KYBERNETIKA — VOLUME 39 (2003), NUMBER 1, PAGES 13 – 27

ON UNEQUALLY SPACED AR(1) PROCESS

JAN ŠINDELÁŘ AND JIŘÍ KNÍŽEK

Discrete autoregressive process of the first order is considered. The process is observed at unequally spaced time instants. Both least squares estimate and maximum likelihood estimate of the autocorrelation coefficient are analyzed. We show some dangers related with the estimates when the true value of the autocorrelation coefficient is small. Monte-Carlo method is used to illustrate the problems.

Keywords: AR(1) process, unequally spaced, autocorrelation coefficient, least squares estimate, maximum likelihood estimate

AMS Subject Classification: 60G10, 62M10

INTRODUCTION

The presented paper has been motivated by analysis of time series applied in biological research. Namely, densities of spots on two-dimensional gel electrophoresis maps were considered. Their time series were approximated via general linear regression model with disturbances forming an autoregressive process of the first order (AR(1) process). Observations were performed in fixed time instants according to some protocol. As a rule, the instants were unequally spaced.

Different estimation methods are discussed, for example, in [9, 10, 13]. Many aspects of parametric modelling approach may be found in [3, 14]. The article [15] provides an interesting overview of parametric modelling for growth curve data for both equally and unequally spaced cases. A wide range of bibliography can be found there as well. Various types of models dealing with unequally spaced case can be found in [4, 5, 6, 8], papers [2, 7, 12] are related to the AR(1) case.

The paper is devoted to estimation of parameters of discrete and unequally spaced AR(1) process. Both least squares and maximum likelihood estimates of the autocorrelation coefficient are considered. The estimates are obtained via optimization of the corresponding statistics denoted S_{LS} and S_{ML}. Analytical properties of the statistics rather than the statistical ones are stated. We analyze the dependence of the statistics on values of the autocorrelation coefficient in a neighbourhood of zero. We show that the statistics S_{LS} may reach minimum (resp. maximum) at zero and derive necessary and sufficient conditions characterizing both these situations. We show that the statistics S_{ML} may have minimum (maximum, inflex point respectively) at zero and derive necessary and sufficient conditions characterizing all three possibilities. Consequently, neither the least squares, nor the maximum likelihood estimator of the autocorrelation coefficient is the right one in the unequally spaced case. The stated properties of the estimates are illustrated by means of the Monte Carlo method.

The paper is self contained. We state and briefly prove the known results further applied in the paper. New results deal with the behaviour of the statistics S_{LS} and S_{ML} near zero.

The paper is organized as follows. Unequally spaced autoregressive process of the first order is outlined in Section 1. Least squares estimate of the autocorrelation coefficient is discussed in Section 2. Auxiliary results related to maximum likelihood estimation of the parameters of the process are stated in Section 3. Maximum likelihood estimates of the parameters are discussed in Section 4. Section 5 is devoted to Monte Carlo simulations.

1. THE AR(1) PROCESS

We consider a discrete AR(1) process

$$X_t = \rho X_{t-1} + v_t \qquad t = \ldots - 2, -1, 0, 1, 2, \ldots. \tag{1}$$

Hence the random variables v_t are independent and identically distributed, they have zero mean and a positive variance σ^2. The autocorrelation coefficient satisfies $|\rho| < 1$. We assume that the process is observed at time instants $t_0 < t_1 < \ldots < t_n$, $1 \le n$. Thus the random vector

$$\mathbf{Y} := (Y_0, \ldots, Y_n)$$

is obtained, where each Y_i denotes the member X_{t_i} of the process. Time increments

$$k_i := t_i - t_{i-1}$$

play an important role rather than the individual times due to stationarity of the process.

Majority of the results stated below holds almost surely (a. s.). This is not explicitly stated, except in cases where misunderstanding may occur.

2. LEAST SQUARES ESTIMATE OF AUTOCORRELATION COEFFICIENT

Least squares estimate of the autocorrelation coefficient is considered. The case in which all time increments are greater than one is the main theme of the section. It is shown that the estimate behaves badly if the true value of the autocorrelation coefficient is small. Necessary and sufficient conditions characterizing such situations are derived.

The following lemma leads to least squares estimate of the autocorrelation coefficient.

Lemma 2.1. The random variables

$$A_i := \frac{\left(Y_i - \rho^{k_i} Y_{i-1}\right) \cdot \sqrt{1-\rho^2}}{\sqrt{1-\rho^{2k_i}}}$$

are independent and identically distributed. They have zero mean and variance σ^2.

Proof. Consider the differences $Y_i - \rho^{k_i} Y_{i-1} = X_{t_i} - \rho^{k_i} X_{t_{i-1}}$. They equal $\rho^{k_i - 1} v_{t_{i-1}+1} + \ldots + \rho v_{t_i - 1} + v_{t_i}$ according to (1). Hence the differences have zero mean. Their variance equals $\sigma^2 \cdot \left(\rho^{2(k_i-1)} + \ldots + \rho^2 + 1\right) = \sigma^2 \cdot \frac{1-\rho^{2k_i}}{1-\rho^2}$. Thus the random variables A_i have zero mean and variance σ^2. Moreover, each A_i depends only on $v_{t_{i-1}+1}, \ldots, v_{t_i}$, hence $A_1, \ldots, A_n$ are mutually independent. □

An estimate of the autocorrelation coefficient may be found by means of (non-linear) least squares method. It minimizes the sum $\sum_{i=1}^{n} A_i^2$, i.e. it minimizes the statistics

$$S_{LS} := \sum_{i=1}^{n} \frac{\left(Y_i - \rho^{k_i} Y_{i-1}\right)^2 \cdot (1-\rho^2)}{1-\rho^{2k_i}} . \tag{2}$$

The same estimate is obtained by means of maximum likelihood method assuming that the random variables v_t are normally distributed and the value of the random variable Y_0 is known.

We introduce the following notation. We write $f \sim_x g$ if $f = g\theta$ holds in some neighbourhood of x, where θ is a continuous function with $\theta(x) = 1$. Hence if $f \sim_x g$, then the functions f and g are of the same sign and of the same magnitude in a neighbourhood of x. Clearly, the relation $\sim_x$ is an equivalence. For instance, we have $1 - x^{2k} \sim_1 k(1-x^2)$.

The statistics S_{LS} is a convex function of ρ when all time increments equal one. Its behaviour may be more complicated in unequally spaced case. To illustrate this fact we investigate behaviour of the statistics in a neighbourhood of zero.

Lemma 2.2. a) If $k_i > 1$, then we have $\frac{\partial A_i^2}{\partial \rho}(0) = 0$.

b) We have[1] $\frac{\partial S_{LS}}{\partial \rho}(0) = -2 \sum_{k_i=1} Y_i Y_{i-1}$.

c) If $k_i > 1$ for all i, then in some neighbourhood of zero we have

$$\frac{\partial S_{LS}}{\partial \rho} \sim_0 -2\rho \left[2 \sum_{k_i=2} Y_i Y_{i-1} + \sum_{i=1}^{n} Y_i^2 \right] .$$

Proof. a) Let us rewrite A_i^2 in the form

$$A_i^2 = \frac{\left(Y_i - \rho^{k_i} Y_{i-1}\right)^2}{1 + \rho^2 + \ldots \rho^{2(k_i-1)}} .$$

[1]The operator $\sum_{\substack{i=1 \\ k_i=k}}^{n}$ is abbreviated by $\sum_{k_i=k}$ for any $k \in \{1, \ldots, n\}$.

Assume that $k_i > 1$. Then we have

$$-\frac{\partial A_i^2}{\partial \rho} = \rho \cdot \left\{ \frac{2k_i \left[\rho^{k_i-2} Y_i Y_{i-1} - \rho^{2k_i-2} Y_{i-1}^2\right]}{1+\rho^2+\ldots+\rho^{2(k_i-1)}} \right. \\ \left. + \frac{\left[Y_i^2 - 2\rho^{k_i} Y_i Y_{i-1} + \rho^{2k_i} Y_{i-1}^2\right] \cdot 2 \cdot \left[1+2\rho^2+\ldots+(k_i-1)\rho^{2(k_i-1)-2}\right]}{\left[1+\rho^2+\ldots+\rho^{2(k_i-1)}\right]^2} \right\}. \tag{3}$$

This derivative evaluates to zero for $\rho = 0$. Lemma 2.2 a) is proved.

b) We have $\frac{\partial S_{LS}}{\partial \rho}(0) = \sum_{k_i=1} \frac{\partial A_i^2}{\partial \rho}(0)$ by part a) of the lemma. If $k_i = 1$, then we have $-\frac{\partial A_i^2}{\partial \rho} = 2(Y_i - \rho Y_{i-1})Y_{i-1}$, thus $-\frac{\partial A_i^2}{\partial \rho}(0) = 2Y_iY_{i-1}$. Summarizing these results we obtain $\frac{\partial S_{LS}}{\partial \rho}(0) = -\sum_{k_i=1} 2Y_iY_{i-1}$.

c) The expression in braces of (3) is denoted by $f_i(\rho)$. Assume that $k_i > 1$ for all i. If $k_i = 2$, then $f_i(0) = 4Y_iY_{i-1} + 2Y_i^2$ is valid, as follows from definition of the function f_i. If $k_i > 2$, then $f_i(0) = 2Y_i^2$. Hence in some neighbourhood of zero we have $-\frac{\partial S_{LS}}{\partial \rho} \sim_0 \rho \left[(\sum_{k_i=2} 4Y_iY_{i-1} + 2Y_i^2) + (\sum_{k_i>2} 2Y_i^2)\right]$, therefore part c. of the lemma is valid. □

The statistics S_{LS} reaches a local extreme at zero when all time increments are greater than one, as shown in

Theorem 2.1. Let $k = \min_{i=1,\ldots,n} k_i$. We denote $B := 2\sum_{k_i=2} Y_iY_{i-1} + \sum_{i=1}^n Y_i^2$.

a) Suppose that all v_t are normally distributed. If $k = 1$, then S_{LS} is strictly monotone at zero a. s.

b) Assume that $k > 1$. Properties of S_{LS} at zero are summarized in the following table:

$k = 2,\ B < 0$	$k = 2,\ B > 0$	$k > 2$
S_{LS} reaches a local minimum at zero	S_{LS} reaches a local maximum at zero	S_{LS} reaches a local maximum at zero.

Proof. a) Let us apply Lemma 2.2 b). If $k = 1$, then the sum $\sum_{k_i=1} Y_iY_{i-1} = -\frac{1}{2}\frac{\partial S_{LS}}{\partial \rho}(0)$ contains at least one summand. Suppose that all v_t are normally distributed. Then the sum differs from zero a. s. Thus the partial derivative $\frac{\partial S_{LS}}{\partial \rho}(0)$ differs from zero a. s. Hence the statistics S_{LS} is strictly monotone at zero a. s.

b) Firstly, let k equal two and B be negative. Then the sign of the derivative $\frac{\partial S_{LS}}{\partial \rho}$ equals the sign of ρ in some neighbourhood of zero, as follows from Lemma 2.2 c). Hence $S_{LS}(\rho)$ reaches a local minimum at zero. The sign of the derivative $\frac{\partial S_{LS}}{\partial \rho}$ equals the sign of $-\rho$ in the remaining two cases, hence $S_{LS}(\rho)$ reaches a local maximum at zero in both cases. □

Least squares estimate of the autocorrelation coefficient minimizes the statistics S_{LS}. But Theorem 2.1 shows that the statistics usually reaches a local maximum at zero when all time increments are greater than one. Consequently, the estimate behaves badly for small values of the autocorrelation coefficient. This statement is demonstrated by means of simulations in Section 5. The behaviour of the statistics S_{LS} is illustrated in Figure 1.

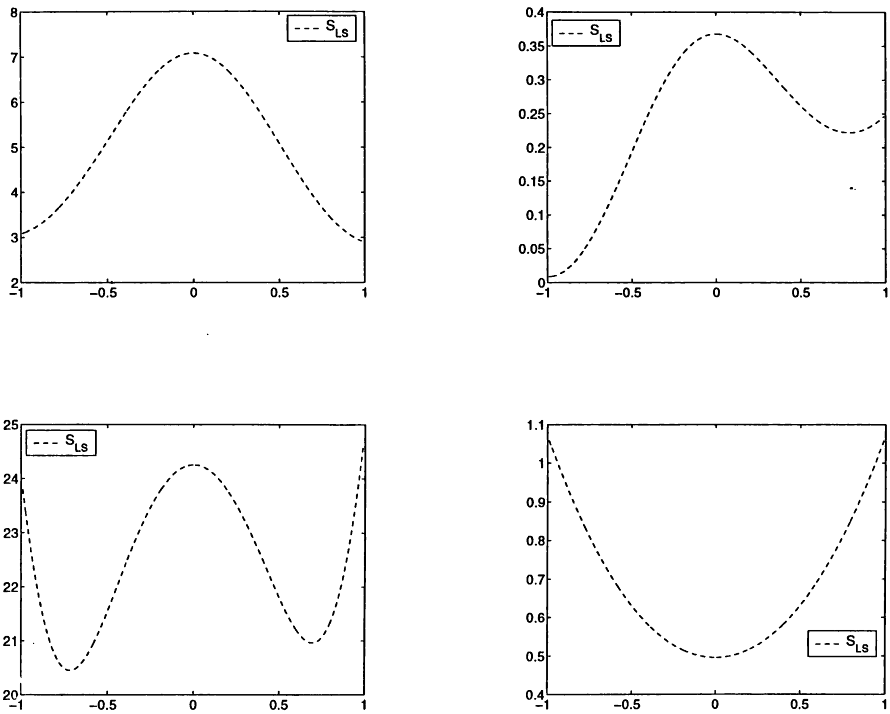

Fig. 1. Graphs of $S_{LS}(\rho)$ obtained from simulations with $k > 1$. Typical behaviour (local maximum at zero) is illustrated on the first three pictures.

3. ON COVARIANCE MATRIX OF AR(1) PROCESS

Auxiliary results related to maximum likelihood estimation of parameters of AR(1) process are stated in this section.

Let us state a form of the covariance matrix of the random vector $\mathbf{Y}$. Covariances of AR(1) process satisfy $\operatorname{cov}(X_t, X_{t+k}) = \frac{\sigma^2}{1-\rho^2} \cdot \rho^{|k|}$. It means that for $j > i$ we have $\operatorname{cov}(Y_i, Y_j) = \frac{\sigma^2}{1-\rho^2} \cdot \rho^{k_i+\ldots+k_{j-1}}$. Hence the covariance matrix of the random vector $\mathbf{Y}$ has the form

$$\boldsymbol{\Sigma} = \frac{\sigma^2}{1-\rho^2} \cdot \begin{pmatrix} 1 & \rho^{k_1} & \rho^{k_1+k_2} & \dots & & \rho^{k_1+\dots+k_n} \\ & 1 & \rho^{k_2} & & & \rho^{k_2+\dots+k_n} \\ & & 1 & & & \rho^{k_3+\dots+k_n} \\ & \text{symm.} & & \ddots & & \\ & & & & 1 & \rho^{k_n} \\ & & & & & 1 \end{pmatrix} . \tag{4}$$

Hereafter "symm." indicates that the matrix is symmetrical. Assumptions stated above guarantee that the matrix $\boldsymbol{\Sigma}$ is positive definite.

Let us evaluate the determinant of the covariance matrix and find its inverse.

Lemma 3.1. a) The determinant of the covariance matrix $\boldsymbol{\Sigma}$ is given by

$$|\boldsymbol{\Sigma}| = \frac{\sigma^{2(n+1)}}{(1-\rho^2)^{n+1}} \cdot \prod_{i=1}^{n} \left(1-\rho^{2k_i}\right) . \tag{5}$$

b) The inverse of the covariance matrix $\boldsymbol{\Sigma}$ equals

$$\boldsymbol{\Sigma}^{-1} = \frac{1-\rho^2}{\sigma^2} \cdot \begin{pmatrix} a_1 & b_1 & 0 & \dots & & 0 \\ & c_1 & b_2 & & & 0 \\ & & c_2 & & & \\ & \text{symm.} & & \ddots & & \\ & & & & c_{n-1} & b_n \\ & & & & & a_n \end{pmatrix} \tag{6}$$

with the elements a_i, b_i and c_i given by

$$a_i = \frac{1}{1-\rho^{2k_i}} \qquad b_i = \frac{-\rho^{k_i}}{1-\rho^{2k_i}} \qquad c_i = -1 + a_i + a_{i+1} \, . \tag{7}$$

P r o o f. a) The matrix expanded on the right hand side of (4) is denoted $\mathbf{Q}$. Let us take

$$[\, i\text{th row} \,] \; := \; [\, i\text{th row} \,] \; - \; \rho^{k_i} \cdot [\, (i+1)\text{th row} \,]$$

in the matrix. We obtain a lower triangular matrix with diagonal elements $1-\rho^{2k_1}, 1-\rho^{2k_2}, \dots, 1-\rho^{2k_n}$ and 1. Hence the determinant of the matrix $\mathbf{Q}$ equals $\prod_{i=1}^{n} \left(1-\rho^{2k_i}\right)$. Thus the determinant of $\boldsymbol{\Sigma}$ evaluates according to (5).

b) It can be confirmed directly that the inverse of the matrix $\boldsymbol{\Sigma}$ equals (6). □

Let us evaluate the quadratic form $\mathbf{y}'\boldsymbol{\Sigma}^{-1}\mathbf{y}$ and find some of its properties.

Lemma 3.2. Consider a vector $\mathbf{y} := (y_0, y_1, \ldots, y_n)'$ of reals. Let us denote

$$D(\mathbf{y}) := \frac{\sigma^2}{1-\rho^2} \cdot \mathbf{y}' \boldsymbol{\Sigma}^{-1} \mathbf{y} . \tag{8}$$

a) We have

$$D(\mathbf{y}) = y_0^2 + \sum_{i=1}^{n} \frac{\left(y_i - \rho^{k_i} y_{i-1}\right)^2}{1-\rho^{2k_i}} . \tag{9}$$

b) We have $D(\mathbf{y}) > 0$ if and only if $\mathbf{y}$ differs from zero vector.

P r o o f. a) The form (6) of the inverse of $\boldsymbol{\Sigma}$ stated in Lemma 3.1 gives us

$$D(\mathbf{y}) = a_1 y_0^2 + a_n y_n^2 + \sum_{i=1}^{n-1} c_i y_i^2 + 2 \sum_{i=1}^{n} b_i y_i y_{i-1} . \tag{10}$$

We have $c_i = -1 + a_i + a_{i+1}$ for all i, thus

$$\begin{aligned} D(\mathbf{y}) &= -\sum_{i=1}^{n-1} y_i^2 + \sum_{i=1}^{n} a_i \left(y_i^2 + y_{i-1}^2\right) + 2 \sum_{i=1}^{n} b_i y_i y_{i-1} && (11) \\ &= -\sum_{i=1}^{n-1} y_i^2 + \sum_{i=1}^{n} \left(\frac{y_i^2 + y_{i-1}^2}{1-\rho^{2k_i}} - \frac{2 y_i y_{i-1} \rho^{k_i}}{1-\rho^{2k_i}} \right) , && (12) \end{aligned}$$

where the coefficients a_i and b_i are given by (7). Finally, the right hand side of (9) equals (12).

b) The matrix $\boldsymbol{\Sigma}^{-1}$ is positive definite. Hence $D(\mathbf{y}) = \frac{\sigma^2}{1-\rho^2} \cdot \mathbf{y}' \boldsymbol{\Sigma}^{-1} \mathbf{y}$ is positive if and only if $\mathbf{y}$ is different from zero vector. □

The statistics S_{LS} leads to least squares estimate of the autocorrelation coefficient. Let us express the statistics by means of the statistics $D(\mathbf{Y})$.

Lemma 3.3. We have

$$S_{LS} = \left[D(\mathbf{Y}) - Y_0^2\right] \cdot (1-\rho^2) . \tag{13}$$

P r o o f. We have $\sum_{i=1}^{n} \frac{\left(Y_i - \rho^{k_i} Y_{i-1}\right)^2}{1-\rho^{2k_i}} = D(\mathbf{Y}) - Y_0^2$ by Lemma 3.2 a). Let us multiply both sides of the equation by $(1-\rho^2)$. The left hand side of the new equation equals the statistics S_{LS}, as follows from definition (2) of the statistics. □

4. MAXIMUM LIKELIHOOD ESTIMATION AND AR(1) PROCESS

Maximum likelihood estimates of parameters of AR(1) process are considered. The case in which all time increments are greater than one is the main theme of this section. We show that the maximum likelihood estimate of the autocorrelation coefficient may not be the appropriate one when the true value of the coefficient is small.

Let us start with likelihood function of the random vector $\mathbf{Y}$. We assume that the random variables v_t have normal distribution. Hence the random vector $\mathbf{Y}$ has normal distribution $N(\mathbf{0},\mathbf{\Sigma})$. It means that the likelihood function $\mathcal{L}=\mathcal{L}(\mathbf{Y}|\rho,\sigma)$ satisfies

$$\begin{aligned}\mathcal{L}^2 &= C_1\cdot|\mathbf{\Sigma}|^{-1}\cdot\exp\left(-\mathbf{Y}'\mathbf{\Sigma}^{-1}\mathbf{Y}\right)\\ &= C_1\cdot|\mathbf{\Sigma}|^{-1}\cdot\exp\left(-\frac{1-\rho^2}{\sigma^2}\cdot D(\mathbf{Y})\right), \qquad (14)\end{aligned}$$

where $C_1>0$ does not depend on parameters of the process.

Maximum likelihood estimates of the variance σ^2 and of the autocorrelation coefficient ρ may be obtained by means of the following proposition.

Proposition 4.1. a) The maximum likelihood estimate $\hat{\sigma}^2$ of variance is given by

$$\hat{\sigma^2}(\rho) = \frac{D(\mathbf{Y})\cdot(1-\rho^2)}{n+1}. \qquad (15)$$

b) Assume that $\sigma=\hat{\sigma}$. Then we have

$$\mathcal{L}^{-\frac{2}{n+1}}(\mathbf{Y}|\rho,\hat{\sigma}) = C_2\cdot D(\mathbf{Y})\cdot\left[\prod_{i=1}^{n}\left(1-\rho^{2k_i}\right)\right]^{\frac{1}{n+1}}, \qquad (16)$$

where $C_2>0$ does not depend on ρ.

P r o o f. a) The likelihood function satisfies

$$2\frac{\partial}{\partial\sigma^2}\ln\mathcal{L} = -\frac{\partial}{\partial\sigma^2}[\ln|\mathbf{\Sigma}|] - (1-\rho^2)\cdot D(\mathbf{Y})\cdot\frac{\partial}{\partial\sigma^2}\frac{1}{\sigma^2},$$

as follows from (14) and the fact that $D(\mathbf{Y})$ does not depend on σ by (9). Determinant of the covariance matrix $\mathbf{\Sigma}$ has been found in Lemma 3.1 a). We have $\ln|\mathbf{\Sigma}| = (n+1)\ln(\sigma^2)+C_3$, where C_3 does not depend on σ^2. Hence $\frac{\partial}{\partial\sigma^2}\ln|\mathbf{\Sigma}| = (n+1)\sigma^{-2}$. Summarizing these results we obtain

$$2\frac{\partial}{\partial\sigma^2}\ln\mathcal{L} = \sigma^{-4}\cdot\left[\,-(n+1)\cdot\sigma^2+D(\mathbf{Y})\cdot(1-\rho^2)\,\right].$$

Thus, for $D(\mathbf{Y})>0$ the likelihood function reaches its unique maximum at $\frac{D(\mathbf{Y})\cdot(1-\rho^2)}{n+1}$. Moreover, the random vector $\mathbf{Y}$ differs from zero vector a. s. Hence $D(\mathbf{Y})$ is positive a. s. by Lemma 3.2 b), which leads us to (15).

b) Let us take $\sigma = \hat{\sigma}$. The determinant $|\Sigma|$ has been evaluated in (5). Substituting $\sigma^2 = \hat{\sigma}^2$ we obtain

$$|\Sigma| = (n+1)^{-(n+1)} \cdot \left[D^{n+1}(\mathbf{Y}) \cdot \prod_{i=1}^{n} \left(1 - \rho^{2k_i}\right) \right] .$$

Moreover, part a) of the proposition gives us

$$\frac{1-\rho^2}{\hat{\sigma}^2} \cdot D(\mathbf{Y}) = n+1 .$$

Without loss of generality we can assume that $D(\mathbf{Y})$ is different from zero, since it is positive a.s. It results that the likelihood function satisfies $\mathcal{L}^2(\mathbf{Y}|\rho,\hat{\sigma}) = C_4 \left[D^{n+1}(\mathbf{Y}) \cdot \prod_{i=1}^{n} \left(1-\rho^{2k_i}\right) \right]^{-1}$, where $C_4 > 0$ does not depend on ρ. We take $-\frac{1}{n+1}$th power of both sides of the equation and obtain the desired result. □

Proposition 4.1 shows that the maximum likelihood estimate of the autocorrelation coefficient minimizes the statistics

$$S_{ML} := D(\mathbf{Y}) \cdot \left[\prod_{i=1}^{n} \left(1 - \rho^{2k_i}\right) \right]^{\frac{1}{n+1}} . \tag{17}$$

Maximum likelihood estimate of the autocorrelation coefficient is a solution of the likelihood equation $\frac{\partial}{\partial\rho} \ln \mathcal{L}(\mathbf{Y}|\rho,\hat{\sigma}) = 0$. Moreover, the left hand side of the equation satisfies $-\frac{2}{n+1} \cdot \frac{\partial}{\partial\rho} \ln \mathcal{L}(\mathbf{Y}|\rho,\hat{\sigma}) = \frac{\partial}{\partial\rho} \ln S_{ML}(\rho)$. Let us evaluate the last stated derivative. It enables us to state some properties of the statistics S_{ML}.

Lemma 4.1. a) We have

$$\frac{\partial}{\partial\rho} \ln S_{ML}(\rho) = \frac{1}{D(\mathbf{Y})} \cdot \frac{\partial}{\partial\rho} D(\mathbf{Y}) + \frac{1}{n+1} \cdot \sum_{i=1}^{n} \frac{-2k_i\rho^{2k_i-1}}{1-\rho^{2k_i}} .$$

b) Moreover, it holds

$$\frac{\partial}{\partial\rho} D(\mathbf{Y}) = \sum_{i=1}^{n} \frac{2k_i\rho^{k_i-1}}{\left(1-\rho^{2k_i}\right)^2} \cdot \left[\left(Y_i^2 + Y_{i-1}^2\right) \cdot \rho^{k_i} - Y_i Y_{i-1} \left(1+\rho^{2k_i}\right) \right] . \tag{18}$$

Proof. Validity of the first equation of the lemma follows from definition (17) of the statistics S_{ML}. The second equation follows from Lemma 3.2 a). □

Assume for a moment that all time increments equal one. Then the maximum likelihood estimate of the autocorrelation coefficient is a root of a cubic equation (see [1]). The root lies in the interval $(-1,1)$. Let us turn to case of arbitrary time increments. Now we show that the statistics S_{ML} reaches its minimum in the interval $(-1,1)$.

Proposition 4.2. We have $S_{ML} \sim_0 \sum_{i=0}^{n} Y_i^2$. Moreover, S_{ML} converges to infinity if ρ tends to any of the values 1^- and -1^+.

Proof. Recall that we have

$$S_{ML} = D(\mathbf{Y}) \cdot \left[\prod_{i=1}^{n} \left(1 - \rho^{2k_i}\right) \right]^{\frac{1}{n+1}} . \tag{19}$$

Assume that ρ equals zero. Then S_{ML} equals $D(\mathbf{Y})$. Moreover, $D(\mathbf{Y})$ equals $\sum_{i=0}^{n} Y_i^2$ by Lemma 3.2 a). Thus $S_{ML} \sim_0 \sum_{i=0}^{n} Y_i^2$.

Suppose that ρ is close to one of the values ± 1. Then we have

$$D(\mathbf{Y}) \quad \sim_{\pm 1} \quad Y_0^2 + \left(1 - \rho^2\right)^{-1} \cdot \sum_{i=1}^{n} \frac{\left(Y_i - \rho^{k_i} Y_{i-1}\right)^2}{k_i} ,$$

as follows from Lemma 3.2 a) and the fact that $1 - \rho^{2k_i} \sim_{\pm 1} k_i \cdot (1 - \rho^2)$ holds. Moreover, we have $\left[\prod_{i=1}^{n} \left(1 - \rho^{2k_i}\right) \right]^{\frac{1}{n+1}} \sim_{\pm 1} \left(\prod_{i=1}^{n} k_i\right)^{\frac{1}{n+1}} \cdot \left(1 - \rho^2\right)^{1 - \frac{1}{n+1}}$. Let us use these results and the form (19) of the statistics S_{ML}. We obtain

$$S_{ML} \sim_{\pm 1} \left(Y_0^2 (1 - \rho^2) + \sum_{i=1}^{n} \frac{\left(Y_i - \rho^{k_i} Y_{i-1}\right)^2}{k_i} \right) \cdot \left(\prod_{i-1}^{n} k_i \right)^{\frac{1}{n+1}} \cdot \left(1 - \rho^2\right)^{-\frac{1}{n+1}} .$$

The last multiplier $\left(1 - \rho^2\right)^{-\frac{1}{n+1}}$ converges to infinity if ρ tends to any of the values 1^- and -1^+. As a result, S_{ML} converges to infinity as well. □

We investigate behaviour of the statistics S_{ML} near zero.

Theorem 4.1. Let $k = \min_{i=1,\dots,n} k_i$ be minimum of all time increments.

a) If $k = 1$, then S_{ML} is strictly monotone at zero a. s.

b) Assume that $k > 1$. Properties of the statistics S_{ML} at zero are summarized in the following table:

k is odd	k is even	k is even
	$\sum_{k_i = k} Y_i Y_{i-1} < 0$	$\sum_{k_i = k} Y_i Y_{i-1} > 0$
S_{ML} has an inflex point at zero	S_{ML} reaches a local minimum at zero	S_{ML} reaches a local maximum at zero.

Proof. Let us approximate the derivative $\frac{\partial}{\partial \rho} \ln S_{ML}(\rho)$ in a small neighbourhood of zero. We have $D(\mathbf{Y}) \sim_0 \sum_{i=0}^{n} Y_i^2$, as follows from Lemma 3.2 a). Moreover, Lemma 4.1 b) gives $\frac{\partial}{\partial \rho} D(\mathbf{Y}) \sim_0 \sum_{i=1}^{n} 2 k_i \rho^{k_i - 1} \cdot [-Y_i Y_{i-1}]$. It results that

$\frac{\partial}{\partial \rho} D(\mathbf{Y}) \sim_0 \sum_{k_i=k} 2k_i \rho^{k_i-1} \cdot [-Y_i Y_{i-1}]$, because $\sum_{k_i=k} Y_i Y_{i-1} \neq 0$ holds a. s. Finally, we have $\sum_{i=1}^{n} \frac{-2k_i \rho^{2k_i-1}}{1-\rho^{2k_i}} \sim_0 \sum_{k_i=k} -2k_i \rho^{2k_i-1}$.

Using these approximations and Lemma 4.1 a) we obtain

$$\frac{\partial}{\partial \rho} \ln S_{ML}(\rho) \sim_0 \left(\sum_{i=0}^{n} Y_i^2 \right)^{-1} \cdot \left(-2k\rho^{k-1} \sum_{k_i=k} Y_i Y_{i-1} \right) + \frac{1}{n+1} \left(-2k\rho^{2k-1} \sum_{k_i=k} 1 \right) .$$

The second summand can be disregarded, as ρ^{2k-1} is substantially smaller than ρ^{k-1}. Thus

$$\frac{\partial}{\partial \rho} \ln S_{ML}(\rho) \sim_0 \; - \rho^{k-1} \cdot \left\{ 2k \cdot \left(\sum_{i=0}^{n} Y_i^2 \right)^{-1} \right\} \cdot \sum_{k_i=k} Y_i Y_{i-1} \, . \tag{20}$$

Assume that $k = 1$. Then ρ^{k-1} equals one. Hence the derivative $\frac{\partial}{\partial \rho} \ln S_{ML}(\rho)$ is positive (resp. negative) in some neighbourhood of zero a. s. Thus S_{ML} is strictly monotone in the neighbourhood.

Suppose that $k > 1$. Then the derivative $\frac{\partial}{\partial \rho} \ln S_{ML}(\rho)$ evaluates to zero for $\rho = 0$. Moreover, its sign equals the sign of $-\rho^{k-1} \cdot \sum_{k_i=k} Y_i Y_{i-1}$ in a neighbourhood of zero, as follows from (20). Parsing the possibilities "k is odd/even" and "$\sum_{k_i=k} Y_i Y_{i-1}$ is positive/negative" stated in Theorem 4.1 b) we find that the theorem is valid. □

Maximum likelihood estimate of the autocorrelation coefficient minimizes the statistics S_{ML}. Theorem 4.1 b) shows that the statistics may have an inflex point or it may reach a local maximum at zero when all time increments are greater than one. Consequently, the maximum likelihood estimate may not be the appropriate one when the true value of the autocorrelation coefficient is small. It should be noted that maximum likelihood estimators are not always efficient, or the best possible, as pointed out in [11]. Behaviour of the statistics S_{ML} and S_{LS} is illustrated in Figure 2.

Estimates of the autocorrelation coefficient stated in the paper minimize the following statistics:

maximum likelihood estimate minimizes	$D(\mathbf{Y})$	$\cdot \left[\prod_{i=1}^{n} (1-\rho^{2k_i}) \right]^{\frac{1}{n+1}}$,
least squares estimate minimizes	$(D(\mathbf{Y}) - Y_0^2)$	$\cdot [1-\rho^2]$.

The statistics seem to be alike. But the behaviour of the estimates is different, as suggested by Theorems 2.1 and 4.1. This claim is supported by simulations performed in the next section.

5. MONTE CARLO SIMULATIONS

We deal with small sample properties of the estimators stated in previous sections. The length of the series considered in simulations is 100, 10 000 repetitions are performed. Sample means (SM) and sample standard deviations (SSD) of the estimates are considered. Variance σ^2 equals one in all the experiments.

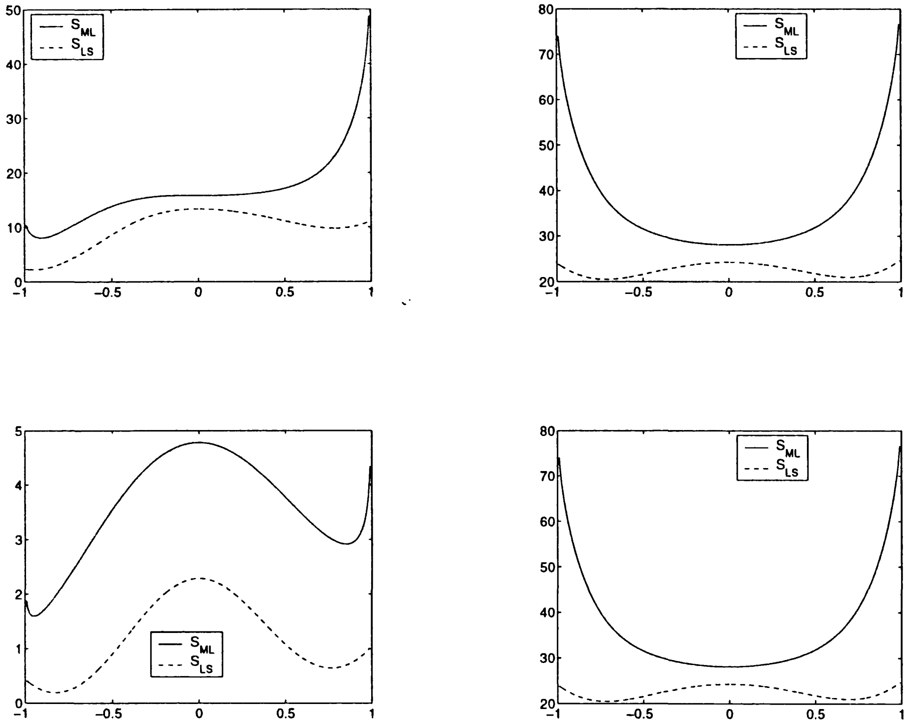

Fig. 2. Graphs of $S_{ML}(\rho)$ and $S_{LS}(\rho)$ obtained from simulations with $k > 1$. The statistics $S_{ML}(\rho)$ may have an inflex point, local minimum or maximum at zero.

Two types of simulations are stated. Firstly, time increments equal either two or three, about 50 % for each. The value of the autocorrelation coefficient varies from -0.9 to 0.9. Secondly, fixed value -0.5 of the autocorrelation coefficient is taken. We use the same type of time increments, but replace some of them by ones. From 10 % to 90 % of the increments are replaced. For instance, if 20 % of them are replaced, then about 40 % of the increments equal two and the remaining approximately 40 % equal three.

Our experiments show that the absolute value of the autocorrelation coefficient may be estimated correctly, but the sign of the estimate may be wrong. Because of this, absolute values of the estimates are also stated. Subscript $_a$ is used to mark the case of absolute values.

We start with the case where all time increments are greater than one and the value of the autocorrelation coefficient varies.

Table 1 shows the properties of maximum likelihood estimator of the autocorrelation coefficient. The estimator does not behave well for small values of ρ, which is in accordance with Theorem 4.1. Reasonable absolute values of the estimates are obtained if the value of ρ equals −0.5, −0.3, 0.3 and 0.5.

Table 1. Properties of maximum likelihood estimator of the autocorrelation coefficient.

ρ	−0.9	−0.7	−0.5	−0.3	−0.1	0.1	0.3	0.5	0.7	0.9
SM	−0.889	−0.680	−0.315	−0.052	−0.005	−0.003	0.048	0.312	0.681	0.889
SSD	0.033	0.132	0.390	0.351	0.284	0.284	0.352	0.394	0.124	0.033
SM_a	0.889	0.689	0.483	0.296	0.204	0.203	0.297	0.484	0.689	0.889
SSD_a	0.033	0.066	0.135	0.196	0.198	0.199	0.196	0.134	0.066	0.033

Table 2 shows the properties of maximum likelihood estimator of variance. It suggests that, even in case where the sign of the autocorrelation coefficient has been wrongly estimated, reasonable estimates of the variance can still be obtained.

Table 2. Properties of least squares estimator of the autocorrelation coefficient.

ρ	−0.9	−0.7	−0.5	−0.3	−0.1	0.1	0.3	0.5	0.7	0.9
SM	0.998	0.998	0.992	0.960	0.930	0.929	0.959	0.990	0.998	0.999
SSD	0.149	0.163	0.177	0.173	0.159	0.161	0.171	0.176	0.164	0.148

Table 3 shows the properties of least squares estimator of the autocorrelation coefficient. A brief look at the last two rows dealing with absolute values of the estimates shows that the estimator is a bad one. This is also suggested by Theorem 2.1.

Table 3. Properties of least squares estimator of the autocorrelation coefficient.

ρ	−0.9	−0.7	−0.5	−0.3	−0.1	0.1	0.3	0.5	0.7	0.9
SM	−0.958	−0.870	−0.516	−0.124	−0.009	−0.004	0.121	0.508	0.872	0.959
SSD	0.013	0.157	0.644	0.783	0.779	0.779	0.784	0.651	0.146	0.013
SM_a	0.958	0.884	0.825	0.792	0.778	0.778	0.792	0.825	0.884	0.959
SSD_a	0.013	0.023	0.027	0.030	0.032	0.033	0.030	0.027	0.022	0.013

Let us turn over to fixed value -0.5 of the autocorrelation coefficient. The number of time increments equal to one varies from 10 % to 90 %.

Table 4 shows the properties of maximum likelihood estimator of the autocorrelation coefficient. It indicates that the behaviour of the estimator is improved, even in cases when only a small number of time increments is equal to one.

Table 4. Properties of maximum likelihood estimator of the autocorrelation coefficient (the first row – percentage of time increments equal one).

	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %
SM	−0.428	−0.472	−0.486	−0.488	−0.487	−0.489	−0.490	−0.490	−0.489
SSD	0.283	0.161	0.112	0.099	0.095	0.091	0.090	0.090	0.087
SM_a	0.481	0.480	0.487	0.488	0.487	0.489	0.490	0.490	0.489
SSD_a	0.178	0.137	0.110	0.099	0.095	0.091	0.090	0.090	0.087

We have found in our simulations that there is an improvement in the behaviour of least squares estimator of the autocorrelation coefficient in cases where more than 60 % of time increments equal one.

ACKNOWLEDGEMENT

This research has been partially supported by the Ministry of Health of the Czech Republic through the project 6458–3 of and by the grant project K1019101 of the Grant Agency of the Academy of Sciences of the Czech Republic.

(Received November 1, 2002.)

REFERENCES

[1] T. W. Anderson: The Statistical Analysis of Time Series. Wiley, New York 1971.

[2] B. H. Baltagi and P. X. Wu: Unequally spaced panel data regressions with AR(1) disturbances. Econometric Theory *15* (1999), 814–823.

[3] P. J. Diggle, K. Y. Liang, and S. L. Zeger: Analysis of Longitudinal Data. Oxford University Press, New York 1994.

[4] M. A. Hauser: Maximum likelihood estimators for ARMA and ARFIMA models: a Monte Carlo study. J. Statist. Plann. Inference *80* (1999), 229–255.

[5] R. H. Jones: Longitudinal Data with Serial Correlation: A State-space Approach. Chapman & Hall, New York 1993.

[6] R. H. Jones and L. M. Ackerson: Serial correlation in unequally spaced longitudinal data. Biometrika *77* (1990), 721–731. 1

[7] R. H. Jones and F. Boadi-Boateng: Unequally spaced longitudinal data with AR(1) serial correlation. Biometrics *47* (1991), 161–175.

[8] R. H. Jones and A. V. Vecchia: Fitting continuous ARMA models to unequally spaced spatial data. J. Amer. Statist. Assoc. *88* (1993), 947–954.
[9] S. M. Kay: Fundamendals of Statistical Signal Prosessing: Estimation Theory. Prentice Hall, Englewood Cliffs, NJ 1993.
[10] D. Kazakos and P. Papantoni-Kazakos: Detection and Estimation. Computer Science Press, New York 1990.
[11] L. Le Cam: Maximum likelihood: An introduction. Internat. Statist. Rev. *58* (1990), 2, 153–171.
[12] D. J. Mckenzie: Estimation of AR(1) models with unequally spaced pseudo-panels. Econometrics J. *4* (2001), 1, 89–108.
[13] H. W. Sorenson: Parameter Estimation: Principles and Problems. Marcel Dekker, New York 1980.
[14] G. Verbeke and G. Molenberghs: Linear Mixed Models for Longitudinal Data. Springer Verlag, New York 2000.
[15] D. L. Zimmerman and V. Núñez-Antón: Parametric modelling of growth curve data: An overview. Test, Sociedad de Estadística e Investigación Operativa *10* (2001), 1, 1–73.

RNDr. Jan Šindelář, CSc. and Ing. Jiří Knížek, CSc., Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mails: sindelar@utia.cas.cz, knizek@utia.cas.cz