СТАТИСТИЧЕСКОЕ ОТНЕСЕНИЕ РЕАЛИЗАЦИЙ СТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ К ЗАДАННЫМ АВТОРЕГРЕССИОННЫМ МОДЕЛЯМ

Е. Е. Жук

Белорусский государственный университет
Минск, Беларусь
E-mail: zhukee@mail.ru

Исследуется проблема статистического отнесения реализаций стационарных в широком смысле временных рядов к заданным классам на основе их описания авторегрессионными моделями. Предлагается использовать решающее правило в пространстве коэффициентов авторегрессии. В качестве меры эффективности принимаемых решений используется риск (вероятность ошибочного отнесения), который вычислен аналитически в асимптотике растущих длительностей реализаций. Рассмотрен случай двух классов.

Ключевые слова: стационарный временной ряд, разложение Вольда, авторегрессионные модели, классы, реализация, отнесение, решающее правило, риск.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И ПОСТАНОВКА ЗАДАЧИ

Пусть $\{x_t\}_{t\in Z}$ – стационарный в широком смысле временной ряд (ВР) [1, 5] ($Z=\{0,\pm1,\pm2,\ldots\}$ – множество целых чисел), отсчеты $x_t\in R$, $t\in Z$, которого имеют нулевое математическое ожидание:

$$\mathrm{E}\{x_t\}=0,\ \ \forall t\in Z, \tag{1}$$

и ковариационную функцию:

$$\sigma(\tau)=\mathrm{Cov}\{x_t,x_{t+\tau}\}=\mathrm{E}\{x_t x_{t+\tau}\}\quad \left(\mathrm{D}\{x_t\}=\sigma(0)<+\infty\right),\ \ \forall t,\tau\in Z. \tag{2}$$

Как известно [1, 5], такой ВР, с учетом (1), при весьма общих условиях регулярности можно представить в виде так называемого разложения Вольда:

$$x_t+\sum_{j=1}^{+\infty}\beta_j x_{t-j}=u_t,\ \ t\in Z;\qquad \sum_{j=1}^{+\infty}\beta_j^2<+\infty, \tag{3}$$

где случайные величины $\{u_t\}_{t\in Z}$ некоррелированы и имеют нулевые математические ожидания и одинаковую ограниченную дисперсию:

$$\mathrm{E}\{u_t\}=0,\ \ \mathrm{D}\{u_t\}=\mathrm{E}\{u_t^2\}=\sigma^2<+\infty; \tag{4}$$
$$\mathrm{E}\{u_t u_l\}=0,\ \ \forall t,l\in Z,\ \ l\neq t.$$

Разложение (3), (4) является авторегрессией бесконечного порядка [1, 5] и однозначно определяется коэффициентами авторегрессии $\beta=(\beta_j)_{j=1}^{+\infty}$ и дисперсией σ^2 случайных величин (ошибок наблюдений [1, 5]) $\{u_t\}_{t\in Z}$, что обозначим как АР($+\infty,\beta,\sigma^2$).

Пусть теперь заданы $L \geq 2$ классов $\{\Omega_i\}_{i\in S}$ стационарных в широком смысле временных рядов ($S = \{1,\ldots,L\}$ – множество номеров классов). Класс Ω_i описывается моделью типа (3), (4) АР($+\infty, \beta^{(i)}, \sigma^2_{(i)}$), которая задается своими значениями коэффициентов авторегрессии $\beta^{(i)} = (\beta^{(i)}_j)_{j=1}^{+\infty}$ и дисперсии ошибок наблюдений $\sigma^2_{(i)}$ ($i \in S$).

Наблюдается реализация $X = \{x_t\}_{t=1}^{T}$ длительности T произвольного стационарного ВР, описанного выше моделью АР($+\infty, \beta, \sigma^2$), параметры $\beta = (\beta_j)_{j=1}^{+\infty}$ и σ^2 которой неизвестны. Задача заключается в отнесении [3, 4] данного временного ряда к одному из классов $\{\Omega_i\}_{i\in S}$ (к одной из авторегрессионных моделей АР($+\infty, \beta^{(i)}, \sigma^2_{(i)}$), $i \in S$) по имеющейся реализации $X = \{x_t\}_{t=1}^{T}$. При этом необходимо определить принцип (критерий), по которому будет производиться отнесение, построить решающее правило (РП) и оценить его эффективность.

РЕШАЮЩЕЕ ПРАВИЛО В ПРОСТРАНСТВЕ КОЭФФИЦИЕНТОВ АВТОРЕГРЕССИИ И ЕГО РИСК

Отметим сразу, что при описании классов $\{\Omega_i\}_{i\in S}$ моделями АР($+\infty, \beta^{(i)}, \sigma^2_{(i)}$), $i \in S$, задать дисперсии ошибок $\{\sigma^2_{(i)}\}_{i\in S}$ проблематично. Более того, если наблюдения производятся в одних и тех же условиях, то чаще всего $\sigma^2_{(i)} = \sigma^2$, $i \in S$. Поэтому далее при принятии решений дисперсии ошибок учитывать не будем.

Из (3) также следует, что в модели АР($+\infty, \beta, \sigma^2$) для коэффициентов авторегрессии $\beta = (\beta_j)_{j=1}^{+\infty}$ выполняется: $\beta_j \to 0$, $j \to +\infty$. Что позволяет, выбрав “достаточно большой” порядок авторегрессии p, считать что в $\beta = (\beta_j)_{j=1}^{+\infty}$: $\beta_j = 0$, $j > p$, и временной ряд $\{x_t\}_{t\in Z}$ описывается моделью АР($p, \beta_{(p)}, \sigma^2$) [1, 2, 5]:

$$x_t + \sum_{j=1}^{p} \beta_j x_{t-j} = u_t, \quad t \in Z, \tag{5}$$

где $\beta_{(p)} = (\beta_1,\ldots,\beta_p)' \in R^p$ – первые p коэффициентов авторегрессии из $\beta = (\beta_j)_{j=1}^{+\infty}$ (“′” – символ транспонирования), а случайные величины $\{u_t\}_{t\in Z}$ определены в (4).

Сделанное выше предположение позволяет оценить первые p коэффициентов авторегрессии из $\beta = (\beta_j)_{j=1}^{+\infty}$, построив по реализации $X = \{x_t\}_{t=1}^{T}$ длительности $T > p$ МНК-оценку [5] $\hat{\beta}_{(p)} = (\hat{\beta}_1,\ldots,\hat{\beta}_p)' \in R^p$:

$$\hat{\beta}_{(p)} = -\left(\sum_{t=p+1}^{T} X_t X_t'\right)^{-1} \sum_{t=p+1}^{T} x_t X_t, \tag{6}$$

где обозначено: $X_t = (x_{t-1},\ldots,x_{t-p})' \in R^p$, $t = \overline{p+1,T}$.

По аналогии с [2] для отнесения реализации X к одному из классов $\{\Omega_i\}_{i\in S}$ можно предложить следующее РП:

$$d = d(X) = \arg\min_{i\in S}\left|\hat{\beta} - \beta^{(i)}\right| = \arg\min_{i\in S}\left|\hat{\beta} - \beta^{(i)}\right|^2 = \arg\min_{i\in S}\left\{\left|\hat{\beta}_{(p)} - \beta^{(i)}_{(p)}\right|^2 + \sum_{j=p+1}^{+\infty}\left(\beta^{(i)}_j\right)^2\right\}, \tag{7}$$

где $\hat{\beta} := (\hat{\beta}'_{(p)}, 0, \ldots, 0, \ldots)'$, $\beta^{(i)}_{(p)} = (\beta^{(i)}_1, \ldots, \beta^{(i)}_p)'$ ($i \in S$), а $|y| = \sqrt{y'y} = \sqrt{\Sigma_{j=1}^N y_j^2}$ – евклидова норма вектора $y = (y_j)_{j=1}^N \in R^N$ ($N \le +\infty$).

Содержательный смысл РП (7) прост: оно относит временной ряд $\{x_t\}_{t\in Z}$ по его реализации $X = \{x_t\}_{t=1}^T$ к тому классу с номером $d(X) \in S$, к коэффициентам авторегрессии которого ближе всего в метрике Евклида оценка коэффициентов (6), построенная по X.

Отметим также, что если авторегрессионные модели АР($+\infty, \beta^{(i)}, \sigma^2_{(i)}$), $i \in S$, определяющие классы $\{\Omega_i\}_{i\in S}$, таковы что $\beta^{(i)} = ((\beta^{(i)}_{(p)})', 0, \ldots, 0, \ldots)'$, $i \in S$, то РП (7) упрощается и принимает вид:

$$d = d(X) = \arg\min_{i\in S} \left| \hat{\beta}_{(p)} - \beta^{(i)}_{(p)} \right|. \tag{8}$$

В качестве меры эффективности РП (7) (или РП (8)) будем, как и в [3, 4], использовать обобщение традиционного риска [2, 5]:

$$r_T = \mathrm{P}\{d(X) \notin D^o\}, \quad D^o = \{k : \left|\beta - \beta^{(k)}\right| = \min_{i\in S}\left|\beta - \beta^{(i)}\right|\}. \tag{9}$$

Здесь $D^o \subseteq S$ – множество номеров тех классов, к которым ВР $\{x_t\}_{t\in Z}$ из (3), (4) ближе всего по коэффициентам авторегрессии в смысле расстояния между ними в метрике Евклида (учтено, что могут быть совпадающие по значению расстояния). Сам риск r_T имеет смысл вероятности ($0 \le r_T \le 1$) не отнести временной ряд по его реализации $X = \{x_t\}_{t=1}^T$ длительности T к одному их ближайших классов. Чем меньше значение риска (9), тем эффективнее РП (7).

Отметим также, что если $D^o = S$, то $r_T = 0$ и выносимое РП (7) решение не принципиально. Если множество D^o состоит из одного элемента (один ближайший класс), то риск (9) принимает вид:

$$r_T = \mathrm{P}\{d(X) \ne d^o\}, \quad d^o = \arg\min_{i\in S}\left|\beta - \beta^{(i)}\right|. \tag{10}$$

АСИМПТОТИЧЕСКОЕ ВЫЧИСЛЕНИЕ РИСКА

Исследуем асимптотическое поведение риска (9) РП (7) в условиях “достаточно большой” длительности T подлежащей отнесению реализации $X = \{x_t\}_{t=1}^T$:

$$T \to +\infty. \tag{11}$$

Приведем сначала вспомогательный результат, основанный на известных свойствах оценки (6) [1, 5]. Введем обозначения:

$$n_N(y \mid \mu, \Sigma) = (2\pi)^{-\frac{N}{2}} \, |\det(\Sigma)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y-\mu)'\,\Sigma^{-1}(y-\mu)\right), \quad y \in R^N,$$

– плотность N-мерного нормального (гауссовского) закона распределения вероятностей (закон $N_N(\mu, \Sigma)$ [5]) с вектором математического ожидания $\mu \in R^N$ и невырожденной ковариационной $(N \times N)$-матрицей Σ ($|\det(\Sigma)| \ne 0$).

Лемма. Пусть $X = \{x_t\}_{t=1}^T$ – реализация длительности T временного ряда, определяемого моделью АР($p, \beta_{(p)}, \sigma^2$) из (5), где случайные величины $\{u_t\}_{t\in Z}$ независимы в совокупности и одинаково распределены по нормальному закону $N_1(0,\sigma^2)$ $(\sigma^2 < +\infty)$, а корни характеристического уравнения

$$\sum_{j=1}^{p}\beta_j z^{p-j} = 0$$

лежат внутри единичного круга ($|z|<1$), тогда $\hat{\beta}_{(p)}$ из (6) состоятельная по вероятности оценка для $\beta_{(p)} = (\beta_1,\ldots,\beta_p)' \in R^p$: $\hat{\beta}_{(p)} \xrightarrow{\text{P}} \beta_{(p)}$, $T \to +\infty$, и в асимптотике (11) она также и нормально распределена:

$$L\left\{\sqrt{T}(\hat{\beta}_{(p)} - \beta_{(p)})\right\} \to N_p(0_p, \sigma^2\Sigma_{p,p}^{-1}), \tag{12}$$

где 0_p – нулевой p-вектор, а $\Sigma_{p,p} = (\sigma(i-j))_{i,j=1}^{p}$ – ковариационная $(p\times p)$-матрица, элементы которой определяются ковариационной функцией (2), вычисляемой по $\beta_{(p)}$ из уравнений Юла – Уокера ($\Sigma_{p,p} = \Sigma_{p,p}(\beta_{(p)})$):

$$\sigma(\tau) + \sum_{j=1}^{p}\beta_j\sigma(\tau - j) = 0, \quad \tau = \overline{1, p}.$$

Рассмотрим случай, когда у ВР $\{x_t\}_{t\in Z}$ один ближайший класс среди $\{\Omega_i\}_{i\in S}$:

$$\exists d^o: \left|\beta - \beta^{(d^o)}\right| < \left|\beta - \beta^{(i)}\right|, \quad \forall i \in S, \quad i \neq d^o, \tag{13}$$

чему соответствует риск в виде (10).

Теорема. Пусть в условиях леммы выполняется (13), тогда риск r_T РП (7) в асимптотике (11) удовлетворяет соотношению:

$$r_T / \tilde{r}_T \to 1, \quad T \to +\infty; \tag{14}$$

$$\tilde{r}_T = 1 - \int_{R^p}\prod_{\substack{j\in S\\ j\neq d^o}} U\left(\left|y - \beta_{(p)}^{(j)}\right|^2 - \left|y - \beta_{(p)}^{(d^o)}\right|^2 + \sum_{l=p+1}^{+\infty}\left((\beta_l^{(j)})^2 - (\beta_l^{(d^o)})^2\right)\right) n_p\left(y \middle| \beta_{(p)}, \frac{1}{T}\sigma^2\Sigma_{p,p}^{-1}\right)dy,$$

где $U(z) = \{1, если\, z \geq 0; 0, если\, z < 0\}$ – единичная функция Хэвисайда, а

$$d^o = \arg\min_{i\in S}\left\{\left|\beta_{(p)} - \beta_{(p)}^{(i)}\right|^2 + \sum_{l=p+1}^{+\infty}(\beta_l^{(i)})^2\right\} \tag{15}$$

– номер ближайшего класса из (10).

Доказательство. Для риска r_T из (10) с учетом вида РП (7) справедлива цепочка равенств:

$$r_T = \text{P}\{d(X) \neq d^o\} = 1 - \text{P}\{d(X) = d^o\} = 1 - \text{P}\left\{\bigcap_{\substack{j\in S\\ j\neq d^o}}\left\{\left|\hat{\beta} - \beta^{(j)}\right|^2 \geq \left|\hat{\beta} - \beta^{(d^o)}\right|^2\right\}\right\} =$$

$$= 1 - \text{P}\left\{\bigcap_{\substack{j\in S\\ j\neq d^o}}\left\{\left|\hat{\beta}_{(p)} - \beta_{(p)}^{(j)}\right|^2 - \left|\hat{\beta}_{(p)} - \beta_{(p)}^{(d^o)}\right|^2 + \sum_{l=p+1}^{+\infty}\left((\beta_l^{(j)})^2 - (\beta_l^{(d^o)})^2\right) \geq 0\right\}\right\},$$

откуда с учетом (12) и получаем (14). Соотношение (15) для d^o очевидно, поскольку в условиях леммы: $\beta_j = 0$, $j > p$, и из (10):

$$d^o = \arg\min_{i\in S} \left|\beta - \beta^{(i)}\right| = \arg\min_{i\in S} \left|\beta - \beta^{(i)}\right|^2 = \arg\min_{i\in S} \left\{ \left|\beta_{(p)} - \beta_{(p)}^{(i)}\right|^2 + \sum_{l=p+1}^{+\infty} (\beta_l^{(i)})^2 \right\}.$$

СЛУЧАЙ ДВУХ КЛАССОВ

Пусть теперь число классов равно двум ($L = 2$, $S = \{1,2\}$). Не ограничивая общности, будем считать, что порядок авторегрессии p подлежащего отнесению ВР $\{x_t\}_{t\in Z}$ и коэффициенты авторегрессионных моделей, определяющих классы, таковы что

$$\beta_j^{(i)} = 0, \quad j > p, \quad i \in S. \tag{16}$$

Следствие. Пусть в условиях теоремы $L = 2$ и имеет место (16), тогда в условиях асимптотики (11) ($T \to +\infty$):

$$\frac{r_T}{\tilde{r}_T} \to 1, \quad \tilde{r}_T = \Phi\left(-\sqrt{T}\, \frac{\left| \left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|^2 - \left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|^2 \right|}{2\sigma\Delta} \right), \tag{17}$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{w^2}{2}\right) dw$, $z \in R$ – функция распределения вероятностей стандартного нормального закона $N_1(0,1)$, а величина

$$\Delta = \sqrt{(\beta_{(p)}^{(1)} - \beta_{(p)}^{(2)})' \Sigma_{p,p}^{-1} (\beta_{(p)}^{(1)} - \beta_{(p)}^{(2)})} \tag{18}$$

– аналог межклассового расстояния Махаланобиса [5].

Доказательство. С учетом (16) и того, что $L = 2$, из (14), (15) получаем:

$$\tilde{r}_T = \begin{cases} 1 - \mathrm{P}\left\{ \left|\xi - \beta_{(p)}^{(2)}\right|^2 - \left|\xi - \beta_{(p)}^{(1)}\right|^2 \geq 0 \right\}, \text{ если } d^o = 1; \\ \mathrm{P}\left\{ \left|\xi - \beta_{(p)}^{(2)}\right|^2 - \left|\xi - \beta_{(p)}^{(1)}\right|^2 \geq 0 \right\}, \text{ если } d^o = 2, \end{cases} \tag{19}$$

где случайный p-вектор $\xi \in R^p$ распределен по нормальному закону $N_p(\beta_{(p)}, \sigma^2 \Sigma_{p,p}^{-1} / T)$.

Преобразуем левую часть неравенства из (19):

$$\left|\xi - \beta_{(p)}^{(2)}\right|^2 - \left|\xi - \beta_{(p)}^{(1)}\right|^2 = 2\left(\xi - \frac{\beta_{(p)}^{(1)} + \beta_{(p)}^{(2)}}{2} \right) \left(\beta_{(p)}^{(1)} - \beta_{(p)}^{(2)}\right) \in R \tag{20}$$

– случайная величина, линейная по ξ, и потому являющаяся нормальной случайной величиной со следующими математическим ожиданием и дисперсией:

$$\mathrm{E}\{\xi\} = \left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|^2 - \left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|^2, \quad \mathrm{D}\{\xi\} = 4\sigma^2 (\beta_{(p)}^{(1)} - \beta_{(p)}^{(2)})' \Sigma_{p,p}^{-1} (\beta_{(p)}^{(1)} - \beta_{(p)}^{(2)}) / T = 4\sigma^2 \Delta^2 / T,$$

где Δ – величина из (18).

Из (19) с учетом нормировки случайной величины (20) до стандартного нормального закона $N_1(0,1)$ получаем:

$$\tilde{r}_T = \begin{cases} \Phi\left(-\sqrt{T}\,\dfrac{\left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|^2 - \left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|^2}{2\sigma\Delta} \right), \text{ если } d^o = 1; \\ \Phi\left(\sqrt{T}\,\dfrac{\left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|^2 - \left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|^2}{2\sigma\Delta} \right), \text{ если } d^o = 2, \end{cases} \quad (21)$$

где учтено известное свойство функции распределения вероятностей стандартного нормального закона: $1 - \Phi(z) = \Phi(-z)$, $z \in R$.

Но, исходя из (15), с учетом условия (16):

$$d^o = \begin{cases} 1, \text{ если } \left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|^2 < \left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|^2; \\ 2, \text{ если } \left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|^2 > \left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|^2, \end{cases}$$

что позволяет записать $\tilde{r}_T$ из (21) в виде выражения из (17).

Практическая значимость полученных асимптотических соотношений (14), (17) состоит в том, что они при больших длительностях реализаций ($T \to +\infty$) позволяют приближенно вычислить риск: $r_T \approx \tilde{r}_T$. Из (17) также следует, что $r_T \to 0$, $T \to +\infty$, то есть с увеличением длительности подлежащей отнесению реализации эффективность принимаемых решений повышается (значение риска уменьшается).

Проанализировав дополнительно поведение $\tilde{r}_T$ из (17) относительно расстояний $\left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|$, $\left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|$ и расстояния Δ между $\beta_{(p)}^{(1)}$ и $\beta_{(p)}^{(2)}$ из (18), заключаем, что риск уменьшается с ростом Δ и увеличением различия между собой расстояний $\left|\beta_{(p)} - \beta_{(p)}^{(1)}\right|$ и $\left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|$ между коэффициентами авторегрессии $\beta_{(p)}$ подлежащего отнесению ВР и коэффициентами $\beta_{(p)}^{(1)}$ и $\beta_{(p)}^{(2)}$ авторегрессионных моделей, задающих классы.

Отметим, что при $\left|\beta_{(p)} - \beta_{(p)}^{(1)}\right| = \left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|$ риск заведомо равен нулю: $r_T = 0$, $\forall T$, а соотношение (17) приводит к неточному результату, поскольку получено в предположении (13): $\left|\beta_{(p)} - \beta_{(p)}^{(1)}\right| \neq \left|\beta_{(p)} - \beta_{(p)}^{(2)}\right|$.

ЛИТЕРАТУРА

1. *Андерсон, Т.* Статистический анализ временных рядов / Т. Андерсон. М. : Мир, 1976. 755 с.

2. *Жук, Е. Е.* Кластер-анализ реализаций временных рядов авторегрессии / Е. Е. Жук // Автоматика и телемеханика. 2003. № 1. С. 74–85.

3. *Жук, Е. Е.* Отнесение многомерных наблюдений к заданным классам методом максимального правдоподобия и его эффективность / Е. Е. Жук // Весцi НАН Беларусi. Сер. фiз.-мат. навук. 2012. № 4. С. 37–41.

4. *Жук, Е. Е.* Статистическое определение ближайших классов по многомерным обучающим выборкам / Е. Е. Жук // Теория вероятностей, случайные процессы, математическая статистика и их приложения. Мн., 2014. С. 59–63.

5. *Харин, Ю. С.* Математическая и прикладная статистика / Ю. С. Харин, Е. Е. Жук. Мн. : БГУ, 2005. 279 с.