Acta Math. Univ. Comenianae
Vol. LXIII, 1(1994), pp. 107–116

ON ESTIMATION OF A COVARIANCE FUNCTION OF STATIONARY ERRORS IN A NONLINEAR REGRESSION MODEL

F. ŠTULAJTER

Abstract. A nonlinear regression model with correlated, normally distributed stationary errors is investigated. Limit properties of an approximate estimator of an unknown covariance function of stationary errors are studied and sufficient conditions under which this estimator is consistent are shown.

1. Introduction

The theory of estimation in a nonlinear regression model has been extensively studied by many authors (see Jennrich (1969), Rattkowsky (1983), Gallant (1987) and others). The main effort was devoted to the study of problems of estimation of unknown regression parameters by least squares method under the assumption that errors are independent and identically distributed with some unknown variance. Under these assumptions the limit properties of an approximate least squares estimator of regression parameters and variance were derived. In this connection the classical results are given by Jennrich (1969), Box (1971), Clarke (1980), Pázman (1984), Wu (1981) and others. The case of correlated errors was studied by Hannan (1971), Gallant and Goebel (1976), Gallant (1987) and Štulajter (1992) and was devoted mainly to problems of estimation of regression parameters and their limit properties. Cook and Tsai (1985) studied properties of residuals in a nonlinear regression model with uncorrelated errors.

The aim of this article is to study the problem of estimation of parameters of random errors which are assumed to be a finite part of a stationary gaussian random process with an unknown covariance function which should be estimated.

Let us consider a random process y following a nonlinear regression model

$$y_t = f(x_t, \theta) + \varepsilon_t; \quad t = 1, 2, \dots \tag{1}$$

where f is a model function, x_t; $t = 1, 2, \dots$ are assumed to be known k-dimensional vectors, $\theta = (\theta_1, \dots, \theta_p)'$ is an unknown vector of regression parameters which

Received November 16, 1992.
1980 *Mathematics Subject Classification* (1991 *Revision*). Primary 62J02.
Key words and phrases. Nonlinear regression, least squares estimator, residuals, correlated errors, covariance function estimation.

belongs to some open set Θ. Further we'll assume that the vector $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ has $N(0, \Sigma)$ distribution, where $\Sigma_{ij} = R(|i-j|)$; $i, j = 1, 2, \dots, n$ and $R(\cdot)$ is a covariance function of a stationary stochastic process $\varepsilon = \{\varepsilon(t);\ t = 1, 2, \dots\}$. This covariance function should be estimated, using the vector $y = (y_1, \dots, y_n)'$ of observations following the model (1). We'll derive an approximate consistent estimator of this covariance function. This estimator can be used e.g. in a kriging method of prediction of a stochastic process (see Stein (1988)).

2. An Approximate Least Squares Estimator and Approximate Residuals

The problem of estimation of the covariance function $R(\cdot)$ will be solved using a stochastic approximation for the least squares estimator $\hat{\theta}$ given by

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \sum_{t=1}^{n} [y_t - f(x_t; \theta)]^2 .$$

Using the idea of Box (1971) it was shown in Štulajter (1992) that the estimator $\hat{\theta}$ can be approximated by the estimator $\tilde{\theta}$ given by

$$\tilde{\theta} = \theta + A\varepsilon \ + (J'J)^{-1} \left[(\varepsilon' N \varepsilon) - \frac{1}{2} J'(\varepsilon' A' H A \varepsilon) \right] . \tag{2}$$

Here J is the $n \times p$ matrix of derivatives of f with $J_{ij} = \frac{\partial f(x_i;\theta)}{\partial \theta_j}$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, p$, $A = (J'J)^{-1}J'$ is the $p \times n$ matrix, $(\varepsilon' A' H A \varepsilon)$ denotes the $n \times 1$ random vector with components $\varepsilon' A' H_j A \varepsilon$; $j = 1, 2, \dots, n$, where the $p \times p$ matrix H_j is given by

$$(H_j)_{kl} = \frac{\partial^2 f(x_j; \theta)}{\partial \theta_k \partial \theta_l}; \quad k, l = 1, 2, \dots, p; \quad j = 1, 2, \dots, n .$$

All the derivatives are assumed to be continuous and are computed at the true value of the parameter θ. Next, $(\varepsilon' N \varepsilon)$ denotes the $p \times 1$ random vector with components $\varepsilon' N_j \varepsilon$; $j = 1, 2, \dots, p$, where the $n \times n$ matrix N_j is given by

$$(N_j)_{kl} = \sum_{i=1}^{n} (H_i A)_{jk} M_{il} .$$

Here M is a projection matrix, $M = I - J(J'J)^{-1}J'$.

Using a part of the Taylor expansion of f at the point $\tilde{\theta}$ we get

$$f(\tilde{\theta}) = f(\theta) + J(\tilde{\theta} - \theta) + \frac{1}{2} (\tilde{\theta} - \theta)' H (\tilde{\theta} - \theta) \tag{3}$$

where $f(\theta) = ((f(x_1;\theta), \ldots, f(x_n;\theta))'$ and $(\tilde{\theta}-\theta)'H(\tilde{\theta}-\theta)$ denotes the $n \times 1$ random vector with components $(\tilde{\theta}-\theta)'H_j(\tilde{\theta}-\theta)$; $j = 1, 2, \ldots, n$.

For the residuals $\hat{e} = y - f(\tilde{\theta})$ we get, using (2) and (3), the expression

$$\begin{aligned}\hat{e} = y - f(\tilde{\theta}) = \varepsilon - J\left\{A\varepsilon\ + (J'J)^{-1}\left[(\varepsilon' N\varepsilon) - \frac{1}{2}J'(\varepsilon' A' H A\varepsilon)\right]\right\}\\ -\frac{1}{2}\left(A\varepsilon + (J'J)^{-1}\left[(\varepsilon' N\varepsilon) - \frac{1}{2}J'(\varepsilon' A' H A\varepsilon)\right]\right)' H\\ \left(A\varepsilon + (J'J)^{-1}\left[(\varepsilon' N\varepsilon) - \frac{1}{2}J'(\varepsilon' A' H A\varepsilon)\right]\right).\end{aligned}$$

Using only the linear and quadratic (in components of ε) terms we can approximate the residuals $\hat{\varepsilon}$ by $\tilde{\varepsilon}$ given by

$$\tilde{e} = M\varepsilon - A'(\varepsilon' N\varepsilon) - \frac{1}{2}M(\varepsilon' A' H A\varepsilon). \tag{4}$$

These residuals will be used for estimation the unknown covariance function $R(\cdot)$. Some properties of residuals for the case of uncorrelated errors were studied by Cook and Tsai (1985).

3. Estimation of a Covariance Function

As we have told in the introduction, we'll assume that the vector ε has the $N_n(0,\Sigma)$ distribution with $\Sigma_{ij} = R(|i-j|)$. Now, let us consider the random matrix $\tilde{\Sigma}$ given by

$$\begin{aligned}\tilde{\Sigma} = \tilde{e}\tilde{e}' = M\varepsilon\varepsilon' M - A'(\varepsilon' N\varepsilon)\varepsilon' M - M\varepsilon(\varepsilon' N\varepsilon)' A + A'(\varepsilon' N\varepsilon)(\varepsilon' N\varepsilon)' A\\ -\frac{1}{2}M\varepsilon(\varepsilon' A' H A\varepsilon)' M - \frac{1}{2}M(\varepsilon' A' H A\varepsilon)\varepsilon' M\\ +\frac{1}{2}A'(\varepsilon' N\varepsilon)(\varepsilon' A' H A\varepsilon)' M + \frac{1}{2}M(\varepsilon' A' H A\varepsilon)(\varepsilon' N\varepsilon)' A\\ +\frac{1}{4}M(\varepsilon' A' H A\varepsilon)(\varepsilon' A' H A\varepsilon)' M\,.\end{aligned} \tag{5}$$

The estimators $\hat{R}(t)$ of $R(t)$; $t = 0, 1, \ldots, n-1$ given by

$$\hat{R}(t) = \frac{1}{n-t}\sum_{s=1}^{n-t}\left(y_{s+t} - f(x_{s+t};\hat{\theta})\right)\left(y_s - f(x_s;\hat{\theta})\right) \tag{5a}$$

are the natural generalizations of the estimators of $R(\cdot)$ for the case when the mean value follows a linear regression model which were studied in Štulajter (1989) and (1991).

The estimators $\hat{R}(t)$ can be approximated by the estimators

$$\tilde{R}(t) = \frac{1}{n-t} \sum_{s=1}^{n-t} \tilde{e}_{t+s} \tilde{e}_s; \quad t = 0, 1, \dots, n-1$$

with $\tilde{e}$ given by (4), which can be written in the form (they depend on n)

$$\tilde{R}_n(t) = \frac{1}{n-t} \operatorname{tr}(B_t \tilde{\Sigma}); \quad t = 0, 1, \dots, n-1 \,. \tag{6}$$

Here B_t; $t = 0, 1, \dots, n-1$ are the block $n \times n$ matrices,

$$B_t = \frac{1}{2} \left[\begin{pmatrix} 0 & I_t \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ I_t & 0 \end{pmatrix} \right]$$

with I_t being the $(n-t) \times (n-t)$ identity matrix and tr denotes the trace of a matrix.

In the sequal we shall need the notion of the Euclidean inner product defined for any $n \times n$ matrices A and B by $(A, B) = \sum_{i,j=1}^{n} A_{ij} B_{ij}$ which can be written as $(A, B) = \operatorname{tr}(AB')$.
Thus the Schwarz inequality can be written as

$$|\operatorname{tr}(AB')| \leq \|A\| \, \|B\|, \quad \text{where } \|A\| = (A, A)^{1/2}.$$

It is easy to prove that $\|AB\| \leq \|A\| \, \|B\|$ for any matrices A and B and $\|AB_t\| \leq \|A\|$; $t = 0, 1, \dots, n-1$ for any matrix A and the matrices B_t defined in (6) (for the proof of the last inequality we refer to Štulajter (1991)). Next, using the equality $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, which holds for any matrices for which the products AB and BA are defined and are square matrices, we can write:

$$\begin{aligned} \operatorname{tr}(B_t \tilde{\Sigma}) = {} & \varepsilon' M B_t M \varepsilon - 2\varepsilon' M B_t A' (\varepsilon' N \varepsilon) - (\varepsilon' N \varepsilon)' A B_t A' (\varepsilon' N \varepsilon) \\ & - (\varepsilon' A' H A \varepsilon)' M B_t M \varepsilon \; + (\varepsilon' A' H A \varepsilon)' M B_t A' (\varepsilon' N \varepsilon) \\ & + \frac{1}{4} (\varepsilon' A' H A \varepsilon)' M B_t M (\varepsilon' A' H A \varepsilon). \end{aligned} \tag{7}$$

Now we shall study limit properties, as n tends to infinity, of the estimators $\tilde{R}_n(t)$; $t = 0, 1, \dots, n-1$ given by (6). The matrices Σ, M, H and others and also their norms depend on n but this will not be announced later on.

Theorem. *Let in the nonlinear regression model* (1)

$$(J_n' J_n)^{-1} = \frac{1}{n} G_n \,, \tag{8}$$

where $\lim_{n\to\infty} G_n = G$ *and* G *is a nonnegative definit matrix. Next, let the following limits*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\partial f(x_t;\theta)}{\partial \theta_i} \frac{\partial^2 f(x_t;\theta)}{\partial \theta_j \partial \theta_k} \tag{9}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 f(x_t;\theta)}{\partial \theta_i \partial \theta_j} \frac{\partial^2 f(x_t;\theta)}{\partial \theta_k \partial \theta_l} \tag{10}$$

exist and are finite for every $i,\ j,\ k,\ l$. *Let the errors* ε *have* $N_n(0,\Sigma)$ *distribution with* $\Sigma_{ij} = R(|i-j|);\ i,j = 1,2,\dots,n$ *and let*

$$\lim_{n\to\infty} \frac{1}{n} \|\Sigma\| = 0\,. \tag{11}$$

Then the estimators $\tilde{R}_n(t)$ *given by* (6) *converges for every fixed* t *in probability to* $R(t)$ *as* n *tends to infinity.*

Proof. It was shown in Štulajter (1991) that

$$\lim_{n\to\infty} E\left[\frac{1}{n}\varepsilon' M B_t M \varepsilon - R(t)\right]^2 = 0 \ \text{ if only } \quad \lim_{n\to\infty} \frac{1}{n}\|\Sigma\| = 0$$

and thus $\frac{1}{n}\varepsilon' M B_t M\varepsilon$ converges in probability to $R(t)$ for every t as n tends to infinity. Thus the theorem will be proved if we show that all the members appearing in (7) and multiplied by $\frac{1}{n}$ converge in probability to zero. Let us consider the term $\frac{1}{n}\varepsilon' M B_t A'(\varepsilon' N\varepsilon)$. We can write:

$$\left|\frac{1}{n}\varepsilon' M B_t A'(\varepsilon' N \varepsilon)\right|^2 \le \frac{1}{n^2}\|M\varepsilon\|^2 \|B_t A'(\varepsilon' N\varepsilon)\|^2 \le \frac{1}{n}\varepsilon' M\varepsilon \frac{1}{n}\|A'(\varepsilon' N\varepsilon)\|^2.$$

Now we shall prove that $\frac{1}{n}\varepsilon' M\varepsilon \xrightarrow{P} R(0)$ (converges in probability to $R(0)$) and $\frac{1}{n}\|A'(\varepsilon' N\varepsilon)\|^2 \xrightarrow{P} 0$ and thus their product converges in probability to zero. But $\frac{1}{n}\varepsilon' M\varepsilon = \frac{1}{n}\varepsilon' M B_0 M\varepsilon$ (since $B_0 = I$ and $M^2 = M$) and it was already shown that this term converges to $R(0)$. Next we have:

$$(A'(\varepsilon' N\varepsilon))_i = \sum_{t=1}^{n} (A' H_t A\varepsilon)_i (M\varepsilon)_t \quad \text{for } i = 1,2,\dots,n \qquad \text{and}$$

$$\begin{aligned}\frac{1}{n}\|A'(\varepsilon' N\varepsilon)\|^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{t=1}^{n}(A' H_t A\varepsilon)_i (M\varepsilon)_t\right)^2 \\ &\le \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{n}(A' H_t A\varepsilon)_i^2 \sum_{t=1}^{n}(M\varepsilon)_t^2 = \sum_{t=1}^{n}\varepsilon' A' H_t A A' H_t A\varepsilon \frac{1}{n}\varepsilon' M\varepsilon\,.\end{aligned}$$

Since $\frac{1}{n}\varepsilon' M\varepsilon \xrightarrow{P} R(0)$, it remains to prove that $\sum_{t=1}^n \varepsilon' A' H_t A A' H_t A\varepsilon \xrightarrow{P} 0$. But $E\left[\sum_{t=1}^n \varepsilon' A' H_t A A' H_t A\varepsilon\right] = \sum_{t=1}^n \operatorname{tr}(A' H_t A A' H_t A\Sigma)$, since $E[\varepsilon' C\varepsilon] = \operatorname{tr}(C\Sigma)$ for any symmetric matrix C and thus

$$\left| E\left[\sum_{t=1}^n \varepsilon' A' H_t A A' H_t A\varepsilon\right]\right| \le \sum_{t=1}^n |\operatorname{tr}(A' H_t A A' H_t A\Sigma)| \le \sum_{t=1}^n \|A' H_t A\|^2 \|\Sigma\| .$$

Next,

$$\sum_{t=1}^n \|A' H_t A\|^2 = \sum_{t=1}^n \operatorname{tr}(A' H_t A A' H_t A) = \operatorname{tr}\left(\sum_{t=1}^n H_t (J'J)^{-1} H_t (J'J)^{-1}\right),$$

since $AA' = (J'J)^{-1}$. Thus we have:

$$\lim_{n\to\infty} E\left[\sum_{t=1}^n \varepsilon' A' H_t A A' H_t A\varepsilon\right] = 0$$

if the assumptions (8), (10) and (11) of the theorem are fulfilled. From the same reasons, using the expression $\operatorname{Var}(\varepsilon' C\varepsilon) = 2\operatorname{tr}(C\Sigma C\Sigma)$, which holds for any symmetric matrix C and any random vector ε having $N_n(0,\Sigma)$ distribution, we get:

$$\lim_{n\to\infty} \operatorname{Var}\left[\sum_{t=1}^n \varepsilon' A' H_t A A' H_t A\varepsilon\right] = 0 .$$

Next, $|\frac{1}{n}(\varepsilon' N\varepsilon)' A B_t A'(\varepsilon' Ne)| \le \frac{1}{n}\|A'(\varepsilon' N\varepsilon)\|^2 \xrightarrow{P} 0$ as we have just shown. Further, denote $P = J(J'J)^{-1}J' = I - M$. Then

$$\begin{aligned} \left|\frac{1}{n}(\varepsilon' A' HA\varepsilon)' MB_t M\varepsilon\right| &= \frac{1}{n}|(\varepsilon' A' HA\varepsilon)(I-P)B_t M\varepsilon| \\ &\le \frac{1}{n}\left(|\varepsilon' A' HA\varepsilon)' B_t M\varepsilon| + |(\varepsilon' A' HA\varepsilon) PB_t M\varepsilon|\right) \quad \text{and} \\ \left|\frac{1}{n}(\varepsilon' A' HA\varepsilon)' B_t M\varepsilon\right| &\le \frac{1}{n}\|(\varepsilon' A' HA\varepsilon\|^2 \frac{1}{n}\varepsilon' M\varepsilon \\ &\le \varepsilon' A' A\varepsilon \frac{1}{n}\sum_{t=1}^n \varepsilon' A' H_t^2 A\varepsilon \frac{1}{n}\varepsilon' M\varepsilon . \end{aligned} \tag{12}$$

It is easy to prove that the mean values and variances of $\varepsilon' A' A\varepsilon$ and $\varepsilon' A' \frac{1}{n}\sum_{t=1}^n H_t^2 A\varepsilon$ converge to zero under the assumptions of the theorem and thus these random variables converge to zero in probability. For the second term of the right hand side of the inequality (12) we have:

$$\begin{aligned}\left|\frac{1}{n}(\varepsilon' A' H A\varepsilon)' P B_t M\varepsilon\right|^2 &\leq \frac{1}{n^2}\|(\varepsilon' A' H A\varepsilon)\|^2 \|P B_t M\varepsilon\|^2 \\ &\leq \frac{1}{n^2}\|(\varepsilon' A' H A\varepsilon)\|^2 \|P\|^2 \|B_t M\varepsilon\|^2 \\ &\leq \frac{p}{n}\|(\varepsilon' A' H A\varepsilon)\|^2 \frac{1}{n}\varepsilon' M\varepsilon,\end{aligned}$$

since $P^2 = P$, $P = P'$ and thus $\|P\|^2 = \operatorname{tr}(P) = \operatorname{rank}(P) = p$.

Let us consider the last two terms of (7). We get, as before:

$$\begin{aligned}&|(\varepsilon' A' H A\varepsilon)' M B_t A'(\varepsilon' N\varepsilon)| \\ &\qquad \leq |(\varepsilon' A' H A\varepsilon) B_t A'(\varepsilon' N\varepsilon)| + +|(\varepsilon' A' H A\varepsilon)' P B_t A'(\varepsilon' N\varepsilon)|\,.\end{aligned}$$

Next,

$$\frac{1}{n^2}|(\varepsilon' A' H A\varepsilon)' B_t A'(\varepsilon' N\varepsilon)|^2 \leq \frac{1}{n}\|(\varepsilon' A' H A\varepsilon)\|^2 \frac{1}{n}\|A'(\varepsilon' N\varepsilon)\|^2$$

and we know from our proof that both terms on the right hand side of the last inequality converge to zero. Finally,

$$\frac{1}{n^2}|(\varepsilon' A' H A\varepsilon)' P B_t A'(\varepsilon' N\varepsilon)|^2 \leq \frac{p}{n}\|(\varepsilon' A' H A\varepsilon)\| \frac{1}{n}\|A'(\varepsilon' N\varepsilon)\|^2.$$

The last term of (7) can be bounded by the same way.

The proof of the theorem now follows from the derived results and from the well known facts on convergence in probability:
a) $X_n \to X$ iff $X_n^2 \to X^2$
b) if $X_n \to X$ and $Y_n \to Y$, then $X_n Y_n \to XY$ and $aX_n + bY_n \to aX + bY$
c) if $|X_n| \leq |Y_n|$ and $Y_n \to 0$, then $X_n \to 0$ and
d) if $E[X_n] \to 0$ and $\operatorname{Var}[X_n] \to 0$, then $X_n \to 0$. □

Remarks. 1. The conditions (8), (9) and (10) are similar to those appearing in Jennrich (1969), Wu (1981) and others studying the limit properties of the least squares estimator $\hat{\theta}$ of θ. It was shown in Štulajter (1991) that for consistency of estimators of a covariance function weaker conditions than for consistency of regression parameters are requiared if the regression model is linear. A similar situation occours in the case of nonlinear regression.

2. For estimating $R(0)$ we have $B_0 = I$. Two terms from (7) vanish in this case, since $MA' = 0$.

3. If the errors are uncorrelated with a common variance σ^2 then $\frac{1}{n}\|\Sigma\| = \frac{1}{n^{1/2}}\sigma^2$ and the condition (11) of the theorem is fulfilled.

4. For stationary errors we have:

$$\|\Sigma\| = \left(nR^2(0) + 2\sum_{t=1}^{n}(n-t)R^2(t)\right)^{1/2} \quad \text{and}$$
$$\frac{1}{n}\|\Sigma\| \leq \left(\frac{1}{n}R^2(0) + \frac{2}{n}\sum_{t=1}^{n}R^2(t)\right)^{1/2}.$$

It is easy to prove that $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} R^2(t) = 0$ if $\lim_{t\to\infty} R(t) = 0$. Thus the condition (11) can be replaced by the more natural condition $\lim_{t\to\infty} R(t) = 0$.

4. Simulation Results

Let us consider the random process y following the nonlinear regression model

$$y(t) = \beta_1 + \beta_2 t + \gamma_1 \cos\lambda_1 t + \gamma_2 \sin\lambda_1 t + \gamma_3 \cos\lambda_2 t + \gamma_4 \sin\lambda_2 t + \varepsilon(t);$$

$t = 1, \dots, n$ where $\theta = (\beta_1, \beta_2, \lambda_1, \lambda_2, \gamma_1, \gamma_2, \gamma_3, \gamma_4)'$ is an unknown vector of regression parameters and ε is an $AR(1)$ processs with an autoregression parameter ρ and with a variance σ^2 of a white noise.

The simulation study of the least squares estimates of θ for this model are given in Štulajter (1992). Now we'll illustrate properties of of the estimates $\hat{R}_y$ given by (5a) of the covariance function of y.

We have simulated 3 realizations of the process y of the length 51, 101, 201 with $\theta = (3, 2, 0.75, 0.25, 3, 2, 3, 4)'$, $\sigma^2 = 1$, and with different values of the autoregression parameter ρ. In the following tables corresponding values of the $\hat{R}_y$ and for comparison also values of estimates $\hat{R}_\varepsilon$ computed from realizations of the $AR(1)$ process ε with the mean value zero are given.

$n = 51$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$
t	$\rho = -0.8$		$\rho = -0.4$		$\rho = 0$		$\rho = 0.4$		$\rho = 0.8$	
0	3.11	3.22	1.01	1.18	0.78	0.98	0.82	1.13	1.14	1.72
1	-2.69	-2.69	-0.49	-0.48	-0.07	0.00	0.29	0.41	0.71	1.18
2	2.28	2.30	0.20	0.20	-0.04	-0.03	0.02	0.06	0.41	0.67
3	-2.06	-2.08	-0.19	-0.19	-0.07	-0.12	-0.05	-0.16	0.18	0.23
4	1.92	1.86	0.21	0.07	0.07	-0.09	0.00	-0.25	0.03	-0.06
5	-1.83	-1.85	-0.27	-0.26	-0.17	-0.21	-0.20	-0.33	-0.26	-0.26
6	1.64	1.64	0.17	0.15	0.01	0.00	-0.17	-0.16	-0.41	-0.21

Table 1.

$n=101$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$
t	$\rho=-0.8$		$\rho=-0.4$		$\rho=0$		$\rho=0.4$		$\rho=0.8$	
0	2.02	2.10	0.81	0.88	0.74	0.79	0.87	0.92	1.33	1.44
1	-1.63	-1.67	-0.26	-0.28	0.04	0.03	0.33	0.35	0.92	1.00
2	1.37	1.38	0.06	0.05	-0.08	-0.08	0.00	0.01	0.53	0.54
3	-1.27	-1.26	-0.10	-0.09	-0.08	-0.07	-0.10	-0.11	0.28	0.22
4	1.15	1.09	0.05	0.01	-0.04	-0.07	-0.13	-0.16	0.12	0.01
5	-1.12	-1.07	-0.14	-0.12	-0.11	-0.10	-0.16	-0.18	-0.03	-0.11
6	0.98	0.96	0.06	0.06	-0.01	-0.01	-0.09	-0.11	-0.09	-0.11

Table 2.

$n=201$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$	$\hat{R}_y(t)$	$\hat{R}_\varepsilon(t)$
t	$\rho=-0.8$		$\rho=-0.4$		$\rho=0$		$\rho=0.4$		$\rho=0.8$	
0	1.67	1.74	0.87	0.92	0.79	0.83	0.94	0.97	1.66	1.74
1	-1.21	-1.26	-0.27	-0.30	0.03	0.02	0.37	0.37	1.22	1.27
2	0.87	0.88	0.00	0.01	-0.08	-0.08	0.05	0.06	0.80	0.83
3	-0.69	-0.64	0.02	0.06	0.00	0.02	0.00	0.01	0.52	0.53
4	0.56	0.48	-0.06	-0.11	-0.06	-0.09	-0.07	-0.10	0.25	0.23
5	-0.53	-0.46	0.01	0.04	-0.01	0.00	-0.07	-0.07	0.05	0.05
6	0.49	0.48	0.00	0.00	-0.04	-0.04	-0.11	-0.12	-0.12	-0.11

Table 3.

It can be seen from these tables that the influence of an unknown mean value, following the nonlinear regression model with 8 dimensional vector of regression parameters, on estimation of a covariance function is not very big even for relatively small n ($n = 51$). For $n = 101$ and $n = 201$ the influence of the mean value is negligible for all ρ's.

References

Bates D. M. and Watts D. G., *Relative curvature measures of nonlinearity*, J. Roy. Stat. Soc. B **42** (1980), 1–25.

Box M. J., *Bias in nonlinear estimation*, J. Roy. Stat. Soc. B **33** (1971), 171–201.

Clarke G. P. Y., *Moments of the least squares estimators in a nonlinear regression models*, J. Roy. Stat. Soc. B **42** (1980), 227–237.

Cook R. D. and Tsai C. L., *Residuals in nonlinear regression*, Biometrika **72** (1985), 23–29.

———, *Bias in nonlinear regression*, Biometrika **73** (1985), 615–623.

Gallant A. R. and Goebel J. J., *Nonlinear regression with auto correlated errors*, Jour. of the Amer. Stat. Assoc. **71** (1976), 961–967.

Gallant A. R., *Nonlinear Statistical Modeling*, Wiley, New York, 1987.

Hannan E., *Non-linear time series regression*, Jour. of Appl.Prob. **8** (1971), 767–780.

Hougaard R., *The appropriateness of the asymptotic distribution in nonlinear regression model in relation to curvature*, J. Roy. Stat. Soc. B **47** (1985), 103–114.

Jennrich R. I., *Asymptotic properties of nonlinear least squares estimators*, Ann. Math. Stat. **40** (1969), 633–643.

Pázman A., *Probability distribution of the multivariate nonlinear least squares estimates*, Kybernetika **20** (1984), 209–230.

Rattkowsky D. A., *Nonlinear Regression Modeling*, Marcel Dekker, Inc., New York, 1983.

Stein M. L., *Asymptotically efficient prediction of a random field with a misspecified covariance function*, Ann. Stat. **16** (1989), 55–63.

Štulajter F., *Estimation in Stochastic Processes*, Alfa, Bratislava, 1989. (Slovak)

———, *Consistency of linear and quadratic least squares estimators in regression models with covariance stationary errors*, Appl. Math. **36** (1991), 149–155.

———, *Mean square error matrix of the approximate least squares estimator in a nonlinear regression model with correlated errors*, Acta Math. Univ. Comenianae **LXI**, 2 (1992), 251–261.

Wu C. F., *Asymptotic theory of nonlinear least squares estimators*, Ann. Stat. **9** (1981), 501–513.

F. Štulajter, Faculty of Mathematics and Physics, Comenius University, 842 15 Bratislava, Slovakia