Pliska Stud. Math. Bulgar. 17 (2005), 137–146

PLISKA
STUDIA MATHEMATICA
BULGARICA

SENSITIVITY ANALYSIS OF THE RISK OF FORECASTING FOR AUTOREGRESSIVE TIME SERIES WITH MISSING VALUES

Yu. S. Kharin, A. S. Huryn [1]

The problems of statistical forecasting of vector autoregressive time series with missing values are considered for different levels of prior information on the parameters of the underlying model. The mean square risk of forecasting and the risk sensitivity coefficient are evaluated and analyzed. Results of numerical experiments are presented.

1. Introduction

Missing values is a typical distortion of model assumptions in data analysis [6, 12]. It is fairly common for a time series to have gaps for a variety of reasons [2, 11]: 1) the data do not exist at the frequency we wish to observe them; 2) registration errors; 3) deletion of "outliers".

The autoregressive model (AR or VAR) is often used in practice for statistical analysis of time series. In statistical analysis of autoregressive time series with missing values there are five main problems: P1) evaluation of the likelihood function; P2) construction of the "optimal" (in some sense) forecast or forecasting statistic; P3) interpolation of the missing values; P4) statistical estimation of the model parameters and hypotheses testing; P5) evaluation of the risk (mean square error) of forecasting.

[1]Partially supported by INTAS 03-51-3714

2000 *Mathematics Subject Classification:* 62M20, 62M10, 62-07

Key words: Forecasting, Autoregression, Missing values, Risk, Sensitivity

An approach to analytical solution of the problems P2, P3, P5 is given in [9, 10] for univariate stationary time series: optimal (w.r.t. the minimum of the risk) linear interpolators of missing values $\{x_{n_1}, \ldots, x_{n_m}\}$, $0 = n_1 < n_2 < \ldots < n_m$, based on the knowledge of the infinite past $\{\ldots, x_{-2}, x_{-1}\}$, the finite present $\{x_{k_1}, \ldots, x_{k_r}\}$, $0 < k_1 < \ldots < k_r$, and of the covariance function $\gamma_\tau = \mathrm{cov}\{x_t, x_{t+\tau}\}, \tau = 0, \pm 1, \pm 2, \ldots$, are constructed. An approach to numerical solution of the problems P1 – P5 is presented in [5]. The EM-algorithm for the problems P1, P4, P5 is given in [6]. The problem P4 is solved in [7] and in [8] for the VAR(1), ARMA(2.1), AR(p) time series observed periodically with a known period. Note also, that "outliers"-handling techniques can be modified for missing data situations [12].

This paper is devoted mainly to the analytical solution of the problem P5 in the following aspects: to evaluate the risk of ML-forecasting for vector autoregressive time series under different levels of the prior information on the model parameters (known exactly; known with some misspecification errors; unknown); to evaluate the increment of the risk generated by the effect of missing values; to analyze sensitivity of the risk for different patterns of missing values and different levels of the prior information.

2. Mathematical model

Let the observed p-vector time series be described by the VAR(1) model:

$$Y_t = BY_{t-1} + U_t,\ t \in \mathbf{Z}, \tag{1}$$

where $\mathbf{Z}$ is the set of integers, $Y_t = (y_{t1}, \ldots, y_{tp})' \in \mathbf{R}^p$, B is a $(p \times p)$-matrix of coefficients, $U_t = (u_{t1}, \ldots, u_{tp})' \in \mathbf{R}^p$, $\{U_t\}$ are i.i.d. normal random vectors, $E\{U_t\} = 0_p$ is the zero p-vector, $E\{U_t U_t'\} = \Sigma$, $|\Sigma| \neq 0$, all eigenvalues of the matrix B are inside the unit circle. There are missing values in observations $\{Y_t\}$. For each vector Y_t the binary vector (pattern) $O_t = (o_{t1}, \ldots, o_{tp})'$ is given, where $o_{ti} = \{1$, if y_{ti} is observed; 0, if y_{ti} is a missing value$\}$. Note, that AR(p) and VAR(p) models can be transformed to VAR(1) increasing the number of components [1].

Define the finite set $M = \{(t,i),\ t \in \mathbf{Z},\ i \in \{1, \ldots, p\} : o_{ti} = 1\}$; its elements are assumed to be lexicographically ordered in ascending order; $K = |M|$ is the total number of observed components; $t_- = \min\{t : \sum_{i=1}^{p} o_{ti} > 0\}$ is the minimal time moment with observed components, $t_+ = \max\{t : \sum_{i=1}^{p} o_{ti} > 0\}$ is the maximal time moment with observed components. Define a bijection $M \leftrightarrow \{1, \ldots, K\} : k = \chi(t,i)$ and the inverse function $(t,i) = \bar{\chi}(k)$. Compose the K-vector of all observed components: $X = (x_1, \ldots, x_K)' \in \mathbf{R}^K$, $x_k =$

$y_{\bar{\chi}(k)}$, $k = 1, \ldots, K$. Note, that if $o_{ti} = 1$, $t_- \le t \le t_+$, $1 \le i \le p$, then the process Y_t is observed on $[t_-, t_+]$ without any missing value, $K = (t_+ - t_- + 1)p$, $X = \left(Y'_{t_-}, Y'_{t_-+1}, \ldots, Y'_{t_+}\right)'$, $\chi(t,i) = i + (t - t_-)p$; $\bar{\chi}(k) = ([(k-0.5)/p] + 1, (k-1) \bmod p + 1)$, $k = 1, \ldots, K$.

We will also consider the AR(p)-model as a special case of (1):

$$y_t = \beta' Y_{t-1} + \xi_t, \ t \in \mathbf{Z}, \tag{2}$$

where $Y_t = (y_t, \ldots, y_{t-p+1})' \in \mathbf{R}^p$, $\beta \in \mathbf{R}^p$ is a p-vector of coefficients, $\{\xi_t\}$ are i.i.d. normal variables, $\mathcal{L}(\xi_t) = \mathcal{N}(0, \sigma^2)$; the pattern $O_t = \{1,$ if y_t is observed; 0, if y_t is a missing value$\}$; $M = \{t, \ t \in \mathbf{Z} : O_t = 1\}$, $K = |M| = \sum_{t \in \mathbf{Z}} O_t$, $k = \chi(t) = \sum_{i \le t} O_i$.

Let $Y_{t_++\tau} \in \mathbf{R}^p$ be a "future vector" to be forecasted for $\tau \ge 1$, $\hat{Y}_{t_++\tau} = \hat{Y}_{t_++\tau}(X) : \mathbf{R}^K \to \mathbf{R}^p$ be a forecasting statistic (procedure). Introduce the $(p \times p)$-matrix risk R and the (scalar) risk r of forecasting:

$$R = \mathbf{E}\left\{\left(\hat{Y}_{t_++\tau} - Y_{t_++\tau}\right)\left(\hat{Y}_{t_++\tau} - Y_{t_++\tau}\right)'\right\}, r = \mathrm{tr}(R) \ge 0. \tag{3}$$

It is known [2], that for the case of complete observations and known parameters B, Σ the minimal risk $r_0^* = \mathrm{tr}\left(\sum_{i=0}^{\tau-1} B^i \Sigma (B')^i\right) > 0$ is attained for the forecast: $\tilde{Y}_{t_++\tau} = B^\tau Y_{t_+}$. To evaluate the sensitivity of the risk let us use the risk sensitivity coefficient [3, 4]:

$$\varkappa = (r - r_0^*)/r_0^* \ge 0; \tag{4}$$

it is the relative increment of the risk r generated by missing values w.r.t. the minimal risk r_0^*.

3. ML-forecasting under missing values and known B, Σ

Introduce the matrices:

$$F = (F_{ij}) = \mathrm{cov}\{X, X\}, G = (G_{ij}) = \mathrm{cov}\{X, Y_{t_++\tau}\},$$

$$H = (H_{ij}) = \mathrm{cov}\{Y_{t_++\tau}, Y_{t_++\tau}\}, A_0 = A_0(B, \Sigma) = G' F^{-1}.$$

Lemma 1. *The following expressions for F, G, H hold:*

$$F_{ij} = F_{ji} = \left(B^{\bar{\chi}_1(i) - \bar{\chi}_1(j)} H\right)_{\bar{\chi}_2(i), \bar{\chi}_2(j)}, \ i, j = 1, \ldots, K, \ i \ge j; \tag{5}$$

$$G_{ij} = \left(B^{(t_{+}+\tau)-\bar{\chi}_1(i)}H\right)_{j,\bar{\chi}_2(i)}, \; i=1,\ldots,K, \; j=1,\ldots,p; \; H = \sum_{i=0}^{\infty} B^i\Sigma(B')^i.$$

Proof. Using the expression for the covariance matrix for the VAR(1) model [1]: $\operatorname{cov}\{Y_i, Y_j\} = B^{i-j}H$, $i \geq j$, we find: $F_{i,j} = \operatorname{cov}\{x_i, x_j\} = \operatorname{cov}\{y_{\bar{\chi}(i)}, y_{\bar{\chi}(j)}\} = \left(\operatorname{cov}\{Y_{\bar{\chi}_1(i)}, Y_{\bar{\chi}_1(j)}\}\right)_{\bar{\chi}_2(i),\bar{\chi}_2(j)} = \left(B^{\bar{\chi}_1(i)-\bar{\chi}_1(j)}H\right)_{\bar{\chi}_2(i),\bar{\chi}_2(j)}$. In the same way we find H and G. □

Theorem 1. *If the true values B, Σ are known and $|F| \neq 0$, then the ML-forecasting statistic and its risk functionals are*

$$\hat{Y}_{t_{+}+\tau} = \mathbf{E}\{Y_{t_{+}+\tau}|X\} = A_0X, \tag{6}$$

$$R_0 = H - G'F^{-1}G \succeq 0, \; r_0 = \operatorname{tr}(H) - \operatorname{tr}(F^{-1}GG'). \tag{7}$$

Proof. Denote $Y_{+} = (X', Y'_{t_{+}+\tau})' \in \mathbf{R}^{K+p}$. By the model (1) conditions, the vector Y_{+} has the normal distribution. By the Anderson theorem [1] the likelihood function(w.r.t $Y_{t_{+}+\tau}$) is:

$$l\left(Y_{t_{+}+\tau}; B, \Sigma\right) = n_K\left(X|0_K, F\right) n_p\left(Y_{t_{+}+\tau}|G'F^{-1}X, H - G'F^{-1}G\right), \tag{8}$$

where $n_K\left(X|\mu, \Sigma\right)$ means the K-dimensional normal p.d.f. with the parameters μ, Σ. The ML-forecast is the solution of the extremum problem:
$l\left(Y_{t_{+}+\tau}; B, \Sigma\right) \rightarrow \max_{Y_{t_{+}+\tau}}$. Since the first multiplier in (8) does not depend on $Y_{t_{+}+\tau}$, we come to the unique solution (6): $\hat{Y}_{t_{+}+\tau} = G'F^{-1}X = A_0X$. Using the total mathematical expectation formula and (6) we find the risk (7):
$R_0 = \mathbf{E}\left\{\left(\hat{Y}_{t_{+}+\tau} - Y_{t_{+}+\tau}\right)(\hat{Y}_{t_{+}+\tau} - Y_{t_{+}+\tau})'\right\} = \mathbf{E}\left\{\operatorname{cov}\left\{Y_{t_{+}+\tau}, Y_{t_{+}+\tau}|X\right\}\right\} = \mathbf{E}\left\{H - G'F^{-1}G\right\} = H - G'F^{-1}G$. □

Corollary 1. *The risk sensitivity coefficient for the ML-forecast (6) is* $\varkappa_0 = \left(r_0 - r_0^*\right)/r_0^* = \left(\sum_{i=\tau}^{\infty} \operatorname{tr}\left(B^i\Sigma(B')^i\right) - \operatorname{tr}(G'F^{-1}G)\right) / \sum_{i=0}^{\tau-1} \operatorname{tr}\left(B^i\Sigma(B')^i\right) \geq 0.$

Formulate now important corollaries for the univariate autoregressive model AR(p), defined by (2).

Corollary 2. *Let the AR(p) model (2) takes place and there exists the time moment* $t_* = \max\{t \in M : O_{t_*} = \ldots = O_{t_*+p-1} = 1\}$. *Then the ML-forecast (6) does not depend functionally on the observations* $\{y_t : t < t^*\}$*:* $A_0 = (A_{01}, \ldots, A_{0K})$, $A_{01} = \ldots = A_{0,\chi(t_*-1)} = 0$.

Proof. One can come to the statement using the mathematical induction method. □

This corollary means that the optimal forecast and its risk r_0 are non-sensitive to the "sub-pattern" $\{t \in M : t \leq t_* - 1\}$ for the case when the true model parameters B, Σ are known exactly.

Corollary 3. *If the model (2) under the pattern $M_m = \{t_-, \ldots, t_+\} \setminus \{m\}$ with the only one missing value y_m at a time moment $m \in \{t_+ - p + 1, \ldots, t_+\}$ takes place, and $K = t_+ - t_- \geq 2p$, then the risk sensitivity coefficient for the ML-forecast at $\tau = 1$ is:*

$$\varkappa(m) = \beta^2_{t_+ +1-m}/(1 + \beta^2_1 + \ldots + \beta^2_{t_+ - m}). \tag{9}$$

Proof. Using the Box-Jenkins representation of the likelihood function, integrating it by y_m and maximizing by $y_{t_+ +1}$, one can come to (9). □

One can see from (9), that if the autoregressive coefficient $\beta_{t_+ +1-m} = 0$, then $\varkappa(m) = 0$ and the risk of forecasting is non-sensitive to the missing value y_m. The formula (9) gives also the expression for the variation zone of the risk sensitivity coefficient: $\varkappa_- = \min_{0 \leq j \leq p-1}(s_{j+1} - s_j)/s_j \leq \varkappa \leq \varkappa_+ = \max_{0 \leq j \leq p-1}(s_{j+1} - s_j)/s_j$, where $s_0 = 1$, $s_j = 1 + \beta^2_1 + \ldots + \beta^2_j$. The most influencing is the missing value y_{m^*} at the time moment m^* for which $\varkappa(m^*) = \varkappa_+$.

4. Statistical forecasting in the case of unknown B, Σ

Theorem 2. *If B, Σ are unknown, $|F| \neq 0$, then the ML-forecast of $Y_{t_+ +\tau}$ is*

$$\hat{Y}_{t_+ +\tau} = A_0(\hat{B}, \hat{\Sigma})X, \tag{10}$$

where the ML-estimators $(\hat{B}, \hat{\Sigma})$ are the solution of the minimization problem: $l_1(B, \Sigma) = X'F^{-1}X + \ln|F| + \ln|H - G'F^{-1}G| \to \min_{B,\Sigma}$.

Proof. According to (8), the joint ML-estimators of $Y_{t_+ +\tau}$, B, Σ are the solution of the extremum problem: $l\left(Y_{t_+ +\tau}; B, \Sigma\right) \to \max_{Y_{t_+ +\tau}, B, \Sigma}$. From Theorem 1 we get (10), where $\hat{B}$, $\hat{\Sigma}$ are the solution of the problem: $n_K\left(X|0_K, F\right) n_p\left(G'F^{-1}X|G'F^{-1}X, H - G'F^{-1}G\right) \to \max_{B,\Sigma}$. Taking the logarithm, we come to the statement. □

Because of the computation complexity of the minimization problem in (10), let us adjust the LS-estimators. Consider the situation with "homogeneous observation patterns": $O_t = 0_p$ or $O_t = (1, \ldots, 1) \in \mathbf{R}^p$. Denote $t_0 = t_-$ and represent the sequence of registered observations in the form: $Y_{t_0}, Y_{t_0+1}, \ldots, Y_{t_0+l_0}$,

$\ldots, Y_{t_n}, Y_{t_n+1}, \ldots, Y_{t_n+l_n}$, $\Im = \{t_0+1, \ldots, t_0+l_0, \ldots, t_n+1, \ldots, t_n+l_n\}$, $T_0 = |\Im| = l_0 + \ldots + l_n \geq n+1$, where $n+1$ is the number of series of observations without missing values. Introduce the notation: $\otimes$ is the Kronecker matrix product, $\mathrm{vec}(A)$ is the vectorization of the matrix A by rows. Assume, that $T_0 > p$, $|\sum_{t\in\Im} Y_{t-1}Y'_{t-1}| \neq 0$, and introduce the following modifications of the Anderson estimators [1] for the situation with missing values:

$$\check{B} = \sum_{t\in\Im} Y_t Y'_{t-1} \left(\sum_{t\in\Im} Y_{t-1}Y'_{t-1} \right)^{-1}, \quad \check{\Sigma} = \frac{1}{T_0} \sum_{t\in\Im} \left(Y_t - \check{B}Y_{t-1}\right)\left(Y_t - \check{B}Y_{t-1}\right)'.$$

Theorem 3. *If $T_0 \to \infty$, $T_0/n \to \infty$, and the matrix H is positive defined, then the estimators $\check{B}$, $\check{\Sigma}$ are consistent: $\mathrm{plim}_{T_0\to\infty}\check{B} = B$, $\mathrm{plim}_{T_0\to\infty}\check{\Sigma} = \Sigma$, and the random vector $\mathrm{vec}\left(\sqrt{T_0}\left(\check{B}' - B'\right)\right)$ has the asymptotically normal distribution with the zero mean and the covariance matrix $H^{-1} \otimes \Sigma$.*

P r o o f. At first, let us prove the convergence $\mathrm{plim}_{T_0\to\infty}\left(T_0^{-1}\sum_{t\in\Im} Y_t Y'_t - T_0^{-1}\sum_{t\in\Im} Y_{t-1}Y'_{t-1}\right) = 0_{p\times p}$. It follows from the law of large numbers and the next inequality: $\mathbf{D}\left\{T_0^{-1}\sum_{k=0}^{n}\left(y_{t_k+l_k,i}y_{t_k+l_k,j} - y_{t_k,i}y_{t_k,j}\right)\right\} \leq (n/T_0)^2 *$
$\max_{0\leq k\leq n}\mathbf{D}\left\{y_{t_k+l_k,i}y_{t_k+l_k,j} - y_{t_k,i}y_{t_k,j}\right\} \leq 4\,(n/T_0)^2\,\mathbf{E}\left\{\left(y_{t_0,i}y_{t_0,j}\right)^2\right\}$.

Further, following to the scheme of the proof of this statement for the case with complete data [1], we come to the convergence in probability and to the asymptotic normality. □

Theorem 4. *Let $Y_{t_++\tau}$ be forecasted by the statistic: $\hat{Y}_{t_++\tau} = AX$, where A is a $(p\times K)$-matrix of coefficients: $A = A_0 + a$, a is any $(p\times K)$-matrix of nonrandom misspecification error. Then the $(p\times p)$-matrix risk of forecasting $R_1 = R_0 + aFa'$, where R_0 is the risk determined by (7).*

P r o o f. According to the definition of the risk, the theorem conditions and the result of Theorem 1

$$r_1 = \mathbf{E}\left\{\left((A_0+a)X - Y_{t_++\tau}\right)\left((A_0+a)X - Y_{t_++\tau}\right)'\right\} =$$

$$= \mathbf{E}\left\{aXX'a' + \left(\mathbf{E}\{Y_{t_++\tau}|X\} - Y_{t_++\tau}\right)\left(\mathbf{E}\{Y_{t_++\tau}|X\} - Y_{t_++\tau}\right)'\right\} +$$

$$+\mathbf{E}\left\{aX\left(\mathbf{E}\{Y_{t_++\tau}|X\} - Y_{t_++\tau}\right)' + \left(\mathbf{E}\{Y_{t_++\tau}|X\} - Y_{t_++\tau}\right)X'a'\right\}.$$

Using the total mathematical expectation formula and the equation

$$\mathbf{E}\left\{X'\left(\mathbf{E}\{Y_{t_++\tau}|X\} - Y_{t_++\tau}\right)\right\} =$$

$$= \mathbf{E}\left\{\mathbf{E}\left\{\left(X'\left(\mathbf{E}\{Y_{t_++\tau}|X\} - Y_{t_++\tau}\right)\right)|X\right\}\right\} = 0,$$

we come to the statement. □

Corollary 4. *For the AR(p) model (2) under misspesification the risk sensitivity coefficient (4) is* $\varkappa_1 = \varkappa_0 + a'Fa/\sum_{i=0}^{\tau-1}\mathrm{tr}(B^i\Sigma(B')^i)$.

Consider the case of bounded misspecification error $a \in \mathbf{R}^p$: $0 \le \|a\| \le \gamma$, where $\gamma \ge 0$ is a known upper bound. Then maximizing $\varkappa_1$ w.r.t. a one can find the variation zone for the risk sensitivity coefficient: $\varkappa_0 \le \varkappa_1 \le \varkappa_{1+} = \varkappa_0 + \gamma^2\lambda_{\max}(F)/\sum_{i=0}^{\tau-1}\mathrm{tr}(B^i\Sigma(B')^i)$, where $\lambda_{\max}(F)$ is the maximal characteristic number of the matrix F.

Theorem 5. *Let* $\hat{Y}_{t_++\tau} = \hat{A}X$, *where* $\hat{A} = A_0 + a$ *is an unbiased statistical estimator of* A_0 *with the covariances:* $\mathbf{E}\left\{\left(\hat{A} - A_0\right)_{ij}\left(\hat{A} - A_0\right)_{kl}\right\} = V_{i,j,k,l}, i,k = 1,\ldots,p; j,l = 1,\ldots,K$. *If the estimator* $\hat{A}$ *is independent on* X, *then the matrix risk of forecasting* $R_2 = R_0 + \left(\sum_{s,t=1}^{K} F_{st}V_{i,s,j,t}\right)_{i,j=1,\ldots,p}$.

Proof. According to Theorem 4 we have:

$$(R_2)_{ij} = \left(\mathbf{E}\left\{\mathbf{E}\left\{\left(\hat{Y}_{t_++\tau} - Y_{t_++\tau}\right)\left(\hat{Y}_{t_++\tau} - Y_{t_++\tau}\right)'|a\right\}\right\}\right)_{ij} =$$

$$\left(\mathbf{E}\left\{R_1(a)\right\}\right)_{ij} = \left(\mathbf{E}\left\{R_0 + aFa'\right\}\right)_{ij} =$$

$$(R_0)_{ij} + \mathbf{E}\left\{\sum_{s,t=1}^{K} a_{is}F_{st}a_{jt}\right\} = (R_0)_{ij} + \sum_{s,t=1}^{K} F_{st}V_{i,s,j,t}.$$

□

Note, that the condition of independence of $\hat{A}, X$ is satisfied if the estimator $\hat{A}$ is constructed by the data $\tilde{X}$ independent on X. Such a situation takes place if the "learning stage" is separated from the "forecasting stage" [4].

Corollary 5. *If* $\hat{A}$ *is a consistent estimator:* $\sum_{s,t=1}^{K}|V_{i,s,j,t}| \to 0$, *then* $R_2 \to R_0$.

Proof. Using the inequality for the elements of the covariance matrix F [1]: $|(F)_{ij}| \le c < +\infty$, where c does not depend on i and j, we get the inequality: $\left|(R_2 - R_0)_{ij}\right| = \left|\sum_{s,t=1}^{K} F_{st}V_{i,s,j,t}\right| \le c\sum_{s,t=1}^{K}|V_{i,s,j,t}| \to 0$. □

Corollary 6. *For the AR(p) model (2) under unbiased estimation of A_0 the risk sensitivity coefficient (4) is* $\varkappa_2 = \varkappa_0 + \sum_{i=1}^{p}\left(\sum_{s,t=1}^{K} F_{st}V_{i,s,i,t}\right)$ $/ \sum_{i=0}^{\tau-1} \mathrm{tr}(B^i\Sigma(B')^i)$.

5. Numerical results

To compare the theoretical and experimental results we consider the AR(11) model ($p = 11$) for the classical centered data "The Canadian Lynx data 1821 - 1934" [13]: $y_t = 1.0938y_{t-1} - 0.3571y_{t-2} - 0.1265y_{t-4} + 0.3244y_{t-10} - 0.3622y_{t-11} + \xi_t$, $\sigma^2 = 0.04405$, $t_- = 1$, $t_+ = T = 113$, under a single missing value at the time moment $m \in \{T-p+1, \ldots, T-1\}$ (pattern M_m) and different levels of prior information on model parameters. Note, that in [14] one can find an interesting review of previous analyses of this celebrated set of data. The Monte-Carlo experiments with 100 000 simulations of time series were used to evaluate the experimental value of the risk sensitivity coefficient $\varkappa$ and its 95%-confidence limits.

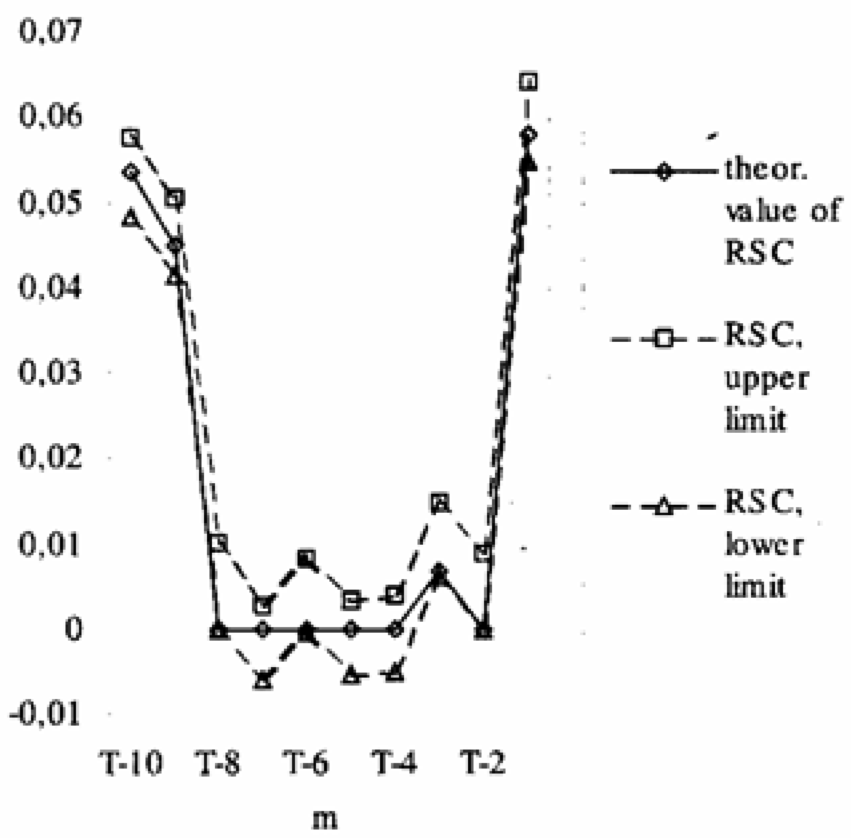

Figure 1: Risk sensitivity coefficient(RSC) in the case of known parameters

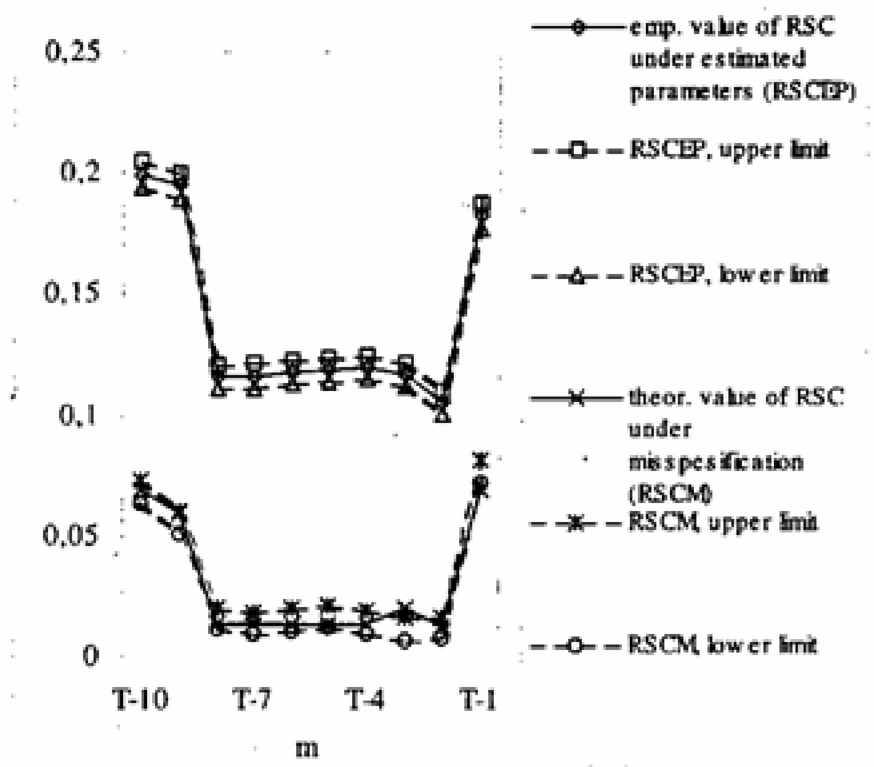

Figure 2: Risk sensitivity coefficient in the case of unknown parameters

Fig.1 presents values of the theoretical risk sensitivity coefficient (9) and its 95%-confidence limits for different values of m for the case of known values of model parameters. Fig. 2 presents theoretical values of the risk sensitivity coefficient $\varkappa_1$ (see Corollary 9) and its 95%-confidence limits under misspecification error $a_{m-p} = a_{m-p+1} = \ldots = a_{t_+-1} = 0.01$, and also the point and interval estimates of the risk sensitivity coefficient $\varkappa_2$ for the case of the "plug-in" forecasting procedure based on the proposed estimators $\check{B}$, $\check{\Sigma}$. Fig. 1,2 show a sufficiently good fit of theoretical and experimental results.

REFERENCES

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

[2] Greene W.H. Econometric Analysis. Macmillan, New York. 2000.

[3] Kharin Yu. Robustness in Statistical Pattern Recognition. Kluwer, Dordrecht. 1996.

[4] KHARIN YU. Robustness Analysis in Forecasting of Time Series, in: R. Dutter et al., eds., Development in Robust Statistics, Springer: Heidelberg / New York. 2000. 180–193.

[5] KOHN R., C. F. ANSLEY Estimation, Prediction and Interpolation for ARIMA Models with Missing Data. *JASA* **81(395)** (1986), 751–761.

[6] LITTLE R., D. RUBIN Statistical Analysis with Missing Data. John Wiley, New York. 1987.

[7] NIJMAN T., F. PALM Parameter Identification in ARMA Processes in the Presence of Regular but Incomplete Sampling. *JTSA* **11(3)** (1990), 239–238.

[8] PANTULA S.G., D. SHIN Testing for a Unit Root in AR Processes with Systematic but Incomplete Sampling. *Stat. & Prob. Letters* **18** (1993), 183–190.

[9] POURAHMADI M. Estimation and Interpolation of Missing Values of a Stationary Time Series. *JTSA* **10(2)** (1989), 149–169.

[10] POURAHMADI M. Foundations of Time Series and Prediction Theory. John Wiley, New York. 2001.

[11] SHAFER J.L. Analysis of Incomplete Data. Chapman and Hall, London. 1997.

[12] STOCKINGER N., R. DUTTER Robust Time Series Analysis: a Survey. *Kybernetika* **23** (1987) 1–91.

[13] TONG H. Some Comments on the Canadian Lynx Data. *J. Roy. Statist. Soc.* Ser. A. **140** 432–436.

[14] TONG H. Non-linear Time Series. Clarendon Press, Oxford. 1999.

Yuriy Kharin
Department of Mathematical Modeling and Data Analysis
Belarusian State University
Fr. Skoriny av. 4,
220050 Minsk, Belarus
e-mail: kharin@bsu.by

Aliaksandr Huryn
Department of Mathematical Modeling and Data Analysis
Belarusian State University
Fr. Skoriny av. 4,
220050 Minsk, Belarus
e-mail: hurynaliaksandr@yahoo.com