УДК 519.24

А.С. ГУРИН

АСИМПТОТИЧЕСКОЕ РАЗЛОЖЕНИЕ РИСКА ПРОГНОЗИРОВАНИЯ AR(1) ВРЕМЕННЫХ РЯДОВ ПРИ НАЛИЧИИ ПРОПУЩЕННЫХ ДАННЫХ

The problem of forecasting of autoregressi ve time series under presence of missing data is considered in the paper. The estimators of model parameters, constructed by method of moments, are used to construct the «plug-in» forecasting statistic. The asymptotic expansion of the risk of the «plug-in» forecasting is found.

Модель авторегрессии часто используется на практике для статистического анализа и моделирования динамических процессов, временных рядов в экономике, медицине, защите окружающей среды и других областях [1, 2]. Наличие пропущенных данных - часто встречающееся явление из-за несовершенства измерительных приборов или малой степени доверия к данным, которое значительно усложняет поиск и реализацию оптимальных алгоритмов оценивания и прогнозирования таких временных рядов [3,4].

В [5] для авторегрессионных временных рядов с «пропусками» построены прогнозы по методу максимального правдоподобия при известных модельных параметрах и оценки модельных параметров по методу моментов. В настоящей статье для «подстановочного» («plug-in») прогнозирования, использующего оценки по методу моментов [5], строится асимптотическое разложение риска прогнозирования на основе методов робастного статистического анализа данных [6, 7].

Математическая модель и постановка задач

Пусть наблюдаемый временной ряд $Y_t \in \mathbf{R}$ описывается моделью авторегрессии первого порядка AR(1) [1]:

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

где B - параметр авторегрессии, $|B| < 1$, $\{U_t\}_{t\in\mathbf{Z}}$ - независимые в совокупности нормально распределенные случайные величины: $L\{U_t\} = N(0,\Sigma)$, $\Sigma > 0$, $t \in \mathbf{Z}$.

В наблюдениях $\{Y_t\}$ имеются пропущенные данные, или «пропуски». Шаблоном «пропусков» назовем двоичную последовательность O_t, где $O_t = 1$, если Y_t наблюдается, $O_t = 0$, если Y_t не наблюдается, $t \in \mathbf{Z}$. Обозначим минимальный и максимальный моменты времени с наблюдаемыми компонентами: $t_- = \min\{\, t : O_t = 1\}$, $t_+ = \max\{\, t : O_t = 1\}$; без ограничения общности $t_- = 1$, $t_+ = T$.

Задача состоит в построении статистического прогноза $\hat{Y}_{T+1}$ и нахождении риска прогнозирования.

«Подстановочное» прогнозирование

Сформулируем дополнительное предположение о шаблоне $\{O_t\}$. При $T \to \infty$ имеют место асимптотики

$$\sum\nolimits_{t=1}^{T-k} O_{t+k}O_t \Big/ (T-k) \to \vartheta_k \in (0,\, 1],$$
$$\sum\nolimits_{t,t'=1}^{T-1} O_{t+k}O_tO_{t'+k'}O_{t'}\delta_{t-t',\tau} \Big/ \left(T - |\tau| - 1\right) \to \tilde{\vartheta}_{\tau,k,k'} \in [0,\, 1], \tag{*}$$

ϑ_k - предельная частота наблюдения пары значений временного ряда в моменты времени, сдвинутые на k, $\tilde{\vartheta}_{\tau,k,k'}$ - предельная частота наблюдения двух пар значений временного ряда в моменты времени, сдвинутые на $\tau + k - k'$ и τ единиц времени от первой пары, где $\tau \in \mathbf{Z}$, $k,\ k' \in \{0,\ 1\}$, $\delta_{a,b}$ — символ Кронекера. Очевидно, что имеет место соотношение $\tilde{\vartheta}_{0,k,k} = \vartheta_k$.

Обозначим ковариации (с учетом свойства стационарности [5]): $G_{t-t'} = \mathbf{Cov}\{Y_t,\ Y_{t'}\}$, $g_{t-t',k,k'} = \mathbf{Cov}\{Y_{t+k}Y_t, Y_{t'+k'}Y_{t'}\}$, $t,\ t' \in \mathbf{Z}$, $k,\ k' \in \{0,\ 1\}$. Очевидно, что $G_t = B^{|t|}G_0$, $G_0 = \dfrac{\Sigma}{1-B^2}$. Назовем

$$T_0 = \inf\left\{T' \in \mathbf{N} : \min\nolimits_{k\in\{0,1\}} \sum\nolimits_{t=1}^{T'-k} O_{t+k}O_t > 0\right\}$$

критическим временем наблюдения для заданного шаблона $\{O_t\}$. Для $T \geq T_0$ определим предельные характеристики шаблона ($\tau \in \mathbf{Z}$, $k, k' \in \{0,\ 1\}$)

$$C_{\tau,k,k'} = \tilde{\vartheta}_{\tau,k,k'} \Big/ \left(\vartheta_k \vartheta_{k'}\right),$$

выборочные ковариации

$$\hat{G}_k = \sum_{t=1}^{T-k} Y_{t+k} Y_t O_{t+k} O_t \Big/ \sum_{t=1}^{T-k} O_{t+k} O_t \quad (2)$$

и при $\hat{G}_0 \neq 0$ статистики, основанные на выборочных ковариациях (2):

$$\hat{B} = \frac{\hat{G}_1}{\hat{G}_0}, \ \hat{\Sigma} = \hat{G}_0 - \frac{\hat{G}_1^2}{\hat{G}_0}. \quad (3)$$

При известном модельном параметре B [1] оптимальным в среднеквадратическом смысле для единичной глубины прогнозирования является прогноз

$$\hat{Y}_{T+1} = BY_T,$$

который имеет риск прогнозирования

$$R_T = \mathbf{E}\left\{\left(\hat{Y}_{T+1} - Y_{T+1}\right)^2\right\} = \Sigma > 0.$$

Для неизвестных модельных параметров построим «подстановочный» прогноз, используя оценки (3):

$$\hat{Y}_{T+1,\text{"plug-in"}} = \hat{B} Y_T.$$

Качество прогнозирования будем также характеризовать среднеквадратическим риском прогнозирования:

$$R_{T,\text{"plug-in"}} = \mathbf{E}\left\{\left(\hat{Y}_{T+1,\text{"plug-in"}} - Y_{T+1,\text{"plug-in"}}\right)^2\right\} > 0. \quad (4)$$

Асимптотическое разложение риска

Докажем вспомогательные утверждения.

Теорема 1. Для модели (1) выполняется $\forall \lambda \in (|B|, 1)$

$$\mathbf{Cov}\left\{Y_{t_1}Y_{t_2}, Y_{t_3}Y_{t_4}\right\} = G_{t_1-t_3}G_{t_2-t_4} + G_{t_1-t_4}G_{t_2-t_3},$$

$$\left|\mathbf{Cov}\left\{Y_{t_1}Y_{t_2}, Y_{t_3}Y_{t_4}\right\}\right| \le \mathrm{Const}_{\bar{t}_1,\bar{t}_2,\bar{t}_3,\bar{t}_4}\left(\lambda^{|t_1-t_3|+|t_2-t_4|} + \lambda^{|t_1-t_4|+|t_2-t_3|}\right), \ t_1,t_2,t_3,t_4 \in \mathbf{Z},$$

$$\left|\mathbf{E}\left\{\left(Y_{t_1}Y_{t_2} - \mathbf{E}\left\{Y_{t_1}Y_{t_2}\right\}\right)\cdots\left(Y_{t_7}Y_{t_8} - \mathbf{E}\left\{Y_{t_7}Y_{t_8}\right\}\right)\right\}\right| \le$$

$$\le \mathrm{Const}_{\bar{t}_1,\ldots,\bar{t}_8} \sum_{(l_1,\ldots,l_8)\in C_8(2,2,2,2)} \lambda^{|t_{l_1}-t_{l_2}|+|t_{l_3}-t_{l_4}|+|t_{l_5}-t_{l_6}|+|t_{l_7}-t_{l_8}|}, \ t_1,\ldots,t_8 \in \mathbf{Z},$$

где

$$C_8(2,\ 2,\ 2,\ 2) = \left\{(l_1,\ \ldots,\ l_8) : l_n \in \{1,\ \ldots,\ 8\},\ l_n \neq l_m \text{ при } n \neq m,\ n,\ m \in \{1,\ \ldots,\ 8\};\right.$$

$$l_1 < l_2,\ l_3 < l_4,\ l_5 < l_6,\ l_7 < l_8,\ l_1 < l_3 < l_5 < l_7;$$

$$\left.(l_1,\ l_2),(l_3,\ l_4),(l_5,\ l_6),(l_7,\ l_8) \notin \{(1,\ 2),\ (3,\ 4),\ (5,\ 6),\ (7,\ 8)\}\right\},$$

$$\left|C_8(2,\ 2,\ 2,\ 2)\right| = 60,$$

$\mathrm{Const}_{\bar{i}_1,\ldots,\bar{i}_n}$ - математический объект (число, вектор, матрица), не зависящий от $i_1,\ldots,i_n$.

Доказательство. См. [5]. ■

Лемма 1. Пусть последовательность а, (Г) ограничена и при $T \to \infty$ сходится:

$$\left|a_t(T)\right| \le a \in (0,\ +\infty),\ a_t(T) \to a_t,\ t \in \mathbf{Z},\ T \in \mathbf{N},$$

и $|b| < 1$. Тогда при $T \to \infty$

$$\sum_{t=-T}^{T} b^{|t|} a_t(T) \to \sum_{t=-\infty}^{\infty} b^{|t|} a_t.$$

Доказательство. Соответствующую разность

$$\left|\sum_{t=-T}^{T} b^{|t|} a_t(T) - \sum_{t=-\infty}^{+\infty} b^{|t|} a_t\right| = \left|\sum_{t=-T}^{T} b^{|t|}\left(a_t(T) - a_t\right) - \sum_{t=-\infty}^{-(T+1)} b^{|t|} a_t - \sum_{t=T+1}^{+\infty} b^{|t|} a_t\right| \le$$

$$\le \left|\sum_{t=-T}^{T} b^{|t|}\left(a_t(T) - a_t\right)\right| + 2a\frac{|b|^{T+1}}{1-|b|} \le$$

$$\leq \left| \sum_{t=-m}^{m} b^{|t|} \left(a_t(T) - a_t \right) \right| + 4a \frac{|b|^{m+1}}{1-|b|} + 2a \frac{|b|^{T+1}}{1-|b|}, \ m \leq T,$$

можно сделать сколь угодно малой, так как последние слагаемые можно уменьшить выбором m, затем при фиксированном m первое - выбором T. ■

Лемма 2. Пусть случайная последовательность y_t, $t \in \mathbf{Z}$, такова, что в некоторой δ-окрестности нуля для некоторой положительной константы C_1 и для любых последовательных T_0 отсчетов существует совместная плотность распределения, не превосходящая этой константы:

$$\exists \delta \in (0,+\infty), \ \exists C_1 \in (0,+\infty), \ \exists T_0 \in \mathbf{N}, \ \forall x_1,...,x_{T_0} \in \mathbf{R}, \ x_1^2 + \cdots + x_{T_0}^2 \leq \delta^2, \ \forall s \in \mathbf{Z},$$

$$P_{y_s,...,y_{s-T_0-1}}\left(x_1,...,x_{T_0}\right) \leq C_1.$$

Тогда существует положительная константа C, такая, что для любых последовательных $T \geq T_0$ отсчетов математическое ожидание случайной величины, обратной к выборочной дисперсии и возведенной в натуральную степень $n < \dfrac{T_0}{2}$, не превосходит этой константы:

$$\exists C_2 \in (0, +\infty), \ \forall T \in \mathbf{N}, \ T \geq T_0, \ \forall n \in \mathbf{N}, \ n < \frac{T_0}{2} \quad \mathbf{E}\left\{ \frac{1}{\hat{G}_0^n} \right\} \leq C_2,$$

где $\hat{G}_0 = \dfrac{y_1^2 + \cdots + y_T^2}{T}$.

Доказательство. Пусть T, T_0, $n \in \mathbf{N}$, $T \geq T_0$, $k = \left[\dfrac{T}{T_0} \right] \in \mathbf{N}$. Справедливо неравенство для функции под знаком математического ожидания:

$$\frac{1}{\hat{G}_0^n} \leq \frac{(2T_0)^n}{k} \left(\left(\frac{1}{y_1^2 + \cdots + y_{T_0}^2} \right)^n + \cdots + \left(\frac{1}{y_{(k-1)T_0+1}^2 + \cdots + y_{kT_0}^2} \right)^n \right).$$

Не ограничивая общности, предположим, что $\delta \in (0, 1)$. Для каждого слагаемого $\forall s \in \mathbf{Z}$ имеем

$$\mathbf{E}\left\{ \left(\frac{1}{y_s^2 + \cdots + y_{s+T_0-1}^2} \right)^n \right\} = \mathbf{E}\left\{ \left(\frac{1}{y_s^2 + \cdots + y_{s+T_0-1}^2} \right)^n ; y_s^2 + \cdots + y_{s+T_0-1}^2 \leq \delta^2 \right\} +$$

$$+ \mathbf{E}\left\{ \left(\frac{1}{y_s^2 + \cdots + y_{s+T_0-1}^2} \right)^n ; y_s^2 + \cdots + y_{s+T_0-1}^2 > \delta^2 \right\} \leq$$

$$\leq \int_{x_1^2 + \cdots + x_{T_0}^2 \leq \delta^2} \left(\frac{1}{x_1^2 + \cdots + x_{T_0}^2} \right)^n P_{y_s,...,y_{s-T_0-1}}\left(x_1,...,x_{T_0}\right) dx_1 \cdots dx_{T_0} + \frac{1}{\delta^{2n}} \leq$$

$$\leq C_1 \int_{r \in [0,\delta], \ \theta_1 \in [0,2\pi], \ \theta_m \in [0,\pi], \ m \in \{2,...,T_0-1\}} r^{T_0-1-2n} \Pi_{k=1}^{T_0-1} \sin^{k-1}\theta_k \, dr d\theta_1 \cdots d\theta_{T_0-1} + \delta^{-T_0} \leq$$

$$\leq 2C_1 \pi^{T_0-1} \frac{\delta^{T_0-2n}}{T_0-2n} + \delta^{-T_0} \leq 2C_1 \pi^{T_0-1} + \delta^{-T_0},$$

где использовалась сферическая замена:

$$\begin{cases} x_1 = r \sin\theta_1 \cdots \sin\theta_{T-1}, \\ x_m = r \cos\theta_{m-1} \Pi_{k=1}^{T-1} \sin\theta_k, \ m \in \{2, ..., T-1\}, \\ x_T = r \cos\theta_{T-1}, \end{cases}$$

$x_1,\ ...,\ x_T \in \mathbf{R},\ \ r\in[0,\ +\infty),\ \ \theta_1\in[0,\ 2\pi],\ \ \theta_m\in[0,\pi],\ \ m\in\{2,\ ...,\ T-1\},\ \ x_1^2+\cdots+x_T^2 =$ $= r^2$, якобиан: $\dfrac{D(x_1,\ ...,\ x_T)}{D(r,\theta_1,\ ...,\ \theta_{T-1})} = r^{T-1}\Pi_{k=1}^{T-1}\sin^{k-1}\theta_k$.

Окончательно имеем

$$\mathbf{E}\left\{\frac{1}{\hat{G}_0^n}\right\} \le (2T_0)^{\frac{T_0}{2}}\left(2C_1\pi^{T_0-1}+\delta^{-T_0}\right) = C_2. \ \blacksquare$$

Лемма 3. Пусть для модели (1) выполнено (*) и имеет место одномерный случай $d = 1$. Тогда определитель ковариационной матрицы отсчетов в моменты времени $i_1,...,i_n\in\mathbf{Z},\ \ i_1<\cdots<i_n,\ \ n\in\mathbf{N}\setminus\{1\}$

$$\det\mathbf{Cov}\left\{\left(Y_{i_1},...,Y_{i_n}\right)',\ \left(Y_{i_1},...,Y_{i_n}\right)'\right\} = G_0^n\left(1-B^{2(i_2-i_1)}\right)\cdots\left(1-B^{2(i_n-i_{n-1})}\right).$$

Доказательство. Проводится очевидным образом по индукции. ■

С использованием доказанных результатов построим асимптотическое разложение риска (4).

Теорема 2. Если для модели (1) выполнено (*), то справедливо асимптотическое разложение среднеквадратического риска прогнозирования (4)

$$R_{T,\text{"plug-in"}} = \Sigma + \frac{1}{T}\frac{\Sigma}{1-B^2}\left(\frac{2B^2}{\vartheta_0^2}\sum_{\tau=-\infty}^{+\infty}B^{2|\tau|}\tilde{\vartheta}_{\tau,0,0} - \frac{4B}{\vartheta_0\vartheta_1}\sum_{\tau=-\infty}^{+\infty}B^{|\tau-1|+|\tau|}\tilde{\vartheta}_{\tau,0,1} + \right.$$
$$\left. + \frac{1}{\vartheta_1^2}\sum_{\tau=-\infty}^{+\infty}\left(B^{2|\tau|}+B^{|\tau+1|+|\tau-1|}\right)\tilde{\vartheta}_{\tau,1,1}\right) + o\left(\frac{1}{T}\right). \qquad (5)$$

Доказательство. Преобразуем риск прогнозирования:

$$R_{T,\text{"plug-in"}} = \mathbf{E}\left\{\left(\hat{B}Y_T - BY_T - U_{T+1}\right)^2\right\} = \Sigma + \mathbf{E}\left\{\left(\hat{B}-B\right)^2 Y_T^2\right\} =$$
$$= \Sigma + \mathbf{E}\left\{\left(\frac{\hat{G}_1}{\hat{G}_0}-\frac{G_1}{G_0}\right)^2 Y_T^2\right\} = \Sigma + \mathbf{E}\left\{f\left(\hat{G}_0,\hat{G}_1\right)Y_T^2\right\},$$
$$f(x,y) = \left(\frac{y}{x}-\frac{y_0}{x_0}\right)^2,\ \ x=\hat{G}_0\ge 0,\ \ y=\hat{G}_1,\ \ x_0=G_0>0,\ \ y_0=G_1.$$

Воспользуемся формулой Тейлора с остаточным членом в форме Лагранжа:

$$f(x,y) = f(x_0,y_0) + \frac{\partial f(x_0,y_0)}{\partial x}(x-x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y-y_0) +$$
$$+\frac{\partial^2 f(x_0,y_0)}{2!\partial x^2}(x-x_0)^2 + \frac{2\partial^2 f(x_0,y_0)}{2!\partial x\partial y}(x-x_0)(y-y_0) + \frac{\partial^2 f(x_0,y_0)}{2!\partial y^2}(y-y_0)^2 +$$
$$+\frac{\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{3!\partial x^3}(x-x_0)^3 +$$
$$+\frac{3\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{3!\partial x^2\partial y}(x-x_0)^2(y-y_0) +$$
$$+\frac{3\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{3!\partial x\partial y^2}(x-x_0)(y-y_0)^2 +$$
$$+\frac{\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{3!\partial y^3}(y-y_0)^3,\ \ \theta\in(0,\ 1).$$

Вычислим все с.

$$f(x_0,y_0)=0,\ \ \frac{\partial f(x_0,y_0)}{\partial x}=0,\ \ \frac{\partial f(x_0,y_0)}{\partial y}=0,$$

$$\frac{\partial^2 f(x_0,y_0)}{\partial x^2}=2\frac{y_0^2}{x_0^4},\quad \frac{\partial^2 f(x_0,y_0)}{\partial x\partial y}=-2\frac{y_0}{x_0^3},\quad \frac{\partial^2 f(x_0,y_0)}{\partial y^2}=\frac{2}{x_0^2},$$

$$\frac{\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{\partial x^3}=$$

$$=-12\frac{\left(y_0+\theta(y-y_0)\right)^2}{\left(x_0+\theta(x-x_0)\right)^5}-12\left(\frac{y_0+\theta(y-y_0)}{x_0+\theta(x-x_0)}-\frac{y_0}{x_0}\right)\frac{y_0+\theta(y-y_0)}{\left(x_0+\theta(x-x_0)\right)^4},$$

$$\frac{\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{\partial x^2\partial y}=$$

$$=8\frac{y_0+\theta(y-y_0)}{\left(x_0+\theta(x-x_0)\right)^4}+4\left(\frac{y_0+\theta(y-y_0)}{x_0+\theta(x-x_0)}-\frac{y_0}{x_0}\right)\frac{1}{\left(x_0+\theta(x-x_0)\right)^3},$$

$$\frac{\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{\partial x\partial y^2}=-\frac{4}{\left(x_0+\theta(x-x_0)\right)^3},$$

$$\frac{\partial^3 f\left(x_0+\theta(x-x_0),y_0+\theta(y-y_0)\right)}{\partial y^3}=0.$$

Получаем выражение для риска: $R_{T,\text{"plug-in"}}=\Sigma+\mathbf{E}\left\{\bar{M}_T+\bar{R}_T\right\}$, где

$$\bar{M}_T=\frac{B^2}{G_0^2}\left(\hat{G}_0-G_0\right)^2Y_T^2-\frac{2B}{G_0^2}\left(\hat{G}_0-G_0\right)\left(\hat{G}_1-G_1\right)Y_T^2+\frac{1}{G_0^2}\left(\hat{G}_1-G_1\right)^2Y_T^2,$$

$$\bar{R}_T=\left(-2\frac{\left(G_1+\theta\left(\hat{G}_1-G_1\right)\right)^2}{\left(G_0+\theta\left(\hat{G}_0-G_0\right)\right)^5}-2\left(\frac{G_1+\theta\left(\hat{G}_1-G_1\right)}{G_0+\theta\left(\hat{G}_0-G_0\right)}-\frac{G_1}{G_0}\right)\frac{G_1+\theta\left(\hat{G}_1-G_1\right)}{\left(G_0+\theta\left(\hat{G}_0-G_0\right)\right)^4}\right)\left(\hat{G}_0-G_0\right)^3Y_T^2+$$

$$+\left(4\frac{G_1+\theta\left(\hat{G}_1-G_1\right)}{\left(G_0+\theta\left(\hat{G}_0-G_0\right)\right)^4}+2\left(\frac{G_1+\theta\left(\hat{G}_1-G_1\right)}{G_0+\theta\left(\hat{G}_0-G_0\right)}-\frac{G_1}{G_0}\right)\frac{1}{\left(G_0+\theta\left(\hat{G}_0-G_0\right)\right)^3}\right)\left(\hat{G}_0-G_0\right)^2\left(\hat{G}_1-G_1\right)Y_T^2-$$

$$-\frac{2}{\left(G_0+\theta\left(\hat{G}_0-G_0\right)\right)^3}\left(\hat{G}_0-G_0\right)\left(\hat{G}_1-G_1\right)^2Y_T^2.$$

Используя формулу для математического ожидания произведения компонент нормально распределенного случайного вектора $L\{Z\}=L\left\{\left(Z_1,\dots,Z_6\right)'\right\}=$ $=N_6\left(0_6,\ \tilde{\Sigma}\right)$, $\tilde{\Sigma}=\left(\tilde{\Sigma}\right)_{ij}$:

$$\mathbf{E}\left\{Z_1Z_2Z_3Z_4Z_5Z_6\right\}=\tilde{\Sigma}_{12}\tilde{\Sigma}_{34}\tilde{\Sigma}_{56}+\tilde{\Sigma}_{12}\tilde{\Sigma}_{35}\tilde{\Sigma}_{46}+\tilde{\Sigma}_{12}\tilde{\Sigma}_{36}\tilde{\Sigma}_{45}+$$
$$+\tilde{\Sigma}_{13}\tilde{\Sigma}_{24}\tilde{\Sigma}_{56}+\tilde{\Sigma}_{13}\tilde{\Sigma}_{25}\tilde{\Sigma}_{46}+\tilde{\Sigma}_{13}\tilde{\Sigma}_{26}\tilde{\Sigma}_{45}+\tilde{\Sigma}_{14}\tilde{\Sigma}_{23}\tilde{\Sigma}_{56}+\tilde{\Sigma}_{14}\tilde{\Sigma}_{25}\tilde{\Sigma}_{36}+\tilde{\Sigma}_{14}\tilde{\Sigma}_{26}\tilde{\Sigma}_{35}+$$
$$+\tilde{\Sigma}_{15}\tilde{\Sigma}_{23}\tilde{\Sigma}_{46}+\tilde{\Sigma}_{15}\tilde{\Sigma}_{24}\tilde{\Sigma}_{36}+\tilde{\Sigma}_{15}\tilde{\Sigma}_{26}\tilde{\Sigma}_{34}+\tilde{\Sigma}_{16}\tilde{\Sigma}_{23}\tilde{\Sigma}_{45}+\tilde{\Sigma}_{16}\tilde{\Sigma}_{24}\tilde{\Sigma}_{35}+\tilde{\Sigma}_{16}\tilde{\Sigma}_{25}\tilde{\Sigma}_{34},$$

предположение ($*$) и лемму 1, получаем асимптотику $\bar{M}_T$:

$$\mathbf{E}\left\{\bar{M}_T\right\}T\to\frac{\Sigma}{1-B^2}\left(\frac{2B^2}{\vartheta_0^2}\sum_{\tau=-\infty}^{+\infty}B^{2|\tau|}\tilde{\vartheta}_{\tau,0,0}-\frac{4B}{\vartheta_0\vartheta_1}\sum_{\tau=-\infty}^{+\infty}B^{|\tau-1|+|\tau|}\tilde{\vartheta}_{\tau,0,1}+\right.$$
$$\left.+\frac{1}{\vartheta_1^2}\sum_{\tau=-\infty}^{+\infty}\left(B^{2|\tau|}+B^{|\tau+1|+|\tau-1|}\right)\tilde{\vartheta}_{\tau,1,1}\right).$$

Используя соотношение $\frac{1}{G_0+\theta\left(\hat{G}_0-G_0\right)} \le \frac{1}{G_0}+\frac{1}{\hat{G}_0}$, $\theta \in (0,\ 1)$, теорему 1, леммы 2 и 3, получаем асимптотику $\overline{R}_T$: $\mathbf{E}\left\{\overline{R}_T\right\} T \to 0$.

Следствие. Пусть выполнены условия теоремы 2 и количество пропусков ограничено: $\exists T_{\text{obs}} \in \mathbf{N}$, что $O_t = 1$, $t \ge T_{\text{obs}}$. Тогда справедливо асимптотическое разложение риска «подстановочного» прогнозирования

$$R_{T,\text{"plug-in"}} = \Sigma + \frac{\Sigma}{T} + o\left(\frac{1}{T}\right).$$

Доказательство. Из теоремы 2 следует, что риск прогнозирования

$$R_{T,\text{"plug-in"}} = \Sigma + \frac{\Sigma}{T} C(B) + o\left(\frac{1}{T}\right),$$

где $C(B)$ - константа, зависящая от B. Используя условие об ограниченности пропусков и преобразовывая выражение для риска (5), находим, что $C(B) = 1$. ■

Результаты этих исследований были частично поддержаны проектом ГКПНИ «Инфотех».

1. Андерсон Т. Статистический анализ временных рядов. М., 1976.
2. Box G.E.P., Jenkins G.M. Time Scries Analysis: Forecasting and Control. San Francisco, 1976.
3. Little R.J.A., Rubin D. B. Statistical Analysis with Missing Data. New York, 2002.
4. Schafcr J.L. Analysis of Incomplete Multivariate Data. London, 1977.
5 Kharin Yu.S., Huryn A. S .//Austrian Journal of Statistics. 2005. Vol. 34. №2. P. 163.
6. Харин Ю.С. Робастность в статистическом распознавании образов. Мн., 1992.
7. Kharin Yu.S. Robustness in Statistical Pattern Recognition. Dordrecht; Boston; London, 1996.

Поступила в редакцию 14.06.06.

Александр Сергеевич Гурин - аспирант кафедры математического моделирования и анализа данных. Научный руководитель - член-корреспондент НАН Беларуси, доктор физико-математических наук, профессор Ю.С. Харин.