

А.П. Сарычев

КЛАССИФИКАЦИЯ СОСТОЯНИЙ ОБЪЕКТА, ФУНКЦИОНИРОВАНИЕ КОТОРОГО ОПИСЫВАЕТСЯ ОДНОМЕРНОЙ АВТОРЕГРЕССИЕЙ

Рассмотрена задача классификации состояний динамического объекта, который может находиться в двух классах состояний. Функционирование объекта в классах описывается различными авторегрессионными моделями. Построено правило классификации и исследованы его свойства.

Введение. Задача обнаружения изменения свойств динамических объектов часто возникает в таких областях, как техническая и медицинская диагностика, контроль технологических процессов, мониторинг, обработка сигналов [1]. Распространённым классом моделей, описывающих функционирование стохастических динамических объектов, является класс одномерных авторегрессионных уравнений. Примерами решения задач обнаружения для этого класса моделей могут служить работы [2]–[3]. Так, в [2] рассмотрена задача обнаружения перехода объекта из одного класса состояний в другой класс, в каждом из которых объект описывается своей одномерной авторегрессионной моделью с известными параметрами. В [3] задача обнаружения рассматривается для одномерной авторегрессионной модели совместно с задачей структурно-параметрической идентификации, т. е. определением наилучшей структуры в заданном классе моделей. Примером успешного решения задачи обнаружения на основе регрессионных уравнений является работа [4], где предложен и обоснован поход к построению математических моделей контроля технического состояния силовых и энергетических установок в длительной эксплуатации.

Основная задача данной работы – обнаружение изменения свойств динамического объекта, который может находиться в двух клас-

сах состояний, и в каждом из них поведение объекта описывается своей авторегрессионной моделью с априорно неизвестными параметрами. В [5] для определения коэффициентов авторегрессионной модели построена итерационная схема оценивания, эффективность которой подтверждена методом статистических испытаний. Опираясь на этот результат, можно сформулировать следующую задачу статистической классификации: на основе обучающих выборок наблюдений двух классов состояний, каждый из которых описывается своей авторегрессией, требуется построить решающее правило, устанавливающее принадлежность анализируемых наблюдений к одному из двух классов.

1. Априорные предположения о динамическом объекте

Пусть функционирование динамического объекта подчиняется закону в виде авторегрессионного уравнения

$$\begin{pmatrix} * \\ x_1 \\ * \\ x_2 \\ \vdots \\ * \\ x_i \\ \vdots \\ * \\ x_n \end{pmatrix} = \begin{bmatrix} * & * & \cdots & * \\ x_0 & x_{-1} & \cdots & x_{1-p} \\ * & * & \cdots & * \\ x_1 & x_0 & \cdots & x_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \\ x_{i-1} & x_{i-2} & \cdots & x_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \\ x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{bmatrix} \begin{pmatrix} 0 \\ \theta_1 \\ 0 \\ \theta_2 \\ \vdots \\ 0 \\ \theta_p \end{pmatrix} + \begin{pmatrix} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_{i-1} \\ \vdots \\ \zeta_{n-1} \end{pmatrix}, \quad (1)$$

которое в матрично-векторной форме можно записать в виде:

$$\overset{*}{\mathbf{x}} = \overset{*}{\mathbf{Z}}(p) \overset{0}{\boldsymbol{\theta}} + \zeta(-1), \quad (2)$$

где $\overset{*}{\mathbf{x}}$ – ненаблюдаемый $(n \times 1)$ -вектор значений выходной переменной в дискретные моменты времени $t = t_i, i = 1, 2, \dots, n$; n – общее число наблюдений; p – число предыдущих значений переменной, которые влияют

на её текущее значение; $\overset{*}{\mathbf{Z}}(p)$ – $(n \times p)$ -матрица p предыдущих ненаблюдаемых значений, в обозначении этой матрицы p означает, что в (1)–(2)

при формировании величины x_i участвуют величины $(x_{i-1}, x_{i-2}, \dots, x_{i-p})$;

$\overset{0}{\boldsymbol{\theta}}$ – $(p \times 1)$ -вектор неизвестных детерминированных коэффициентов;

$\zeta(-1)$ – ненаблюдаемый случайный $(n \times 1)$ -вектор, в его обозначении ”-1”

означає, що в (1)–(2) при формуванні величини x_i аддитивно учас-
тує величина ζ_{i-1} .

В (1)–(2) передполагается, что в формуванні текущего значения
выходной переменной участвуют все p её предыдущих значений. В об-
щем случае не все они могут участвовать в этом формуванні. Для за-
писи моделей в общем случае введем структурные матрицы, смысл кото-
рых покажем на примере. Пусть на текущее значение выходной пере-
менной влияют первое, второе и третье предыдущие её значения из за-
данного числа влияющих предыдущих значений $p = 5$. Тогда вместо

матрицы $\mathbf{Z}^*(p)$ в (1)–(2) следует записать

$$\mathbf{Z}^*(p)\mathbf{S} = \begin{bmatrix} * & * & * & * & * \\ x_0 & x_{-1} & x_{-2} & x_{-3} & x_{-4} \\ * & * & * & * & * \\ x_1 & x_0 & x_{-1} & x_{-2} & x_{-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & * & * & * \\ x_{i-1} & x_{i-2} & x_{i-3} & x_{i-4} & x_{i-5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & * & * & * \\ x_{n-1} & x_{n-2} & x_{n-3} & x_{n-4} & x_{n-5} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} * & * & * \\ x_0 & x_{-1} & x_{-2} \\ * & * & * \\ x_1 & x_0 & x_{-1} \\ \vdots & \vdots & \vdots \\ * & * & * \\ x_{i-1} & x_{i-2} & x_{i-3} \\ \vdots & \vdots & \vdots \\ * & * & * \\ x_{n-1} & x_{n-2} & x_{n-3} \end{bmatrix}, \quad (3)$$

где (5×3) -матрица \mathbf{S} отражает влияние первого, второго и третьего пре-
дыдущих значений переменной на её текущее значение.

Априорная информация о значении p и о том, какие именно пре-
дыдущие значения определяют текущее значение переменной в законе
функционирования объекта (1)–(2), представляется структурной $(p \times m)$ -
матрицей \mathbf{S} (m – число столбцов в матрице \mathbf{S} , которое равно числу не-
известных коэффициентов в модели). Будем предполагать, что эта
структурная матрица задана. С учетом (3) закон (2) для общего случая
формування выходной переменной имеет вид

$$\mathbf{x} = \mathbf{Z}^*(p)\mathbf{S}\overset{\circ}{\boldsymbol{\theta}} + \zeta(-1) = \overset{=}{\mathbf{x}} + \zeta(-1), \quad (4)$$

где $\overset{\circ}{\boldsymbol{\theta}}$ – $(m \times 1)$ -вектор неизвестных детерминированных коэффициентов;
 $\overset{=}{\mathbf{x}}$ – ненаблюдаемая составляющая $(n \times 1)$ -вектора значений выходной
переменной.

Пусть для наблюдений выходной переменной объекта выполняется

$$x_i^* = x_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (5)$$

где x_i – наблюдаемое значение переменной в момент времени $t = t_i$, x_i^* – ненаблюдаемое значение; ε_i – случайная ненаблюдаемая ошибка измерения.

Запишем с учетом (5) модель наблюдения объекта в векторной форме

$$\mathbf{x} = \mathbf{x} + \boldsymbol{\varepsilon}. \quad (6)$$

Пусть относительно $\zeta(-1)$ выполнено

$$E\{\zeta(-1)\} = \mathbf{0}_n, \quad E\{\zeta(-1)\zeta^T(-1)\} = \sigma_\zeta \cdot \mathbf{I}_n, \quad (7)$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям случайного вектора $\zeta(-1)$; $\mathbf{0}_n$ – нулевой $(n \times 1)$ -вектор; σ_ζ – ограниченная дисперсия величины $\zeta_i(-1)$, $i = 1, 2, \dots, n$; \mathbf{I}_n – единичная $(n \times n)$ -матрица.

Пусть относительно $\boldsymbol{\varepsilon}$ выполнено

$$E\{\boldsymbol{\varepsilon}\} = \mathbf{0}_n, \quad E\{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\} = \sigma_\varepsilon \cdot \mathbf{I}_n, \quad (8)$$

где $E\{\cdot\}$ – знак математического ожидания по возможным реализациям случайных векторов $\boldsymbol{\varepsilon}$; σ_ε – ограниченная дисперсия величины ε_i , $i = 1, 2, \dots, n$.

Пусть в результате наблюдения в моменты времени $t = t_i$, $i = 1 - 2p, 2 - 2p, \dots, 0, 1, 2, \dots, n$, получен $(n + 2p)$ -вектор значений переменной

$$(x_{1-2p}, x_{2-2p}, \dots, x_0, x_1, x_2, \dots, x_n)^T = \begin{pmatrix} \mathbf{x}(0) \\ \mathbf{x} \end{pmatrix}, \quad (9)$$

где $(2p \times 1)$ -вектор $\mathbf{x}(0)$ будет использован в качестве начальных условий.

Пусть заданы: 1) p – число предыдущих значений выходной переменной, которые влияют на её текущее значение; 2) структурная $(p \times n)$ -матрица \mathbf{S} , которая показывает, какие именно предыдущие значения переменной определяют её текущее значение.

2. Оценивание коэффициентов в авторегрессионной модели

Для $(n \times 1)$ -вектора \mathbf{x} выполняется

$$\mathbf{x} = \bar{\mathbf{Z}}(p) \mathbf{S} \overset{\circ}{\boldsymbol{\theta}} + \boldsymbol{\xi}, \quad (10)$$

где матрица $\bar{\mathbf{Z}}(p)$ – $(n \times p)$ -матрица ненаблюдаемых значений переменной объекта, по своей структуре она аналогична матрице $\mathbf{Z}^*(p)$ в (1)–(2):

$$\bar{\mathbf{Z}}(p) = \begin{bmatrix} \overset{=}{x_0} & \overset{=}{x_{-1}} & \cdots & \overset{=}{x_{1-p}} \\ \overset{=}{x_1} & \overset{=}{x_0} & \cdots & \overset{=}{x_{2-p}} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{=}{x_{n-1}} & \overset{=}{x_{n-2}} & \cdots & \overset{=}{x_{n-p}} \end{bmatrix}; \quad (11)$$

$\boldsymbol{\xi}$ – случайный $(n \times 1)$ -вектор с нулевым математическим ожиданием

$$\boldsymbol{\xi} = \boldsymbol{\varepsilon} + \mathbf{\Gamma}(-2; Z) \mathbf{S} \overset{\circ}{\boldsymbol{\theta}} + \boldsymbol{\zeta}(-1), \quad E\{\boldsymbol{\xi}\} = \mathbf{0}_n; \quad (12)$$

$$\mathbf{\Gamma}(-2; Z) = \begin{bmatrix} \zeta_{-1} & \zeta_{-2} & \cdots & \zeta_{-p} \\ \zeta_0 & \zeta_{-1} & \cdots & \zeta_{1-p} \\ \vdots & \vdots & \ddots & \vdots \\ \zeta_{n-2} & \zeta_{n-3} & \cdots & \zeta_{n-1-p} \end{bmatrix} - \quad (13)$$

$(n \times p)$ -матрица ненаблюдаемых случайных величин.

Введем обозначения

$$\mathbf{y} = \mathbf{x}, \quad \mathbf{R} = \bar{\mathbf{Z}}(p) \mathbf{S}, \quad (14)$$

где \mathbf{R} – $(n \times m)$ -матрица регрессоров для выходной переменной.

Учитывая (14), регрессионную модель (10) запишем в виде

$$\mathbf{y} = \mathbf{R} \overset{\circ}{\boldsymbol{\theta}} + \boldsymbol{\xi} = \overset{\circ}{\mathbf{y}} + \boldsymbol{\xi}, \quad (15)$$

где \mathbf{y} – $(n \times 1)$ -вектор наблюдаемых зашумленных значений выходной переменной; $\overset{\circ}{\mathbf{y}}$ – $(n \times 1)$ -вектор ненаблюдаемых значений; $\overset{\circ}{\boldsymbol{\theta}}$ – $(m \times 1)$ -вектор неизвестных детерминированных коэффициентов; $\boldsymbol{\xi}$ – $(n \times 1)$ -вектор ненаблюдаемых случайных аддитивных составляющих.

Согласно [6], для оценки коэффициентов $\overset{\circ}{\boldsymbol{\theta}}$ выполняется

$$\hat{\mathbf{d}} = \mathbf{C} \mathbf{y}, \quad (16)$$

где для $(m \times n)$ -матрицы \mathbf{C} выполняется

$$\mathbf{C} = (\mathbf{R}^T \boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{-1} \mathbf{R})^{-1} \mathbf{R}^T \boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{-1}, \quad (17)$$

а Σ_{ξ} – ковариационная $(n \times n)$ -матрица введённого в (15) $(n \times 1)$ -вектора ненаблюдаемых аддитивных случайных составляющих ξ .

Для ковариационной матрицы Σ_{ξ} выполняется [6]:

$$\Sigma_{\xi} = \sigma_{\varepsilon} \cdot \mathbf{I}_n + \Psi + \sigma_{\zeta} \cdot \mathbf{I}_n, \quad (18)$$

где σ_{ε} – дисперсия в модели наблюдения в (8); σ_{ζ} – дисперсия в модели функционирования в (7); Ψ – $(n \times n)$ -матрица имеет вид

$$\Psi = \begin{bmatrix} \psi(0) & \psi(+1) & \dots & \psi(p-1) & 0 & \dots & 0 & 0 \\ \psi(-1) & \psi(0) & \dots & \psi(p-2) & \psi(p-1) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \psi(1-p) & \psi(2-p) & \dots & \psi(0) & \psi(+1) & \dots & 0 & 0 \\ 0 & \psi(1-p) & \dots & \psi(-1) & \psi(0) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & \psi(0) & \psi(+1) \\ 0 & 0 & \dots & 0 & 0 & \dots & \psi(-1) & \psi(0) \end{bmatrix}. \quad (19)$$

В (19) для величин $\psi(\Delta)$, $\Delta = -p+1, -p+2, \dots, p-2, p-1$, выполняется

$$\psi(\Delta) = \text{Cov} \{ \xi_{i_1} \xi_{i_2} \} = \sigma_{\zeta} \cdot \overset{\circ}{\boldsymbol{\theta}}^T \mathbf{S}^T \mathbf{I}_p(i_1 - i_2) \mathbf{S} \overset{\circ}{\boldsymbol{\theta}}, \quad (20)$$

где $\mathbf{I}_p(i_1 - i_2)$ – $(p \times p)$ -матрица, в которой все элементы равны нулю, кроме элементов одной диагонали, равных единице: если $\Delta = i_1 - i_2 = 0$, то это главная диагональ; если $\Delta > 0$, то это диагональ, расположенная выше главной диагонали на Δ строк; если $\Delta < 0$, то это диагональ – ниже главной диагонали на $|\Delta|$ строк.

С учетом (17)–(20) для оценок коэффициентов выполняется

$$\hat{\mathbf{d}} = (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{y}. \quad (21)$$

В формулу (17) для матрицы \mathbf{C} входит ненаблюдаемая матрица регрессоров \mathbf{R} , а в формулу (18) для матрицы Σ_{ξ} – матрица Ψ , элементы которой, как следует из (19)–(20), зависят от неизвестных коэффициентов $\overset{\circ}{\boldsymbol{\theta}}$. Эти обстоятельства использованы [5] для построения итерационной процедуры вычисления неизвестных коэффициентов в виде (21). Процедура является частным случаем итерационных процедур [6]–[7],

разработанных для оценивания коэффициентов в системах авторегрессионных уравнений: первая из них разработана для случая, когда ковариационные матрицы случайных составляющих в модели функционирования и модели наблюдения априорно известны, а вторая – для случая, когда они неизвестны.

Для оценки $\hat{\mathbf{d}}$ в (21) выполняется

$$\hat{\mathbf{d}} = [(\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1}] \mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R} \mathbf{\theta}^0 + \mathbf{C} \xi = \mathbf{\theta}^0 + \mathbf{C} \xi. \quad (22)$$

С учётом (22) регрессионную модель запишем в виде

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{u} = \mathbf{R} \hat{\mathbf{d}} + \mathbf{u}, \quad (23)$$

где $\hat{\mathbf{y}}$ – $(n \times 1)$ -вектор выхода регрессионной модели

$$\hat{\mathbf{y}} = \mathbf{R} \hat{\mathbf{d}} = \mathbf{R} \mathbf{\theta}^0 + \mathbf{R} \mathbf{C} \xi = \hat{\mathbf{y}} + \mathbf{R} \mathbf{C} \xi; \quad (24)$$

\mathbf{u} – $(n \times 1)$ -вектор остатков [8], для которого выполняется

$$\mathbf{u} = \xi - \mathbf{R} \mathbf{C} \xi, \quad E\{\mathbf{u}\} = \mathbf{0}_n, \quad (25)$$

т. е. его математическое ожидание равно нулевому $(n \times 1)$ -вектору.

Для квадратичной формы вектора остатков (25) выполняется

$$\begin{aligned} \omega &= E\{\mathbf{u}^T \mathbf{u}\} = E\{(\xi - \mathbf{R} \mathbf{C} \xi)^T (\xi - \mathbf{R} \mathbf{C} \xi)\} = \\ &= \text{tr} [E\{\xi \xi^T\} - \mathbf{R} \mathbf{C} E\{\xi \xi^T\} - E\{\xi \xi^T\} \mathbf{C}^T \mathbf{R}^T + \mathbf{R} \mathbf{C} E\{\xi \xi^T\} \mathbf{C}^T \mathbf{R}^T] = \\ &= \text{tr} [\Sigma_{\xi} - \mathbf{R} \mathbf{C} \Sigma_{\xi} - \Sigma_{\xi} \mathbf{C}^T \mathbf{R}^T + \mathbf{R} \mathbf{C} \Sigma_{\xi} \mathbf{C}^T \mathbf{R}^T] = \\ &= \text{tr} [\Sigma_{\xi} - \mathbf{R} (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T \Sigma_{\xi}^{-1} \Sigma_{\xi} - \Sigma_{\xi} \Sigma_{\xi}^{-1} \mathbf{R} (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T + \\ &\quad + \mathbf{R} (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T \Sigma_{\xi}^{-1} \Sigma_{\xi} \Sigma_{\xi}^{-1} \mathbf{R} (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T] = \\ &= \text{tr} [\Sigma_{\xi} - \mathbf{R} (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T] = \\ &= n \cdot (\sigma_{\varepsilon} + \psi(0) + \sigma_{\zeta}) - \text{tr} [\mathbf{R} (\mathbf{R}^T \Sigma_{\xi}^{-1} \mathbf{R})^{-1} \mathbf{R}^T]. \quad (26) \end{aligned}$$

Используем результаты (10)–(26) для построения решающего правила, предназначенного для классификации новых наблюдений.

3. Решающее правило классификации на основе двух авторегрессионных моделей

Пусть в процессе своего функционирования исследуемый объект может находиться в одном из двух классов состояний, каждый из которых характеризуется своей авторегрессионной моделью. Пусть $X(I)$ и

$X(II)$ – выборки наблюдений за состоянием объекта, полученные для первого P_I и второго P_{II} класса состояний в соответствии с (1)–(9). Тогда возможно формулирование следующей задачи статистической классификации: на основе обучающих выборок $X(I)$ и $X(II)$ построить решающее правило, которое устанавливает принадлежность анализируемого (существующего или ожидаемого) наблюдения (или ряда наблюдений) к одному из двух классов состояний. Решение задачи предполагает построение двух авторегрессионных моделей для двух классов состояний P_I и P_{II} .

Пусть, в соответствии с формулой (14), $(l \times 1)$ -вектор наблюдений $\mathbf{y}_* = (y_*(1), y_*(2), \dots, y_*(l))^T = (x_{n+1}, x_{n+2}, \dots, x_{n+l})^T$ представляет собой вектор наблюдений, принадлежность которого к первому или второму классу состояний требуется установить.

Пусть $(l \times m)$ -матрица $\mathbf{R}_* = [\mathbf{r}_*(1), \mathbf{r}_*(2), \dots, \mathbf{r}_*(m)]$, представляет собой столбцы регрессоров, сформированные в соответствии с (14) для вектора наблюдений $\mathbf{y}_* = (y_*(1), y_*(2), \dots, y_*(l))^T$.

Рассмотрим случай, в котором исследуемый вектор наблюдений (отрезок временного ряда) принадлежит классу состояний P_I . (В силу симметрии случай принадлежности классу состояний P_{II} может быть рассмотрен аналогично).

Рассмотрим сначала *правильное решение*, состоящее в том, что исследуемый вектор наблюдений отнесён к классу состояний P_I .

Учитывая (15) и (22)–(25), запишем случайный вектор остатков

$$\mathbf{u}_*(I, I) = \mathbf{y}_*(I) - \hat{\mathbf{y}}_*(I) = \boldsymbol{\xi}_*(I) - \mathbf{R}_*(I)\mathbf{C}(I)\boldsymbol{\xi}, \quad (27)$$

или в покомпонентном виде

$$\begin{pmatrix} u_*(I | I, 1) \\ u_*(I | I, 2) \\ \vdots \\ u_*(I | I, l) \end{pmatrix} = \begin{pmatrix} \xi_*(I, 1) \\ \xi_*(I, 2) \\ \vdots \\ \xi_*(I, l) \end{pmatrix} - [\mathbf{R}_*(I)\mathbf{C}(I)] \begin{pmatrix} \xi(I, 1) \\ \xi(I, 2) \\ \vdots \\ \xi(I, n) \end{pmatrix}, \quad (28)$$

где $\mathbf{y}_*(I)$ – $(l \times 1)$ -вектор исследуемых наблюдений выходной переменной для первого класса состояний; $\hat{\mathbf{y}}_*(I) = (\hat{y}_*(I, 1), \hat{y}_*(I, 2), \dots, \hat{y}_*(I, l))^T$ –

$(l \times 1)$ -вектор расчетных выходов авторегрессионной модели для первого класса состояний, который получается после подстановки в модель матрицы регрессоров $\mathbf{R}_*(I) = [\mathbf{r}_*(I,1), \mathbf{r}_*(I,2), \dots, \mathbf{r}_*(I,m)]$;

$$\mathbf{C}(I) = (\mathbf{R}^T(I) \boldsymbol{\Sigma}_\xi^{-1}(I) \mathbf{R}(I))^{-1} \mathbf{R}^T(I) \boldsymbol{\Sigma}_\xi^{-1}(I).$$

Введем случайный вектор $\boldsymbol{\eta}_*(I|I) = (\xi(I,1), \xi(I,2), \dots, \xi(I,n(I)), \xi_*(I,1), \xi_*(I,2), \dots, \xi_*(I,l))^T$, и запишем случайный вектор $\mathbf{u}_*(I|I)$:

$$\begin{aligned} \mathbf{u}_*(I|I) &= [\mathbf{M}_1(I|I) - [\mathbf{R}_*(I)\mathbf{C}(I)]\mathbf{M}_2(I|I)]\boldsymbol{\eta}_*(I|I) = \\ &= [\mathbf{M}_1(I|I) - \mathbf{P}_*(I|I)\mathbf{M}_2(I|I)]\boldsymbol{\eta}_*(I|I) = \mathbf{D}_*(I|I)\boldsymbol{\eta}_*(I|I), \end{aligned} \quad (29)$$

где

$$\mathbf{P}_*(I|I) = \mathbf{R}_*(I)\mathbf{C}(I); \quad \mathbf{D}_*(I|I) = \mathbf{M}_1(I|I) - \mathbf{P}_*(I|I)\mathbf{M}_2(I|I); \quad (30)$$

а для блочных матриц $\mathbf{M}_1(I|I)$ и $\mathbf{M}_2(I|I)$ выполняется:

$$\mathbf{M}_1(I|I) = \begin{bmatrix} \mathbf{O}_{(l \times n(I))} & \mathbf{I}_l \end{bmatrix}, \quad \mathbf{M}_2(I|I) = \begin{bmatrix} \mathbf{I}_{n(I)} & \mathbf{O}_{(n(I) \times l)} \end{bmatrix}; \quad \mathbf{O}_{(l \times n(I))} - \text{нулевая } (l \times n(I))\text{-матрица; } \mathbf{I}_l - \text{единичная } (l \times l)\text{-матрица; } \mathbf{I}_{n(I)} - \text{единичная } (n(I) \times n(I))\text{-матрица; } \mathbf{O}_{(n(I) \times l)} - \text{нулевая } (n(I) \times l)\text{-матрица.}$$

Вычислим ковариационную матрицу случайного вектора $\mathbf{u}_*(I|I)$. Для этого воспользуемся следующим известным результатом многомерного статистического анализа (см., например, теорему 2.4.5 на стр. 41 в [9]).

Теорема а. Если векторная случайная величина $\boldsymbol{\eta}$ распределена по многомерному нормальному закону $\boldsymbol{\eta} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, то векторная случайная величина $\boldsymbol{\gamma} = \mathbf{D}\boldsymbol{\eta}$ распределена по многомерному нормальному закону $\boldsymbol{\gamma} \sim N(\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^T)$.

Для ковариационной матрицы случайного вектора $\boldsymbol{\eta}_*(I|I)$ выполняется

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\eta}}(I|I) &= E\{(\boldsymbol{\eta}_*(I|I) - E\{\boldsymbol{\eta}_*(I|I)\})(\boldsymbol{\eta}_*(I|I) - E\{\boldsymbol{\eta}_*(I|I)\})^T\} = \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_\xi(I) & \mathbf{O}_{(n(I) \times l)} \\ \mathbf{O}_{(l \times n(I))} & \boldsymbol{\Sigma}_*(I) \end{bmatrix}, \end{aligned} \quad (31)$$

где

$$\boldsymbol{\Sigma}_\xi(I) = \sigma_\varepsilon(I) \cdot \mathbf{I}_{n(I)} + \boldsymbol{\Psi}(I) + \sigma_\zeta(I) \cdot \mathbf{I}_{n(I)}, \quad (32)$$

$$\Sigma_*(I) = \sigma_\varepsilon(I) \cdot \mathbf{I}_l + \Psi_*(I) + \sigma_\zeta(I) \cdot \mathbf{I}_l, \quad (33)$$

где дисперсии $\sigma_\varepsilon(I)$ и $\sigma_\zeta(I)$ определены в (7)–(8), $(n(I) \times n(I))$ -матрица $\Psi(I)$ – в (19), а $(l \times l)$ -матрица $\Psi_*(I)$ определяется аналогично (19) для исследуемого $(l \times 1)$ -вектора наблюдений \mathbf{y}_* .

Применяя приведенную выше теорему и учитывая (29) и (31)–(33), для ковариационной матрицы случайного вектора $\mathbf{u}_*(I|I)$ получаем

$$\mathbf{\Omega}_*(I|I) = \mathbf{D}_*(I|I) \Sigma_\eta(I|I) \mathbf{D}_*^T(I|I) = \Sigma_*(I) + \mathbf{P}_*(I|I) \Sigma_\xi(I) \mathbf{P}_*^T(I|I). \quad (34)$$

Квадратичная форма

$$l_*(I|I) = \mathbf{u}_*^T(I|I) \mathbf{\Omega}_*^{-1}(I|I) \mathbf{u}_*(I|I) \quad (35)$$

имеет χ^2 -распределение с h степенями свободы (см. теорему 3.3.3 в [9]).

Рассмотрим теперь *ошибочное решение*, состоящее в том, что исследуемый вектор наблюдений, принадлежащее классу P_I , отнесено к классу P_{II} .

Рассмотрим случайный вектор

$$\begin{aligned} \mathbf{u}_*(II, I) &= \mathbf{y}_*(I) - \hat{\mathbf{y}}_*(II) = \overset{o}{\mathbf{y}}_*(I) + \overset{o}{\boldsymbol{\xi}}_*(I) - \overset{o}{\mathbf{y}}_*(II) - \mathbf{R}_*(I) \mathbf{C}(II) \boldsymbol{\xi} = \\ &= \begin{pmatrix} \overset{o}{y}_*(I,1) - \overset{o}{y}_*(II,1) \\ \overset{o}{y}_*(I,2) - \overset{o}{y}_*(II,2) \\ \vdots \\ \overset{o}{y}_*(I,l) - \overset{o}{y}_*(II,l) \end{pmatrix} + \begin{pmatrix} \overset{o}{\xi}_*(I,1) \\ \overset{o}{\xi}_*(I,2) \\ \vdots \\ \overset{o}{\xi}_*(I,l) \end{pmatrix} - [\mathbf{R}_*(I) \mathbf{C}(II)] \begin{pmatrix} \overset{o}{\xi}(II,1) \\ \overset{o}{\xi}(II,2) \\ \vdots \\ \overset{o}{\xi}(II,l) \end{pmatrix}, \quad (36) \end{aligned}$$

где $\mathbf{y}_*(I)$ – $(l \times 1)$ -вектор исследуемых наблюдений выходной переменной для первого класса состояний;

$\hat{\mathbf{y}}_*(II) = (\hat{y}_*(II,1), \hat{y}_*(II,2), \dots, \hat{y}_*(II,l))^T$ – $(l \times 1)$ -вектор выходов модели для второго класса состояний, который получается после подстановки в модель матрицы регрессоров $\mathbf{R}_*(I) = [\mathbf{r}_*(I,1), \mathbf{r}_*(I,2), \dots, \mathbf{r}_*(I,m)]$;

$$\mathbf{C}(II) = (\mathbf{R}^T(II) \Sigma_\xi^{-1}(II) \mathbf{R}(II))^{-1} \mathbf{R}^T(II) \Sigma_\xi^{-1}(II).$$

Введем случайный вектор $\boldsymbol{\eta}_*(II|I) = (\xi(II,1), \xi(II,2), \dots, \xi(II,n(II)), \xi_*(I,1), \xi_*(I,2), \dots, \xi_*(I,l))^T$, и запишем случайный вектор $\mathbf{u}_*(II|I)$:

$$\begin{aligned} \mathbf{u}_*(II|I) &= \begin{pmatrix} \overset{\circ}{0} & \overset{\circ}{0} \\ y_*(I,1) - y_*(II,1) & \overset{\circ}{0} \\ \overset{\circ}{0} & \overset{\circ}{0} \\ y_*(I,2) - y_*(II,2) & \overset{\circ}{0} \\ & \vdots \\ \overset{\circ}{0} & \overset{\circ}{0} \\ y_*(I,l) - y_*(II,l) & \overset{\circ}{0} \end{pmatrix} + [\mathbf{M}_1(II|I) - \mathbf{P}_*(II|I)\mathbf{M}_2(II|I)]\boldsymbol{\eta}_*(II|I) = \\ &= \overset{\circ}{\mathbf{y}}_*(I) - \overset{\circ}{\mathbf{y}}_*(II) + [\mathbf{M}_1(II|I) - \mathbf{P}_*(II|I)\mathbf{M}_2(II|I)]\boldsymbol{\eta}_*(II|I) = \\ &= \boldsymbol{\delta}_*(II|I) + \mathbf{D}_*(II|I)\boldsymbol{\eta}_*(II|I), \end{aligned} \quad (37)$$

где

$$\mathbf{P}_*(II|I) = \mathbf{R}_*(I)\mathbf{C}(II); \quad \mathbf{D}_*(II|I) = \mathbf{M}_1(II|I) - \mathbf{P}_*(II|I)\mathbf{M}_2(II|I); \quad (38)$$

а для блочных матриц $\mathbf{M}_1(II|I)$ и $\mathbf{M}_2(II|I)$ выполняется

$$\mathbf{M}_1(II|I) = [\mathbf{O}_{(l \times n(II))} \mid \mathbf{I}_l], \quad \mathbf{M}_2(II|I) = [\mathbf{I}_{n(II)} \mid \mathbf{O}_{(n(II) \times l)}];$$

$\mathbf{O}_{(l \times n(II))}$ – нулевая матрица размера $(l \times n(II))$; \mathbf{I}_l – единичная $(l \times l)$ -матрица; $\mathbf{I}_{n(II)}$ – единичная $(n(II) \times n(II))$ -матрица; $\mathbf{O}_{(n(II) \times l)}$ – нулевая матрица размера $(n(II) \times l)$; $\boldsymbol{\delta}_*(II|I) = \overset{\circ}{\mathbf{y}}_*(I) - \overset{\circ}{\mathbf{y}}_*(II)$.

Для ковариационной матрицы случайного вектора $\boldsymbol{\eta}_*(II|I)$ выполняется

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\eta}}(II|I) &= E\{(\boldsymbol{\eta}(II|I) - E\{\boldsymbol{\eta}(II|I)\})(\boldsymbol{\eta}(II|I) - E\{\boldsymbol{\eta}(II|I)\})^T\} = \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_{\xi}(II) & \mathbf{O}_{(n(II) \times l)} \\ \mathbf{O}_{(l \times n(II))} & \boldsymbol{\Sigma}_*(I) \end{bmatrix}, \end{aligned} \quad (39)$$

где

$$\boldsymbol{\Sigma}_{\xi}(II) = \sigma_{\varepsilon}(II) \cdot \mathbf{I}_{n(II)} + \boldsymbol{\Psi}(II) + \sigma_{\zeta}(II) \cdot \mathbf{I}_{n(II)}, \quad (40)$$

где дисперсии $\sigma_{\varepsilon}(II)$ и $\sigma_{\zeta}(II)$ определены в (7)–(8), $(n(II) \times n(II))$ -матрица $\boldsymbol{\Psi}(II)$ – в (19), а матрица $\boldsymbol{\Sigma}_*(I)$ – в (33).

Применяя приведенную выше теорему и учитывая (36) и (39)–(40), для ковариационной матрицы случайного вектора $\mathbf{u}_*(II|I)$ получаем

$$\begin{aligned} \boldsymbol{\Omega}_*(II|I) &= \mathbf{D}_*(II|I)\boldsymbol{\Sigma}_{\boldsymbol{\eta}}(II|I)\mathbf{D}_*^T(II|I) = \\ &= \boldsymbol{\Sigma}_*(I) + \mathbf{P}_*(II|I)\boldsymbol{\Sigma}_{\xi}(II)\mathbf{P}_*^T(II|I). \end{aligned} \quad (41)$$

Квадратичная форма

$$l_*(II | I) = \mathbf{u}_*^T(II | I) \mathbf{\Omega}_*^{-1}(II | I) \mathbf{u}_*(II | I) \quad (42)$$

имеет нецентральное χ^2 -распределение с h степенями свободы и параметром нецентральности (см., например, теорему 5.4.1 в [9])

$$\tau_*^2(II | I) = \mathbf{\delta}_*^T(II | I) \mathbf{\Omega}_*^{-1}(II | I) \mathbf{\delta}_*(II | I). \quad (43)$$

На основе результатов (28)–(43), сформулируем *решающее правило*:

если $l_*(I | I) < l_*(II | I)$, то наблюдение \mathbf{y}_* будем относить к классу P_I ;
если $l_*(I | I) > l_*(II | I)$, то наблюдение \mathbf{y}_* будем относить к классу P_{II} . (44)

Для вычисления теоретического значения вероятности ошибочной классификации наблюдения, фактически принадлежащего первой совокупности, необходимо знать совместную плотность распределения случайных величин $l_*(I | I)$ и $l_*(II | I)$. Запись совместной плотности представляет сложную задачу, поскольку эти величины *статистически зависимы*.

Менее сложной задачей является вычисление вероятности ошибочной классификации наблюдения, фактически принадлежащего первому классу состояний, для конкретного значения случайной величины $l_*(I | I) = l$. В этом случае вероятность отнести анализируемое наблюдение к классу состояний P_{II} , в то время как оно принадлежит P_I , есть не что иное, как вероятность выполнения неравенства $l_*(II | I) < l$. Поскольку случайная величина $l_*(II | I)$ имеет нецентральное χ^2 -распределение, эту вероятность можно вычислить:

$$\text{prob}(II | I) = \text{prob}(l_*(II | I) < l) = \int_0^l f(x; h, \tau_*^2(II | I)) dx, \quad (45)$$

где $f(x; h, \tau_*^2(II | I))$ – функция плотности χ^2 -распределения с h степенями свободы и параметром нецентральности $\tau_*^2(II | I)$, который введен в (43).

Вероятность правильной классификации наблюдения, фактически принадлежащего первой совокупности, можно записать, используя (45)

$$\text{prob}(I | I) = 1 - \text{prob}(II | I). \quad (46)$$

Отметим, что рассмотренная задача обнаружения изменения свойств динамических объектов может быть поставлена и решена в условиях структурной неопределённости, когда структурные матрицы, введённые в первом пункте, априорно неизвестны. В этом случае построение авторегрессионных моделей оптимальной сложности для двух классов состояний может быть проведено на основе результатов [11]–[12].

Заключение. Рассмотрена задача статистической классификации состояний динамического объекта, который может находиться в двух классах состояний, в каждом из которых его функционирование описывается своей авторегрессией с априорно неизвестными параметрами. Предполагается, что выполнены следующие условия: а) модели функционирования в первом и втором классах различны как по коэффициентам, так и по структуре моделей; б) дисперсии случайных величин в моделях функционирования и моделях наблюдения для первого и второго классов могут быть различными. Построено правило классификации и исследованы его свойства.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Обнаружение изменения свойств сигналов и динамических систем: пер. с англ. / М. Бассвиль, А. Вилски, А. Банвенист и др.; под. ред. М. Бассвиль, А. Банвениста. – М. : Мир, 1989. – 278 с.
2. Бородкин Л. И. Алгоритм обнаружения моментов изменения параметров уравнения случайного процесса / Л. И. Бородкин, В. В. Моттль // Автоматика и телемеханика. – 1976. – № 6. – С. 23–31.
3. Карташов В. Я. Обнаружение структурно-параметрических изменений в стохастических системах в реальном масштабе времени алгоритмами непрерывных дробей и структурного анализа / В. Я. Карташов, М. А. Новосельцева // Управление большими системами. – 2011. – Вып. 34. – С. 62–91.
4. Миргород В. Ф. Математические модели процессов управляемого изменения состояния силовых и энергетических установок : дисс. ... д-ра техн. наук / Миргород Владимир Фёдорович. – Днепропетровск, 2012. – 344 с.

5. Сарычев А. П. Линейная авторегрессия на основе метода группового учёта аргументов в условиях квазиповторных наблюдений / А. П. Сарычев // Искусственный интеллект. – 2015. – № 3–4 (69–70). – С. 105–123.
6. Сарычев А. П. Идентификация параметров систем авторегрессионных уравнений при известных ковариационных матрицах // Международный научно-технический журнал “Проблемы управления и информатики”. – 2012. – № 3. – С. 14–30.
7. Сарычев А. П. Исследование методом статистических испытаний итерационной процедуры для идентификации параметров системы авторегрессионных уравнений / А. П. Сарычев // Системні технології. – 2014. – Випуск 3 (92). – С. 77–89.
8. Себер Дж. Линейный регрессионный анализ : пер. с англ. / Дж. Себер. – М. : Мир, 1980. – 456 с.
9. Андерсон Т. Введение в многомерный статистический анализ : пер. с англ. / Т. Андерсон. – М. : Физматгиз. – 1963. – 500 с.
10. Справочник по теории вероятностей и математической статистике / В. С. Корольюк, Н. И. Портенко, А. В. Скороход, А. Ф. Турбин. – М. : Наука, 1985. – 640 с.
11. Сарычев А. П. Моделирование в классе систем авторегрессионных уравнений в условиях структурной неопределенности / А. П. Сарычев // Международный научно-технический журнал “Проблемы управления и информатики”. – 2015. – № 4. – С. 79–103.
12. Александр Сарычев. Моделирование сложных систем в условиях структурной неопределённости: регрессионные и авторегрессионные модели / А. П. Сарычев. LAP LAMBERT Academic Publishing RU, Saarbrücken, Deutschland. – 2016. – 274 с.

REFERENCES

1. Obnaruzhenie izmeneniya svoystv signalov i dinamicheskikh sistem: per. s angl. / M. Bassvil, A. Vilski, A. Banvenist i dr.; pod. red. M. Bassvil, A. Banvenista. – М. : Mir, 1989. – 278 s.

2. Borodkin L. I. Algoritm obnaruzheniya momentov izmeneniya parametrov uravneniya sluchaynogo protsessa / L. I. Borodkin, V. V. Mottl // Avtomatika i telemekhanika. – 1976. – № 6. – S. 23–31.
3. Kartashov V. Ya. Obnaruzhenie strukturno-parametricheskikh izmeneniy v stohasticheskikh sistemah v realnom masshtabe vremeni algoritmami nepreryivnykh drobey i strukturnogo analiza / V. Ya. Kartashov, M. A. Novoseltseva // Upravlenie bolshimi sistemami. – 2011. – Vyip. 34. – S. 62–91.
4. Mirgorod V. F. Matematicheskie modeli protsessov upravlyaemogo izmeneniya sostoyaniya silovyykh i energeticheskikh ustanovok : diss. ... d-ra tehn. nauk / Mirgorod Vladimir Fyodorovich. – Dnepropetrovsk, 2012. – 344 s.
5. Sarychev A. P. Lineynaya avtoregressiya na osnove metoda gruppovogo uchyota argumentov v usloviyakh kvazipovtornykh nablyudeniy / A. P. Sarychev // Iskusstvennyy intellekt. – 2015. – № 3–4 (69–70). – S. 105–123.
6. Sarychev A. P. Identifikatsiya parametrov sistem avtoregressionnykh uravneniy pri izvestnykh kovariatsionnykh matritsah / A. P. Sarychev // Mezhdunarodnyy nauchno-tehnicheskyy zhurnal “Problemy upravleniya i informatiki”. – 2012. – № 3. – S. 14–30.
7. Sarychev A. P. Issledovanie metodom statisticheskikh ispytaniy iteratsionnoy protsedury dlya identifikatsii parametrov sistemy avtoregressionnykh uravneniy / A. P. Sarychev // Sistemni tehnologii. – 2014. – Vipusk 3 (92). – S. 77–89.
8. Seber Dzh. Lineynyy regressionnyy analiz : per. s angl. / Dzh. Seber. – M. : Mir, 1980. – 456 s.
9. Anderson T. Vvedenie v mnogomernyy statisticheskyy analiz : per. s angl. / T. Anderson. – M. : Fizmatgiz. – 1963. – 500 s.
10. Spravochnik po teorii veroyatnostey i matematicheskoy statistike / V. S. Korolyuk, N. I. Portenko, A. V. Skorohod, A. F. Turbin. – M. : Nauka, 1985. – 640 s.
11. Sarychev A. P. Modelirovanie v klasse sistem avtoregressionnykh uravneniy v usloviyakh strukturnoy neopredelennosti / A. P. Sarychev // Mezhdunarodnyy nauchno-tehnicheskyy zhurnal “Problemy upravleniya i informatiki”. – 2015. – № 4. – S. 79–103.

12. Aleksandr Sarychev Modelirovanie slozhnyih sistem v usloviyah strukturnoy neopredelyonnosti: regressionnyie i avtoregressionnyie modeli / A. P. Sarychev LAP LAMBERT Academic Publishing RU, Saarbrücken, Deutschland. – 2016. – 274 с.

Received 15.10.2019.

Accepted 17.10.2019.

**Класифікація станів об'єкту, функціонування якого описується
одновимірною авторегресією**

Розглянуто задачу класифікації станів динамічного об'єкту, який може перебувати у двох класах станів. Функціонування об'єкту в класах описується різними авторегресійними моделями. Побудовано правило статистичної класифікації й досліджені його властивості.

**Classification of the states of object, which functioning is described
by one-dimensional autoregression**

The main purpose of the given work is detection of change of properties of dynamic object that can be in two classes of states, and in each of them the behaviour of object is described by the own autoregression model with unknown parameters.

The problem of detection of change of properties of dynamic objects frequently arises in such areas, as technical and medical diagnostics, the control of technological processes, monitoring, processing of signals.

It is supposed, that there are two groups of observations of functioning of object, which are described by two different autoregression models. Basing on the decision of tasks of identification of two autoregression models, it is necessary to construct deciding rule, which would allow relating new observations to one of two groups of observations. It is supposed, that structures of two models are known, and covariance matrixes of random component of supervision of output variable in two conditions of object can be various.

Basing on the original results on identification autoregression models, the author successfully solves the task of statistical classification. The rule of statistical classification is constructed and its properties are investigated.

Сарычев Александр Павлович – доктор технических наук, ведущий научный сотрудник, Институт технической механики Национальной академии наук Украины и Государственного космического агентства Украины.

Саричев Олександр Павлович - доктор технічних наук, провідний науковий спів-ників, Інститут технічної механіки Національної академії наук України та Державного космічного агентства України.

Sarychev Alexander Pavlovich – doctor of technical sciences, leading scientific researcher, Institute of Technical Mechanics of the National Academy of Sciences of Ukraine and State Space Agency of Ukraine.