} \varphi) \right]^{-1} . \tag{5}$$

A very important class of vectors $\mathbf{X} = (X_1, \ldots, X_N)'$ is the set of finite parts of an infinite random sequence $\{X_t\}_{-\infty}^{\infty}$ where $\mathrm{E}X_t^2 < \infty$, $\operatorname{cov}(X_t, X_s) = \operatorname{cov}(X_{t+r}, X_{s+r})$ for all integers t, s, r. Related problems were considered in mathematical papers from different viewpoints. See e.g. [3], [4], [5]. Such questions as those of the lower limit for e_N and of the asymptotic efficiency of the estimate $\hat{\alpha}$ were solved.

In practice, however, the economic, hydrologic and some other important series usually have not such a length that we may rely on the asymptotic properties of e_N only. It seems to be suitable to determine e_N exactly at least for the most current types of series $\{X_t\}_{-\infty}^{\infty}$ and for the most usual functions φ_t.

In the particular assertions we shall specify the correlation function of the series $\{X_t\}_{-\infty}^{\infty}$ (directly or by means of the spectral density) and the function φ_t occuring in the expectations of the series:

$$\mathrm{E}X_t = \alpha \varphi_t , \quad t = 0, \pm 1, \pm 2, \ldots .$$

We keep the following notation: $\mathbf{G}$ is the covariance matrix of the random variables $X_1, \ldots, X_N$; e_N is the efficiency of the least squares estimate $\hat{\alpha}_N$ where $\hat{\alpha}_N$ is based on $X_1, \ldots, X_N$ only. Let $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_N)'$ and $\mathbf{X} = (X_1, \ldots, X_N)'$.

2. THE ESTIMATE OF THE CONSTANT EXPECTATION IN THE AUTOREGRESSIVE SERIES OF THE FIRST ORDER

Theorem 1. *Let* $\{X_t\}_{-\infty}^{\infty}$ *be a series of the random variables such that*

$$\mathrm{E}X_t = \alpha\,, \quad \operatorname{cov}(X_t, X_s) = (1 - a^2)^{-1}\, a^{|t-s|}$$

where t, s are integers and $0 \neq a \in (-1, 1)$. *Then*

$$e_N = N^2\{[N + 2a/(1 - a)]\,[N - 2a(1 - a^N)/(1 - a^2)]\}^{-1} \tag{6}$$

holds.

Proof. In this case $\varphi_t \equiv 1$. The spectral density

$$f(\lambda) = (2\pi)^{-1}\,|1 - ae^{i\lambda}|^{-2}\,, \quad -\pi \leqq \lambda \leqq \pi$$

corresponds to our correlation function $B(k) = (1 - a^2)^{-1}\, a^{|k|}$ as it is well known. We have

$$\boldsymbol{G} = (1 - a^2)^{-1} \begin{Vmatrix} 1 & a & a^2 & \ldots & a^{N-1} \\ a & 1 & a & \ldots & a^{N-2} \\ a^2 & a & 1 & \ldots & a^{N-3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a^{N-1} & a^{N-2} & a^{N-3} & \ldots & 1 \end{Vmatrix} \tag{7}$$

and with regard to [2], pp. 424, 425, for $N \geqq 3$

$$\boldsymbol{G}^{-1} = \begin{Vmatrix} 1 & -a & 0 & 0 \ldots & 0 & 0 & 0 \\ -a & 1 + a^2 & -a & 0 \ldots & 0 & 0 & 0 \\ 0 & -a & 1 + a^2 & -a \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 \ldots & -a & 1 + a^2 & -a \\ 0 & 0 & 0 & 0 \ldots & 0 & -a & 1 \end{Vmatrix}.$$

As $\varphi_t = 1$, $1 \leqq t \leqq N$, we have $\boldsymbol{\varphi}'\boldsymbol{\varphi} = N$. As a matter of fact, $\boldsymbol{\varphi}'\boldsymbol{G}\boldsymbol{\varphi}$ is the sum of elements of the matrix $\boldsymbol{G}$ and it will be denoted by $\mathrm{S}(\boldsymbol{G})$. Analogously $\boldsymbol{\varphi}'\boldsymbol{G}^{-1}\boldsymbol{\varphi}$ as the sum of elements of $\boldsymbol{G}^{-1}$ will be denoted by $S(\boldsymbol{G}^{-1})$.

For a natural N and $b \neq 1$ it may be proved by induction that the following formulas holds:

$$\sum_{k=1}^{N} kb^k = b(1 - b^N)(1 - b)^{-2} - Nb^{N+1}(1 - b)^{-1}\,, \tag{8}$$

$$\sum_{k=1}^{N} k^2 b^k = 2b^2(1 - b^{N-1})(1 - b)^{-3} + b[1 - (2N - 1)\,b^N]\,(1 - b)^{-2} - N^2 b^{N+1}(1 - b)^{-1}\,. \tag{9}$$

Using (8) we get

$$\mathrm{S}(\boldsymbol{G}) = (1 - a^2)^{-1}\{N + 2a(1 - a)^{-1}\,[N - (1 - a^N)(1 - a)^{-1}]\}\,.$$

Obviously

$$S(\boldsymbol{G}^{-1}) = N + a^2(N - 2) - 2a(N - 1)\,.$$

Substituting into (5) we get after some modification the assertion of Theorem 1.

Theorem 1 implies these **corollaries:**

a) *For fixed a and $N \to \infty$ we have $e_N \to 1$.*

b) *For fixed a and $N \to \infty$*

$$e_N = 1 - 2a^2(1 - a^2)^{-1} N^{-1} + o(N^{-1})$$

holds.

c) *Denote the efficiency e_N given in (6) more explicitly by $e_N(a)$, in order to emphasize its dependence on a. Then*

$$\lim_{a \to 1^-} e_N(a) = 0\,,$$

$$\lim_{a \to -1^+} e_N(a) = 0 \qquad \textit{for } N \textit{ odd}, \; N \geqq 3\,,$$

$$\lim_{a \to -1^+} e_N(a) = \tfrac{1}{2}N(N - 1)^{-1} \qquad \textit{for } N \textit{ even}, \; N \geqq 4\,.$$

The values of $e_N(a)$ are given in Table 1 for various N and a.

Table 1

THE EFFICIENCY $e_N(a)$

a \ N	3	4	5	6	7	8	9	10	15	20	30	50	100	500
−0·95	0·113	0·694	0·157	0·642	0·203	0·626	0·248	0·622	0·369	0·647	0·684	0·750	0·846	0·964
−0·90	0·226	0·722	0·307	0·683	0·385	0·677	0·454	0·683	0·607	0·741	0·793	0·858	0·922	0·983
−0·80	0·439	0·776	0·558	0·759	0·648	0·769	0·711	0·786	0·813	0·858	0·898	0·935	0·966	0·993
−0·70	0·618	0·828	0·729	0·826	0·796	0·842	0·836	0·860	0·894	0·917	0·942	0·964	0·981	0·996
−0·60	0·758	0·874	0·838	0·881	0·878	0·897	0·901	0·911	0·936	0·950	0·965	0·979	0·989	0·998
−0·50	0·857	0·914	0·905	0·923	0·927	0·936	0·940	0·945	0·961	0·970	0·979	0·987	0·993	0·999
−0·40	0·923	0·947	0·947	0·954	0·958	0·962	0·966	0·968	0·978	0·983	0·988	0·993	0·996	0·999
−0·30	0·964	0·972	0·973	0·976	0·978	0·980	0·982	0·984	0·988	0·991	0·994	0·996	0·998	1·000
−0·20	0·987	0·988	0·989	0·990	0·991	0·992	0·993	0·995	0·996	0·996	0·997	0·998	0·999	1·000
−0·10	0·997	0·997	0·997	0·998	0·998	0·998	0·998	0·998	0·999	0·999	0·999	1·000	1·000	1·000
0·10	0·998	0·998	0·998	0·998	0·998	0·998	0·998	0·998	0·999	0·999	0·999	1·000	1·000	1·000
0·20	0·994	0·992	0·992	0·992	0·992	0·993	0·993	0·994	0·995	0·996	0·997	0·998	0·999	1.000
0·30	0·989	0·984	0·983	0·983	0·984	0·984	0·985	0·986	0·989	0·992	0·994	0·996	0·998	1·000
0·40	0·985	0·977	0·973	0·972	0·972	0·973	0·974	0·975	0·981	0·984	0·989	0·993	0·996	0·999
0·50	0·982	0·970	0·963	0·960	0·959	0·959	0·960	0·961	0·968	0·974	0·981	0·988	0·994	0·999
0·60	0·980	0·965	0·955	0·950	0·946	0·945	0·945	0·945	0·952	0·960	0·970	0·980	0·989	0·999
0·70	0·981	0·965	0·952	0·943	0·937	0·933	0·931	0·930	0·933	0·940	0·953	0·968	0·982	0·996
0·80	0·984	0·969	0·956	0·945	0·937	0·930	0·923	0·921	0·913	0·915	0·927	0·946	0·969	0·993
0·90	0·990	0·980	0·970	0·961	0·953	0·945	0·938	0·933	0·912	0·901	0·896	0·906	0·936	0·984
0·95	0·995	0·989	0·983	0·977	0·971	0·966	0·960	0·955	0·935	0·920	0·901	0·887	0·899	0·967

3. THE ESTIMATE OF THE CONSTANT EXPECTATION IN THE AUTOREGRESSIVE SERIES OF THE SECOND ORDER

Theorem 2. *Let $\{X_t\}_{-\infty}^{\infty}$ be a series of the random variables with the expectation* $EX_t \equiv \alpha$ *and with the spectral density*

$$f(\lambda) = (2\pi)^{-1} |a_2 + a_1 e^{i\lambda} + a_0 e^{2i\lambda}|^{-2} . \tag{11}$$

Suppose a_0, a_1, a_2 are real numbers such that the polynomial $a_2 z^2 + a_1 z + a_0 = a_2(z - b_1)(z - b_2)$ has roots b_1, b_2 whose absolute value is greater than 0 and smaller than 1. Then we have for $b_1 \neq b_2$

$$e_N = N^2\{N - 2[(b_1 - b_2)(1 - b_1 b_2)]^{-1} [b_1^2(1 - b_2)^2 (1 - b_1^N)/(1 - b_1^2) - b_2^2(1 - b_1)^2 (1 - b_2^N)/(1 - b_2^2)]\}^{-1} \{N + 2(b_1 - 2b_1 b_2 + b_2) . [(1 - b_1)(1 - b_2)]^{-1}\}^{-1} \tag{12}$$

and for $b_1 = b_2 = b$

$$e_N = N^2[N - 4b(1 + b + b^2)(1 - b)^{-1}(1 + b)^{-3} + 2b^{N+1}(2 + N + 2b + 2b^2 - Nb^2)(1 - b)^{-1}(1 + b)^{-3}]^{-1} . [N + 4b(1 - b)^{-1}]^{-1} . \tag{13}$$

Proof. Without a loss of generality suppose $a_2 = 1$. Hence

$$a_0 = b_1 b_2 , \quad a_1 = -b_1 - b_2 , \quad a_2 = 1 . \tag{14}$$

The spectral density (11) may be written in the form

$$f(\lambda) = (2\pi)^{-1} |(e^{i\lambda} - b_1)(e^{i\lambda} - b_2)|^{-2} , \quad -\pi \leqq \lambda \leqq \pi .$$

From the well-known spectral representation of the correlation function

$$B(k) = \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda)\, d\lambda \tag{15}$$

we evaluate the elements of the covariance matrix $\boldsymbol{G}$.

Suppose $b_1 \neq b_2$. Then we get from (15)

$$B(k) = A_1 b_1^{|k|} + A_2 b_2^{|k|} \tag{16}$$

where

$$A_1 = b_1[(b_1 - b_2)(1 - b_1 b_2)(1 - b_1^2)]^{-1} , \tag{17}$$

$$A_2 = b_2[(b_2 - b_1)(1 - b_1 b_2)(1 - b_2^2)]^{-1} . \tag{18}$$

The covariance matrix $\mathbf{G}$ is

(19) $$\mathbf{G} = A_1\mathbf{G}_1 + A_2\mathbf{G}_2$$

where

(20) $$\mathbf{G}_i = \left\| \begin{matrix} 1 & b_i & b_i^2 & \dots & b_i^{N-1} \\ b_i & 1 & b_i & \dots & b_i^{N-2} \\ \dots & \dots & \dots & \dots & \dots \\ b_i^{N-1} & b_i^{N-2} & b_i^{N-3} & \dots & 1 \end{matrix} \right\|$$

for $i = 1, 2$.

If $N \geqq 5$, then we obtain by [2] this inverse matrix $\mathbf{G}^{-1}$:

$$\left\| \begin{matrix} a_2^2 & a_2a_1 & a_2a_0 & 0 & \dots & 0 & 0 \\ a_2a_1 & a_2^2+a_1^2 & a_2a_1+a_1a_0 & a_2a_0 & \dots & 0 & 0 \\ a_2a_0 & a_2a_1+a_1a_0 & a_2^2+a_1^2+a_0^2 & a_2a_1+a_1a_0 & \dots & 0 & 0 \\ 0 & a_2a_0 & a_2a_1+a_1a_0 & a_2^2+a_1^2+a_0^2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_2^2+a_1^2+a_0^2 & a_2a_1+a_1a_0 & a_2a_0 & 0 \\ 0 & 0 & \dots & a_2a_1+a_1a_0 & a_2^2+a_1^2+a_0^2 & a_2a_1+a_1a_0 & a_2a_0 \\ 0 & 0 & \dots & a_2a_0 & a_2a_1+a_1a_0 & a_2^2+a_1^2 & a_2a_1 \\ 0 & 0 & \dots & 0 & a_2a_0 & a_2a_1 & a_2^2 \end{matrix} \right\| .$$

The sum of elements of the matrix $\mathbf{G}^{-1}$ gives

$$S(\mathbf{G}^{-1}) = Na_2^2 + (N-2)\,a_1^2 + (N-4)\,a_0^2 + \\ + 2[a_2a_1(N-1) + a_1a_0(N-3) + a_2a_0(N-2)] .$$

With respect to (14) we get

(21) $$S(\mathbf{G}^{-1}) = (1-b_1)^2\,(1-b_2)^2\,N + 2(1-b_1)\,(1-b_2)\,(b_1 - 2b_1b_2 + b_2) .$$

The sum of elements of the matrix $\mathbf{G}$ is

$$S(\mathbf{G}) = A_1\{N + 2b_1[N - (1-b_1^N)/(1-b_1)]/(1-b_1)\} + \\ + A_2\{N + 2b_2[N - (1-b_2^N)/(1-b_2)]/(1-b_2)\} .$$

Substituting from (17) and (18) we have

(22) $$S(\mathbf{G}) = N[(1-b_1)^2\,(1-b_2)^2]^{-1} - 2[(b_1-b_2)\,(1-b_1b_2)]^{-1} . \\ . \{b_1^2(1-b_1^N)\,[(1-b_1^2)\,(1-b_1)^2]^{-1} - b_2^2(1-b_2^N)\,[(1-b_2^2)\,(1-b_2)^2]^{-1}\} .$$

According to (5), in our case $e_N = N^2[S(\mathbf{G})\,S(\mathbf{G}^{-1})]^{-1}$ so that (21) and (22) imply after some arrangements formula (12).

Now, we return to the case $b_1 = b_2 = b$, $0 \neq b \in (-1, 1)$. We have

$$f(\lambda) = (2\pi)^{-1} \left|e^{i\lambda} - b\right|^{-4}$$

and from (15) there follows

$$B(k) = (1 + b^2)(1 - b^2)^{-3} b^k + (1 - b^2)^{-2} kb^k, \quad k \geqq 0.$$

Using formulas (8) and (9) we obtain the sum of elements of the matrix $\mathbf{G}$

$$\begin{aligned} S(\mathbf{G}) = N(1 - b)^{-4} - 4b(1 + b + b^2)(1 - b)^{-2}(1 - b^2)^{-3} + \\ + 2b^{N+1}(2 + N + 2b + 2b^2 - Nb^2)(1 - b)^{-2}(1 - b^2)^{-3}. \end{aligned}$$

The sum of elements of the matrix $\mathbf{G}^{-1}$ is given by (21) putting $b_1 = b_2 = b$ because we did not use the assumption $b_1 \neq b_2$ in the proof of (21). Consequently

$$S(\mathbf{G}^{-1}) = (1 - b)^4 N + 4b(1 - b)^3.$$

These results imply (13).

Corollaries of Theorem 2.

a) $\lim_{N\to\infty} e_N = 1$ *in both cases* $b_1 \neq b_2$ *and* $b_1 = b_2$.

b) *If* $N \to \infty$ *we get these asymptotic formulas*:

(23) $$\begin{aligned} e_N = 1 - 2[(b_1^2 + b_2^2)(1 - b_1 b_2) + 2b_1 b_2(1 - b_1 - b_2 + b_1^2 b_2^2)] \,. \\ . [(1 - b_1 b_2)(1 - b_1^2)(1 - b_2^2)]^{-1} N^{-1} + o(N^{-1}) \quad \text{for} \quad b_1 \neq b_2, \end{aligned}$$

(24) $$\begin{aligned} e_N = 1 - 4b^2(2 + 2b + b^2)(1 - b)^{-1}(1 + b)^{-3} N^{-1} + o(N^{-1}) \\ \text{for} \quad b_1 = b_2 = b. \end{aligned}$$

c) (13) *is the limit of* (12) *if we put e.g.* $b_1 = b$, $b_2 \to b$. *Analogously* (24) *is the limit of* (23).

4. THE ESTIMATE OF THE CONSTANT EXPECTATION IN THE AUTOREGRESSIVE SERIES OF THE n-TH ORDER

Theorem 3. *Let* $\{X_t\}_{-\infty}^{\infty}$ *be a series of the random variables with the expectation* $EX_t \equiv \alpha$ *and the sepctral density*

(25) $$f(\lambda) = (2\pi)^{-1} \left|(e^{i\lambda} - b_1) \dots (e^{i\lambda} - b_n)\right|^{-2}$$

where all the numbers $b_1, \dots, b_n$ *are different,* $0 < |b_j| < 1$ *for* $1 \leqq j \leqq n$ *and all the coefficients* $\alpha_0, \dots, \alpha_n$ *of the polynomial*

$$P(z) = (z - b_1) \dots (z - b_n) = \alpha_0 z^n + \alpha_1 z^{n-1} + \dots + \alpha_n$$

are real. Put $a_j = \alpha_{n-j}$, $0 \leqq j \leqq n$.

Then the efficiency e_N of the least squares estimate $\hat{\alpha}$ is

(26) $$e_N = N^2[\mathrm{S}(\boldsymbol{G})\,\mathrm{S}(\boldsymbol{G}^{-1})]^{-1}$$

where $\mathrm{S}(\boldsymbol{G})$ *and* $\mathrm{S}(\boldsymbol{G}^{-1})$ *are the sums of elements of* $\boldsymbol{G}$ *and* $\boldsymbol{G}^{-1}$, *respectively. Further we have*

(27) $$\mathrm{S}(\boldsymbol{G}) = \sum_{j=1}^{n} A_j\{N + 2b_j[N - (1 - b_j^N)/(1 - b_j)]/(1 - b_j)\}$$

where for $1 \leqq j \leqq n$

(28) $$A_j = b_j^{n-1}[(b_j - b_1)\ldots(b_j - b_{j-1})(b_j - b_{j+1})\ldots(b_j - b_n)\,. \\ .\,(1 - \bar{b}_1 b_j)\ldots(1 - \bar{b}_j b_j)\ldots(1 - \bar{b}_n b_j)]^{-1}$$

and for $N \geqq 2n + 1$ *we have*

(29) $$\mathrm{S}(\boldsymbol{G}^{-1}) = N\sum_{j=0}^{n} a_j^2 - 2\sum_{j=1}^{n} j a_{n-j}^2 + 2N\sum_{k=1}^{n}\sum_{j=0}^{n-k} a_{n-j}a_{n-j-k} - \\ - 2\sum_{k=1}^{n}\sum_{j=0}^{n-k}(2j + k)\,a_{n-j}a_{n-j-k}\,.$$

Proof. From (25) and (15) we evaluate the correlation function

$$B(k) = \sum_{j=1}^{n} A_j b_j^k\,, \quad k = 0, 1, \ldots$$

where A_j are given by (28). The sum of elements of the matrix $\boldsymbol{G}$ may be evaluated analogously as in previous sections. This leeds to (27).

We may write the spectral density (25) equivalently in the form

(30) $$f(\lambda) = (2\pi)^{-1}\,\Big|\sum_{k=0}^{n} a_{n-k}e^{ik\lambda}\Big|^{-2}$$

where all the coefficients $a_j = \alpha_{n-j}$ are real and all the roots of the equation $\sum_{k=0}^{n} a_{n-k}z^k = 0$ have the absolute value greater than 1.

Now, we use the following assertion which is proved in [2]. Put $\boldsymbol{G}^{-1}_{(N+1)\times(N+1)} = (g_{ts})_{t,s=0,1,\ldots,N}$. Then for $N \geqq 2n$

$$g_{ts} = \begin{cases} \displaystyle\sum_{i=0}^{\min(N-t,N-s,n-|t-s|)} a_{n-i}a_{n-i-|t-s|} & \text{for } \max(t, s) > N - n\,, \\ \displaystyle\sum_{i=0}^{n-|t-s|} a_{n-i}a_{n-i-|t-s|} & \text{for } n \leqq t, s \leqq N - n\,, \\ \displaystyle\sum_{i=0}^{\min(t,s,n-|t-s|)} a_{n-i}a_{n-i-|t-s|} & \text{for } \min(t, s) < n\,. \end{cases}$$

For $|t - s| > n$ there is $g_{ts} = 0$. If $\min(t, s) < n \leqq N - n < \max(t, s)$ then the first row of the previous formula for g_{ts} gives the same results as the third one. We must be aware of the fact that in our case the matrix $\mathbf{G}^{-1}$ has the order $N \times N$ only. From here we obtain the sum $S(\mathbf{G}^{-1})$ given by formula (29).

In the case of equality of some b'_j we could obtain the results as the limit from (29).

5. THE ESTIMATE OF THE LINEAR EXPECTATION IN THE AUTOREGRESSIVE SERIES OF THE FIRST ORDER

Theorem 4. *Let $\{X_t\}_{-\infty}^{\infty}$ be a sequence of random variables with the expectation $EX_t = \alpha t$, $t = 0, \pm 1, \pm 2, \ldots$, and the correlation function given in Theorem 1. Then the efficiency e_N is given by (5) where*

$$\varphi'\varphi = \tfrac{1}{6}N(N + 1)(2N + 1),$$

$$\varphi'\mathbf{G}\varphi = (1 - a^2)^{-1}\left[\tfrac{1}{6}N(N + 1)(2N + 1)(1 + a)/(1 - a) + a^2(2 + N - aN)(1 - a^N)(1 - a)^{-4} - (1 + N - aN + a^{N+1})\, aN(1 - a)^{-3}\right],$$

$$\varphi'\mathbf{G}^{-1}\varphi = \tfrac{1}{6}N(N + 1)(2N + 1)(1 - a)^2 + a(N + N^2 - a - aN^2).$$

Proof. In this case $\varphi_t = t$, $1 \leqq t \leqq N$ so that $\varphi'\varphi = \sum_{t=1}^{N} t^2$. The derivation of $\varphi'\mathbf{G}\varphi$ and $\varphi'\mathbf{G}^{-1}\varphi$ is obvious as in Section 2 both $\mathbf{G}$ and $\mathbf{G}^{-1}$ are given.

Note a fact that $\lim_{N\to\infty} e_N = 1$ holds, too.

6. THE ESTIMATE OF THE LINEAR EXPECTATION IN THE AUTOREGRESSIVE SERIES OF THE SECOND ORDER

Theorem 5. *Let $\{X_t\}_{-\infty}^{\infty}$ be a series of random variables with the expectation $EX_t = \alpha t$, $t = 0, \pm 1, \pm 2, \ldots$ and the correlation function (16) where $b_1 \neq b_2$. Then $\varphi_t = t$, $1 \leqq t \leqq N$ and the efficiency e_N is given by (5) where*

$$\varphi'\varphi = \tfrac{1}{6}N(N + 1)(2N + 1),$$

$$\varphi'\mathbf{G}\varphi = b_1[(b_1 - b_2)(1 - b_1b_2)(1 - b_1^2)]^{-1}\left[\tfrac{1}{6}N(N + 1)(2N + 1)(1 + b_1) : (1 - b_1) + b_1^2(2 + N - b_1N)(1 - b_1^N)(1 - b_1)^{-4} - b_1N(1 + N - b_1N + b_1^{N+1})(1 - b_1)^{-3}\right] + b_2[(b_2 - b_1)(1 - b_1b_2)(1 - b_2^2)]^{-1}\left[\tfrac{1}{6}N(N + 1)(2N + 1)(1 + b_2) : (1 - b_2) + b_2^2(2 + N - b_2N)(1 - b_2^N)(1 - b_2)^{-4} - b_2N(1 + N - b_2N + b_2^{N+1})(1 - b_2)^{-3}\right],$$

$$\varphi' \mathbf{G}^{-1} \varphi = \tfrac{1}{6} N(N+1)(2N+1)(1-b_1)^2 (1-b_2)^2 + 4b_1^2 b_2 + 4b_1 b_2^2 - b_1^2 - b_2^2 - \\ - 6b_1^2 b_2^2 + N(1 - b_1 b_2)(b_1 + b_2 - 2b_1 b_2) + N^2(b_1 - 2b_1 b_2 + b_2) \,. \\ . (1 - b_1)(1 - b_2).$$

Proof. We get these relations using $\mathbf{G}$ and $\mathbf{G}^{-1}$ from Section 3. The formula $\varphi' \mathbf{G} \varphi$ may be simplified but not substantially. The simplification would concern mainly the coefficient by N^3 which equals to $[3(1 - b_1)^2 (1 - b_2)^2]^{-1}$. We see from here that $\lim_{N \to \infty} e_N = 1$ holds.

The case $b_1 = b_2$ we do not described here. It may be obtain by the limit procedure.

7. THE EFFICIENCY OF THE ESTIMATE OF THE CONSTANT EXPECTATION IN THE AUTOREGRESSIVE SERIES OF THE SECOND ORDER IF THE ESTIMATE IS BASED ON THE ASSUMPTION OF THE AUTOREGRESSIVE SERIES OF THE FIRST ORDER

Theorem 6. *Let $X_1, \ldots, X_N$ be random variables with expectations* $\mathrm{E}X_t = \alpha$, $1 \leqq t \leqq N$ *and with the covariance matrix* (19) *which will be denoted here by* $\mathbf{T}$. *Let $\mathbf{G}$ be the covariance matrix* (7). *Then the efficiency of the best linear unbiased estimate for α based on the covariance matrix $\mathbf{G}$ is*

$$e_N = (\varphi' \mathbf{G}^{-1} \varphi)^2 \, [(\varphi' \mathbf{T}^{-1} \varphi)(\varphi' \mathbf{G}^{-1} \mathbf{T} \mathbf{G}^{-1} \varphi)]^{-1} \tag{31}$$

where $\varphi = (1, 1, \ldots, 1)'$, $\varphi' \mathbf{G}^{-1} \varphi$ *is given by formula* (10), $\varphi' \mathbf{T}^{-1} \varphi$ *is given by* (21) *and*

$$\varphi' \mathbf{G}^{-1} \mathbf{T} \mathbf{G}^{-1} \varphi = (1 - a)^4 \{ b_1 [(b_1 - b_2)(1 - b_1 b_2)(1 - b_1^2)]^{-1} [N + 2b_1(N - \\ - (1 - b_2^N)/(1 - b_1))/(1 - b_1)] + \\ + b_2 [(b_2 - b_1)(1 - b_1 b_2)(1 - b_2^2)]^{-1} \,. \\ . [N + 2b_2(N - (1 - b_2^N)/(1 - b_2))/(1 - b_2)] \} + \\ + 4a(1 - a)^3 [(1 - b_1^N)/(1 - b_1) + (1 - b_2^N)/(1 - b_2)] + \\ + 2a^2(1 - a)^2 (2 + b_1^{N-1} + b_2^{N-1}) \,.$$

Proof. The best linear unbiased estimate based on the matrix $\mathbf{G}$ is

$$\alpha^* = (\varphi' \mathbf{G}^{-1} \varphi)^{-1} \, \varphi' \mathbf{G}^{-1} \mathbf{X}$$

and has the variance

$$\operatorname{var} \alpha^* = (\varphi' \mathbf{G}^{-1} \varphi)^{-2} \, \varphi' \mathbf{G}^{-1} \mathbf{T} \mathbf{G}^{-1} \varphi \,.$$

The best linear unbiased estimate

$$\tilde{\alpha} = (\varphi' \mathbf{T}^{-1} \varphi)^{-1} \, \varphi' \mathbf{T}^{-1} \mathbf{X}$$

has the variance

$$\operatorname{var} \tilde{\alpha} = (\varphi' \mathbf{T}^{-1} \varphi)^{-1} .$$

With respect to the definition of the efficiency

$$e_N = \operatorname{var} \tilde{\alpha} / \operatorname{var} \alpha^*$$

we obtain (31).

As for the evaluation of $\varphi' \mathbf{G}^{-1} \mathbf{T} \mathbf{G}^{-1} \varphi$, we have

$$\varphi' \mathbf{G}^{-1} = ((1-a), (1-a)^2, \ldots, (1-a)^2, (1-a)) .$$

Introduce the vectors ψ and ε:

$$\psi = (1-a)^2 \varphi , \quad \varepsilon' = \varphi' \mathbf{G}^{-1} - \psi' .$$

Obviously

$$\varphi' \mathbf{G}^{-1} \mathbf{T} \mathbf{G}^{-1} \varphi = \psi' \mathbf{T} \psi + 2\psi' \mathbf{T} \varepsilon + \varepsilon' \mathbf{T} \varepsilon .$$

But $(1-a)^{-4} \psi' \mathbf{T} \psi$ is the sum of elements of the matrix $\mathbf{T}$ given in Section 3, formula (22). Further

$$\begin{aligned} \varepsilon' \mathbf{T} \varepsilon &= 2a^2(1-a)^2 (2 + b_1^{N-1} + b_2^{N-1}) , \\ \varepsilon' \mathbf{T} \psi &= 2a(1-a)^3 [(1-b_1^N)/(1-b_1) + (1-b_2^N)/(1-b_2)] . \end{aligned}$$

This concludes the proof.

From (31) there easily follows $\lim_{N \to \infty} e_N = 1$ for an arbitrary $a \in (-1, 1)$, $a \neq 0$, and arbitrary b_1 and b_2, $b_1 \neq b_2$ satisfying assumptions of Theorem 2. For $a = 0$ this assertion was proved as a corollary of Theorem 2.

8. ASYMPTOTIC EFFICIENCY OF SOME ESTIMATES OF THE CONSTANT EXPECTATION

Theorem 7. *Let $\{X_t\}_{-\infty}^{\infty}$ be a stationary autoregressive series of the order m with an unknown expectation $\mathrm{E}X_t \equiv \alpha$. Let $\mathbf{T}$ be the covariance matrix of the random variables $X_1 \ldots, X_N$. Denote by $\mathbf{G}$ the covariance matrix of the random variables $X_1, \ldots, X_N$, supposing $\{X_t\}_{-\infty}^{\infty}$ to be an autoregressive series of the order n with the spectral density (25). Let e_N be the efficiency of the best linear unbiased estimate for α based on the variables $X_1, \ldots, X_N$ when the covariance matrix $\mathbf{G}$ is supposed. Then $\lim_{N \to \infty} e_N = 1$.*

Proof. Analogously to the proof of Theorem 6 we obtain

$$e_N = \varphi' \mathbf{G}^{-1} \varphi \, \varphi' \mathbf{G}^{-1} \varphi (\varphi' \mathbf{T}^{-1} \varphi \, \varphi' \mathbf{G}^{-1} \mathbf{T} \mathbf{G}^{-1} \varphi)^{-1} \tag{32}$$

where $\boldsymbol{\varphi} = (1, 1, \dots, 1)'$. Put $\boldsymbol{\varphi}'\boldsymbol{\varphi} = \mathbf{E}$. From the formula for the elements of the matrix $\mathbf{G}^{-1}$ given in the proof of Theorem 3 it follows that $\mathbf{E}\mathbf{G}^{-1} = q(\mathbf{E} + \mathbf{J})$, where $q = (\sum_{k=0}^{n} a_{n-k})^2$ and the matrix $\mathbf{J}$ may have non-zero elements in the first n and last n columns only. Actually, the sum of the elements of the k-th column $(n + 1 \leqq k \leqq$ $\leqq N - n)$ of the matrix $\mathbf{G}^{-1}$ gives

$$q = 2a_n a_0 + 2(a_n a_1 + a_{n-1} a_0) + \dots + 2(a_n a_{n-1} + \dots + a_1 a_0) + (a_n^2 + \dots + a_0^2) = \\ = (\sum_{k=0}^{n} a_{n-k})^2 .$$

Obviously, all the non-zero elements of the matrix $\mathbf{J}$ may be bounded in the absolute value by a finite constant $K > 0$, where K does not depend on N.

From the assumptions made on $f(\lambda)$ in (25) there follows that the equation $\sum_{k=0}^{n} a_{n-k} z^k = 0$ has all the roots inside of the unit circle. Therefore $z = 1$ is not the root so that $\sum_{k=0}^{n} a_{n-k} \neq 0$ and $q = (\sum_{k=0}^{n} a_{n-k})^2 > 0$.

As $\mathbf{T}$ is a symmetrical matrix, there holds $\mathbf{ET} = \mathbf{TE}$. Using the obvious relation $\mathbf{E}\boldsymbol{\varphi} = N\boldsymbol{\varphi}$ we may write the efficiency (32) in the form

$$(33) \qquad e_N = (N\boldsymbol{\varphi}'\mathbf{G}^{-1}\boldsymbol{\varphi} + \boldsymbol{\varphi}'\mathbf{G}^{-1}\mathbf{J}\boldsymbol{\varphi})(N\boldsymbol{\varphi}'\mathbf{G}^{-1}\boldsymbol{\varphi} + \boldsymbol{\varphi}'\mathbf{T}^{-1}\mathbf{J}\mathbf{T}\mathbf{G}^{-1}\boldsymbol{\varphi})^{-1} .$$

From $q > 0$ we get $\boldsymbol{\varphi}'\mathbf{G}^{-1}\boldsymbol{\varphi} \to \infty$ for $N \to \infty$.

Obviously, the row vector $\boldsymbol{\varphi}'\mathbf{G}^{-1}$ has all the elements bounded in absolute value by some positive constant, say Q, which does not depend on N. This implies $\boldsymbol{\varphi}'\mathbf{G}^{-1}\mathbf{J}\boldsymbol{\varphi}$ $\leqq 2QKnN$.

The existence of a positive constant C (C does not depend on N) such that $\boldsymbol{\varphi}'\mathbf{T}^{-1}\mathbf{J}\mathbf{T}\mathbf{G}^{-1}\boldsymbol{\varphi} \leqq CN$ may be proved.

The proof of this assertion follows from the fact that the row vector $\boldsymbol{\varphi}'\mathbf{T}^{-1}$ as well as the column vectors $\mathbf{G}^{-1}\boldsymbol{\varphi}$ and $\mathbf{T}\mathbf{G}^{-1}\boldsymbol{\varphi}$ have the elements bounded in absolute value by a constant not depending on N. We do not go in the details here.

With respect to this, the assertion of Theorem 7 immediately follows from (33).

References

[1] *U. Grenander, M. Rosenblatt:* Statistical analysis of stationary time series. New York, 1957.

[2] *J. Hájek:* On linear statistical problems in stochastic processes. Czech. Math. J. 12 (87), 1962, 404—444.

[3] *E. J. Hannan:* Анализ временных рядов. Москва 1964.

[4] *T. A. Magness, J. B. McGuire:* Comparison of least squares and minimum variance estimates of regression parameters. Ann. Math. Stat. 33, 1962, 462—470.

[5] *G. S. Watson:* Linear least squares regression. Ann. Math. Stat. 38, 1967, 1679—1699.

Souhrn

EFICIENCE ODHADŮ VE STACIONÁRNÍCH AUTOREGRESNÍCH POSLOUPNOSTECH

JIŘÍ ANDĚL

Předpokládejme, že $\mathbf{X} = (X_1, \ldots, X_N)'$ je náhodný vektor s regulární kovarianční maticí $\mathbf{G}$ takový, že

$$\mathrm{E}X_t = \alpha\varphi_t\,, \quad 1 \leqq t \leqq N\,,$$

kde $\varphi_t(1 \leqq t \leqq N)$ je daná funkce a α neznámý parametr. Budiž $\tilde{\alpha}$ nejlepší nestranný lineární odhad parametru α a $\hat{\alpha}$ budiž odhad α metodou nejmenších čtverců. Eficience odhadu $\hat{\alpha}$ je definována jako

$$e_N = \operatorname{var} \tilde{\alpha}/\operatorname{var} \hat{\alpha}\,.$$

Velmi důležité jsou případy, kdy náhodné veličiny $X_1, \ldots, X_N$ lze pokládat za část nekonečné posloupnosti $\{X_t\}_{-\infty}^{\infty}$, jejíž korelační funkce závisí jen na rozdílu argumentů. Pro nejdůležitější typy těchto posloupností a pro nejužívanější funkce φ_t je v tomto vypočtena eficience e_N. Zde uveďme stručný přehled výsledků.

Typ posloupnosti	φ_t	Věta	Vzorce pro e_N
autoregrese 1. řádu	1	1	(6)
autoregrese 2. řádu	1	2	(12), (13)
autoregrese n-tého řádu	1	3	(26)
autoregrese 1. řádu	t	4	
autoregrese 2. řádu	t	5	

V 7. odstavci článku je vypočtena eficience nejlepšího nestranného lineárního odhadu, který je založen na předpokladu, že $\{X_t\}_{-\infty}^{\infty}$ je autoregrese 1. řádu, ač ve skutečnosti jde o autoregresi 2. řádu. Výsledek je uveden ve větě 6. Nakonec ve větě 7 je dokázáno, že v případě autoregresní posloupnosti m-tého řádu s konstantní střední hodnotou nejlepší nestranný lineární odhad založený na předpokladu autoregrese n-tého řádu je asymptoticky eficientní.

Eficience e_N odhadu konstantní střední hodnoty metodou nejmenších čtverců v autoregresní posloupnosti 1. řádu je uvedena pro některé hodnoty N a některé hodnoty parametru a (který specifikuje korelační funkci) v tabulce 1.

Author's address: RNDr. *Jiří Anděl* CSc., Matematicko-fyzikální fakulta Karlovy university, Sokolovská 83, Praha 8 - Karlín.