



*Российская Академия Наук*

# **А Т АВТОМАТИКА И ТЕЛЕМЕХАНИКА**

**Журнал основан в 1936 году**

**Выходит 12 раз в год**

**4**

**АПРЕЛЬ**

**Москва**

**2025**

Учредители журнала:

Отделение энергетики, машиностроения, механики и процессов управления РАН,  
Институт проблем управления им. В.А. Трапезникова РАН (ИПУ РАН),  
Институт проблем передачи информации им. А.А. Харкевича РАН (ИППИ РАН)

**Главный редактор:**

Галяев А.А.

**Заместители главного редактора:**

Рубинович Е.Я., Хлебников М.В.

**Ответственный секретарь:**

Самохин А.С.

**Редакционный совет:**

Васильев С.Н., Желтов С.Ю., Каляев И.А., Кузнецов О.П., Кулешов А.П.,  
Куржанский А.Б., Мартынюк А.А. (Украина), Пешехонов В.Г., Попков Ю.С.,  
Федосов Е.А., Черноусько Ф.Л.

**Редакционная коллегия:**

Алескеров Ф.Т., Арутюнов А.В., Бахтадзе Н.Н., Бобцов А.А., Виноградов Д.В.,  
Вишневский В.М., Воронцов К.В., Граничин О.Н., Жилиякова Н.Ю.,  
Каравай М.Ф., Кибзун А.И., Краснова С.А., Крищенко А.П., Кузнецов Н.В.,  
Кушнер А.Г., Лазарев А.А., Ляхов А.И., Маликов А.И., Матасов А.И.,  
Меерков С.М. (США), Мещеряков Р.В., Миллер Б.М., Михальский А.И.,  
Моржин О.В., Мунасыпов Р.А., Назин А.В., Немировский А.С. (США),  
Новиков Д.А., Олейников А.Я., Пакшин П.В., Пальчунов Д.Е.,  
Поляков А.Е. (Франция), Протасов В.Ю., Рапопорт Л.Б., Родионов И.В.,  
Сельвесюк Н.И., Соболевский А.Н., Степанов О.А., Филимонюк Л.Ю.,  
Фрадков А.Л., Хоров Е.М., Цыбаков А.Б. (Франция), Чеботарев П.Ю.,  
Чхартишвили А.Г., Щербаков П.С.

Адрес редакции: 117997, Москва, Профсоюзная ул., 65

Тел./факс: 8 (495) 198-17-20, доб. 1443

Электронная почта: redacsia@ipu.ru

Зав. редакцией *Е.А. Мартехина*

Москва

ФГБУ «Издательство «Наука»

© 2025 г. В.О. НИКИФОРОВ, д-р техн. наук (nikiforov\_vo@itmo.ru),  
Д.Н. ГЕРАСИМОВ, канд. тех. наук (dngerasimov@itmo.ru),  
Н.А. ДУДАРЕНКО, канд. тех. наук (dudarenko@itmo.ru)  
(Университет ИТМО, Санкт-Петербург)

## АДАПТИВНАЯ КОМПЕНСАЦИЯ ВНЕШНИХ ВОЗМУЩЕНИЙ В МНОГОКАНАЛЬНЫХ СИСТЕМАХ ПРИ УПРАВЛЕНИИ ПО ВЫХОДУ<sup>1</sup>

Рассматривается задача адаптивной компенсации внешних заранее неизвестных детерминированных возмущений при управлении по выходу (т.е. по вектору выходных переменных) многоканальным линейным объектом. Предлагаемое новое решение основано на адаптивной реализации метода внутренней модели и построении специального наблюдателя, позволяющего получить регрессионную модель многомерного возмущения и синтезировать адаптивный регулятор с числом настраиваемых параметров, равным числу неизвестных коэффициентов характеристического полинома модели возмущения. При определенных условиях динамический порядок такого наблюдателя оказывается существенно ниже, чем у известных решений.

*Ключевые слова:* адаптивная компенсация возмущений, метод внутренней модели, многомерные системы.

DOI: 10.31857/S0005231025040018, EDN: CAZMQB

### 1. Введение

Одним из эффективных и хорошо разработанных методов асимптотической компенсации внешних детерминированных возмущений является метод внутренней модели, что подтверждается практическим опытом его использования в таких приложениях, как системы активной виброзащиты [1–3], двигатели внутреннего сгорания [4, 5], мельничные механизмы [6, 7], процессы непрерывного литья [8], жесткие компьютерные диски [9, 10], системы сейсмозащиты [11], морские краны [12, 13], системы снижения пульсаций в преобразователях тока [14] и в синхронных машинах [15] и др. В соответствии с данным методом внешнее возмущение моделируется в виде выхода динамической автономной системы (генератора возмущения), а для его полной компенсации данная модель соответствующим образом воспроизводится

---

<sup>1</sup> Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (госзадание № 2019-0898).

в контуре замкнутой системы. Изначально метод внутренней модели был разработан для класса линейных систем и линейных моделей внешних возмущений с известными параметрами [16; 17; 18, глава 4]. Однако предположение об известности параметров модели внешнего возмущения снижает практическую ценность такого подхода, так как означает необходимость точного знания формы и частот внешнего возмущения. Случай, когда заранее известен только класс возмущений, но не их точные параметры, имеет большее практическое значение.

Для компенсации заранее неизвестных возмущений была предложена адаптивная реализация метода внутренней модели, при которой параметры модели генератора полагаются неизвестными (для обзора см. [19, секция 1.4; 20] и приведенные в них списки литературы). Данный подход был разработан как для систем со скалярным входом и скалярным выходом [21–25], так и для многоканальных систем с управлением по вектору состояния [19, секция 4.2; 26; 27] или по выходу [20, 28]. Особый класс представляют многоканальные системы с возмущением, приведенным ко входу [29].

Однако известные решения компенсации возмущений на основе принципа внутренней модели в многоканальных системах по выходу (как неадаптивные [17; 18, глава 4], так и адаптивные [20, 28]) предусматривают построение в каждом канале многоканальной системы специального наблюдателя (динамического фильтра, предкомпенсатора), размерность которого равна размерности всей модели генератора возмущений. Таким образом, общая размерность используемых наблюдателей равна произведению размерности модели генератора возмущений на число каналов в системе управления. Это ведет к увеличению динамического порядка регулятора, а также к увеличению числа настраиваемых параметров.

В настоящей статье показано, что при дополнительном неограничительном предположении о блочно-диагональной структуре генератора возмущения возможно построение системы управления с одним наблюдателем, размерность которого равна размерности модели командного генератора. Как следствие, существенно понижается не только динамическая размерность регулятора, но и число настраиваемых параметров.

Предлагаемое решение основано на походе, разработанном в [19, 30, 31], и состоит в построении специального наблюдателя внешнего возмущения, позволяющего получить удобную несмешанную параметризацию возмущения. В свою очередь это позволяет использовать разработанный аппарат матричных уравнений [18] и строго доказать достижение цели управления.

Статья организована следующим образом. Во втором разделе приведена постановка задачи. В третьем разделе представлен наблюдатель возмущения и обсуждены проблемы его построения для многомерных систем. В разделах 4 и 5 синтезирован адаптивный регулятор, а в разделе 6 приводятся результаты моделирования.

В статье используются следующие обозначения:  $|y|$  – евклидова норма вектора  $y$ ;  $s$  – комплексная переменная; запись  $W(s)[\delta]$  означает преобразование сигнала  $\delta$  (возможно, векторного) динамическим блоком с передаточной функцией (матрицей для векторного сигнала)  $W(s)$ .

## 2. Постановка задачи

В статье рассматривается стационарный многоканальный объект вида

$$(1) \quad \begin{cases} \dot{x} = Ax + B(u + \delta), & x(0), \\ y = Cx, \end{cases}$$

где  $x \in \mathbb{R}^n$  – вектор состояния,  $u \in \mathbb{R}^m$  – вектор управляющих воздействий,  $y \in \mathbb{R}^m$  – вектор регулируемых переменных,  $n \geq m$ ,  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{m \times n}$  – известные матрицы постоянных коэффициентов,  $\delta \in \mathbb{R}^m$  – вектор неизмеряемых внешних возмущений.

*Допущение 1.* Относительно объекта управления (1) полагаются выполненными следующие допущения:

**Д.1.1.** тройка матриц  $(A, B, C)$  полностью управляемая и наблюдаемая, а матрицы  $B$  и  $C$  полного ранга, т.е.  $\text{rank } B = \text{rank } C = m$ ;

**Д.1.2.** модель вход–состояние–выход (1) является минимально-фазовой;

**Д.1.3.** прямым измерениям доступен только вектор регулируемых переменных  $y$ , но не вектор состояния  $x$ .

*Замечание 1.* Допущение Д.1.1 означает, что система (1) является управляемой по вектору регулируемых переменных  $y$  (см. определение и критерий в [32]).

*Замечание 2.* В данной статье под минимально-фазовой моделью (1) понимается модель, не имеющая инвариантных нулей или все инвариантные нули которой имеют отрицательные вещественные части. Под *инвариантным нулем* понимается комплексное число  $s_0$ , при подстановке которого в матрицу Розенброка (т.е., при  $s = s_0$ )

$$P(s) = \begin{bmatrix} sI - A & B \\ -C & 0 \end{bmatrix}$$

ее столбцовый ранг понижается [33, с. 237]. Для систем с одинаковым числом управляющих воздействий и регулируемых переменных, у которых матрица Розенброка является квадратной, это означает вырождение данной матрицы при  $s = s_0$ .

*Допущение 2.* Внешнее возмущение  $\delta = [\delta_1, \dots, \delta_m]^\top$  такое, что:

**Д.2.1.** каждая его компонента  $\delta_i$  ( $i = 1, \dots, m$ ) может быть представлена в виде выхода линейного генератора (экзосистемы)

$$(2) \quad \begin{cases} \dot{\xi}_i = \Gamma_i \xi_i, & \xi_i(0), \\ \delta_i = h_i^\top \xi_i, \end{cases}$$

где  $\xi_i \in \mathbb{R}^{q_i}$  – вектор состояния,  $\Gamma_i \in \mathbb{R}^{q_i \times q_i}$  – матрица постоянных коэффициентов, все собственные числа которой являются простыми и лежат на мнимой оси,  $h_i \in \mathbb{R}^{q_i}$  – вектор постоянных коэффициентов;

**Д.2.3.** параметры матриц  $\Gamma_i$  и векторов  $h_i$  являются неизвестными, в то время как размерности  $q_i$  известны;

**Д.2.3.** без потери общности пары  $(h_i^\top, \Gamma_i)$  считаются полностью наблюдаемыми.

*Замечание 3.* Общая экзосистема, генерирующая возмущение  $\delta$ , может быть представлена в виде

$$(3) \quad \begin{cases} \dot{\xi} = \Gamma \xi, & \xi(0), \\ \delta = H \xi, \end{cases}$$

где  $\xi = [\xi_1^\top, \dots, \xi_m^\top]^\top \in \mathbb{R}^q$ , матрицы  $\Gamma = \text{diag}\{\Gamma_i\}$  и  $H = \text{diag}\{h_i^\top\}$  ( $i = 1, \dots, m$ ) являются блочно-диагональными и  $q = \sum_{i=1}^m q_i$ .

В статье решается следующая задача.

*Задача 1.* Требуется синтезировать закон управления в форме обратной связи по вектору регулируемых переменных  $y$ , обеспечивающий ограниченность всех сигналов в замкнутой системе и достижение целевого условия

$$(4) \quad \lim_{t \rightarrow \infty} |y(t)| = 0.$$

Другими словами, требуется стабилизация объекта с одновременной компенсацией внешних возмущений. Так как в соответствии с допущением Д.2.3 внешнее возмущение является заранее неизвестным, то классический метод внутренней модели [16–18] в данном случае является неприменимым, и для решения поставленной задачи будет использована адаптивная модификация метода внутренней модели [19, 21, 31]. Главное отличие от известных решений, разработанных ранее для многоканальных систем [19, 26], состоит в том, что доступными для управления являются только выходные переменные  $y$ , а не весь вектор состояния  $x$ . Как отмечено ниже (см. раздел 3.2), задача компенсации неизвестных возмущений в многоканальных системах при управлении по выходу имеет принципиальную особенность, которая не позволяет непосредственно использовать результаты, полученные для одноканальных систем.

Поставленная задача будет решена в три шага. Сначала с помощью специального наблюдателя будет получена параметризованная модель внешнего возмущения в виде линейной регрессии с постоянной матрицей неизвестных коэффициентов и физически реализуемым регрессором (измеряемым/вычисляемым по управлению и выходу). Затем будет построен алгоритм настройки компенсирующей компоненты закона управления. На последнем шаге будет построен наблюдатель вектора состояния и синтезирована стабилизирующая компонента управления.

### 3. Наблюдатель возмущения

#### 3.1. Каноническая форма экосистемы и параметризация возмущения

При управлении по выходной переменной для построения наблюдателя скалярного возмущения оказывается полезным следующее утверждение, позволяющее выразить возмущение  $\delta_i$  через *фильтрованное возмущение*  $\delta_{fi} = w_i(s)[\delta_i]$ , где  $w_i(s)$  – минимально фазовая асимптотически устойчивая передаточная функция [19, с. 315].

*Лемма 1.* Скалярное возмущение  $\delta_i$  может быть представлено в виде

$$(5) \quad \delta_i = \psi_{fi}^\top \xi_{fi} + \epsilon_i,$$

где  $\psi_{fi} \in \mathbb{R}^{q_i}$  – вектор постоянных неизвестных параметров,  $\epsilon_i$  – экспоненциально затухает, регрессор  $\xi_{fi} \in \mathbb{R}^{q_i}$  является вектором состояния наблюдателя

$$(6) \quad \dot{\xi}_{fi} = G_i \xi_{fi} + l_i \delta_{fi},$$

где  $G_i \in \mathbb{R}^{q_i \times q_i}$  – произвольная гурвицева матрица,  $l_i \in \mathbb{R}^{q_i}$  – постоянный вектор, такой что пара матриц  $(G_i, l_i)$  образует управляемую пару матриц.

В методических целях представим доказательство леммы 1, которое отличается от доказательства, приведенного в [19, с. 315]. На основе его анализа в разделе 3.2 покажем принципиальную проблему, возникающую в случае многоканальных систем.

*Доказательство.* Пусть  $(A_{fi}, b_{fi}, c_{fi})$  – минимальная реализация передаточной функции  $w_i(s)$  и

$$\delta_{fi} = c_{fi}^\top \chi_i, \quad \dot{\chi}_i = A_{fi} \chi_i + b_{fi} \delta_i,$$

где  $\chi_i$  – вектор состояния минимальной реализации фильтра  $w(s)$ . Известно [18, с. 87], что вынужденная составляющая  $\chi_i^*$  вектора состояния  $\chi_i$  в установившемся режиме (т.е. без учета экспоненциально затухающей переходной компоненты) может быть представлена в виде  $\chi_i^* = M_{\xi_i} \xi_i$ , где матрица подобия  $M_{\xi_i}$  является решением уравнения Сильвестра

$$(7) \quad M_{\xi_i} \Gamma_i - A_{fi} M_{\xi_i} = b_{fi} h_i^\top.$$

Так как матрица  $A_{fi}$  гурвицева, а собственные числа матрицы  $\Gamma_i$  лежат на мнимой оси, то данное матричное уравнение имеет единственное решение  $M_{\xi_i}$  [18, с. 370]. Тогда можно записать

$$\delta_{fi} = \bar{h}_i^\top \xi_i + \bar{\epsilon}_i, \quad \dot{\xi}_i = \Gamma_i \xi_i,$$

где  $\bar{h}_i^\top = c_{fi}^\top M_{\xi_i}$ , а  $\bar{\epsilon}_i$  экспоненциально затухает. Так как фильтр  $w_i(s)$  является минимально фазовым, то пара  $(\bar{h}_i^\top, \Gamma_i)$  является наблюдаемой. Для

вынужденной составляющей  $\xi_{fi}^*$  состояния  $\xi_{fi}$  фильтра (6) справедливо соотношение  $\xi_{fi}^* = M_{fi}\xi_i$ , где матрица подобия  $M_{fi}$  является решением уравнения Сильвестра

$$(8) \quad M_{fi}\Gamma_i - G_iM_{fi} = l_i\bar{h}_i^\top.$$

Так как спектры матриц  $\Gamma_i$  и  $G_i$  не пересекаются, пара  $(G_i, l_i)$  является полностью управляемой, а пара  $(\Gamma_i, \bar{h}_i)$  является полностью наблюдаемой, то существует единственная невырожденная матрица  $M_{fi}$ , являющаяся решением уравнения Сильвестра (8) [34, с. 240]. Тогда  $\xi_i = M_{fi}^{-1}\xi_{fi}^*$ , откуда с учетом второго уравнения из (2) непосредственно следует справедливость (5) с  $\psi_{fi}^\top = h_i^\top M_{fi}^{-1}$ .

*Замечание 4.* Особо отметим, что для построения фильтра (6) и регрессионной модели (5) решение матричных уравнений (7) и (8) не требуется и достаточно доказательства факта существования обратимой матрицы  $M_{fi}$ .

### 3.2. Фильтрованное несмешанное возмущение

Введем в рассмотрение динамический блок вида

$$(9) \quad \dot{\hat{x}} = A\hat{x} + Bu + L_y(y - C\hat{x}),$$

где  $\hat{x} \in \mathbb{R}^n$  – вектор состояния с произвольным начальным условием  $\hat{x}(0)$ , а матрица  $L_y$  выбрана так, что матрица  $A_L = A - L_yC$  является гурвицевой и ее корни не совпадают с инвариантными нулями системы (1). Тогда для вектора невязки  $\varepsilon = x - \hat{x}$  имеем

$$(10) \quad \dot{\varepsilon} = A_L\varepsilon + B\delta, \quad \varepsilon(0) = x(0) - \hat{x}(0).$$

Определим *фильтрованное возмущение* выражением

$$(11) \quad \delta_f = y - C\hat{x},$$

для которого справедливо представление  $\delta_f = C\varepsilon$ . Тогда  $\delta_f = W_L(s)[\delta] + \epsilon_f$ , где  $W_L(s) = C(sI - A_L)^{-1}B$ , а  $\epsilon_f = Ce^{A_L t}\varepsilon(0)$  экспоненциально затухает.

Если  $m = 1$ , то на основании выхода наблюдателя (9) может быть сформирован регрессор модели (5). Исследуем, можно ли непосредственно распространить утверждение леммы 1 на случай многомерных систем, заменив фильтр со скалярным входом (6) на фильтр с векторным входом

$$\dot{\xi}_f = G\xi_f + L\delta_f,$$

где  $G = \text{diag}\{G_i\}$ ,  $L = \text{diag}\{l_i\}$ . В рассматриваемом случае векторное фильтрованное возмущение  $\delta_f$  является выходом модели (без учета экспоненциально затухающего члена)

$$\delta_f = C\varepsilon, \quad \dot{\varepsilon} = A_L\varepsilon + B\delta$$



и уравнение Сильвестра (7) принимает вид

$$(12) \quad M_\xi \Gamma - A_L M_\xi = B H,$$

и имеет единственное решение  $M_\xi$  при принятых условиях. Тогда фильтрованное векторное возмущение  $\delta_f$  может быть представлено в виде

$$\delta_f = \bar{H} \xi + \epsilon \quad \dot{\xi} = \Gamma \xi,$$

где  $\epsilon$  экспоненциально затухает, а матрица  $\bar{H} = C M_\xi$  в общем случае не является блочно-диагональной. При этом уравнение (8) принимает вид

$$(13) \quad M_f \Gamma - G M_f = L \bar{H}.$$

Как известно, в случае многомерных систем (см. [34, с. 259] и пример 2.9 на стр. 54 в [26]) условие, что матрицы  $\Gamma$  и  $G$  не имеют общих собственных чисел, а пары  $(G, L)$  и  $(\Gamma, \bar{H})$  являются полностью управляемыми и наблюдаемыми, является необходимым, но не достаточным для существования невырожденного решения  $M_f$ . В случае вырожденности матрицы  $M_f$  не существует параметризации  $\delta = \Psi \xi_f + \epsilon$ , где  $\Psi = H M_f^{-1}$ . Это не позволяет непосредственно расширить лемму 1 на случай векторного возмущения.

Возможным путем преодоления этой проблемы является приведение матрицы  $\bar{H}$  к блочно-диагональному виду. Тогда уравнение (13) распадается на  $t$  независимых уравнений типа (8), гарантированно имеющих невырожденные решения.

Воспользуемся следующим вспомогательным утверждением.

*Лемма 2. При условии выполнения допущений Д.1.1–Д.1.3 существует (возможно, не единственная) физически реализуемая асимптотически устойчивая  $t \times t$  передаточная матрица  $Q(s)$  такая, что передаточная матрица  $D(s) = Q(s) W_L(s)$  является асимптотически устойчивой минимально-фазовой блочно-диагональной. Одна из возможных реализаций матрицы  $Q(s)$  определяется выражением*

$$(14) \quad Q(s) = \text{adj } W_L(s),$$

где  $\text{adj } W_L(s)$  – присоединенная к  $W_L(s)$  матрица. При этом  $D(s) = \text{diag} \left\{ \frac{\beta(s)}{\alpha(s)} \right\}$ , где  $\frac{\beta(s)}{\alpha(s)} = \det W_L(s)$ .

*Доказательство.* Так как объект (1) является полностью управляемым и наблюдаемым, матрицы  $B$  и  $C$  имеют полные ранги, а матрица  $L$  выбрана таким образом, что собственные числа  $A_L$  не совпадают с инвариантными нулями системы (1)<sup>2</sup>, то  $\det W_L(s) \neq 0$  и существует  $W_L^{-1}(s) = \frac{\alpha(s)}{\beta(s)} \text{adj } W_L(s)$ . При этом в силу асимптотической устойчивости  $W_L(s)$

---

<sup>2</sup> Как известно, обратная связь по состоянию – в данном случае это  $L_y C x$  – не влияет на нули замкнутой системы [34, с. 237].

и минимальной фазовости объекта (1) полиномы  $\beta(s)$  и  $\alpha(s)$  являются гурвицевыми, а  $\text{adj } W_L(s)$  – асимптотически устойчивой [35, с. 7]. Тогда из равенства  $W_L^{-1}(s)W_L(s) = \frac{1}{\det W_L(s)}\text{adj } W_L(s)W_L(s) = I$  следует справедливость (14). Непосредственные вычисления матрицы  $Q(s)$  для других блочно-диагональных форм  $D(s)$  показывают, что данное решение является не единственным (см. пример в разделе 6 “Результаты моделирования”).

Выберем матрицу *последовательного компенсатора*  $Q(s)$  (например, в форме (14)) и сформируем *фильтрованное несмешанное возмущение*

$$(15) \quad \bar{\delta}_f = Q(s) [y - C\hat{x}].$$

Для  $\bar{\delta}_f$  можно записать

$$(16) \quad \bar{\delta}_f = \text{diag}\{w_i(s)\}[\delta], \quad i = 1, \dots, m,$$

или

$$\bar{\delta}_{fi} = w_i(s)[\delta_i],$$

где  $w_i(s)$  – минимально фазовые асимптотически устойчивые передаточные функции (например, при  $Q(s) = \text{adj } W_L(s)$  имеем  $w_i(s) = \frac{\beta(s)}{\alpha(s)}$  для всех  $i = 1, \dots, m$ ), а  $\bar{\delta}_{fi}$  –  $i$ -я координата  $\bar{\delta}_f$ .

### 3.3. Параметризация исходного многомерного возмущения

Однако на объект (1) воздействует не фильтрованное несмешанное  $\bar{\delta}_f$ , а исходное возмущение  $\delta$ . С учетом леммы 1 и (16) делаем вывод о справедливости следующего утверждения.

*Лемма 3. При условии выполнения допущений Д.1.1–Д.1.4 и Д.2.1–Д.2.3 возмущение  $\delta$ , действующее на объект (1), может быть представлено в виде регрессионной модели*

$$(17) \quad \delta = \Theta \bar{\xi}_f + \bar{\epsilon},$$

где  $\bar{\epsilon}$  экспоненциально затухает, регрессор  $\bar{\xi}_f$  является вектором состояния наблюдателя

$$(18) \quad \dot{\bar{\xi}}_f = G \bar{\xi}_f + LQ(s) [y - C\hat{x}]$$

с произвольным начальным условием  $\bar{\xi}_f(0)$ , матрицы  $G$ ,  $L$  и  $\Theta$  являются блочно-диагональными

$$G = \text{diag}\{G_i\}, \quad L = \text{diag}\{l_i\}, \quad \Theta = \text{diag}\{\psi_{fi}^\top\},$$

а  $\psi_{fi} \in \mathbb{R}^{q_i}$  являются векторами неизвестных постоянных коэффициентов ( $i = 1, \dots, m$ ).

Действительно, динамический фильтр (18) с многомерным входом  $\bar{\delta}_f$  (15) распадается на  $m$  независимых фильтров типа (6) со скалярными входами  $\bar{\delta}_i = w_i(s)[\delta_i]$  и состояниями  $\bar{\xi}_{fi}$ , что соответствует условиям леммы 1.

Таким образом, лемма 3 сводит неопределенность возмущения  $\delta$  к параметрической неопределенности регрессионной модели (17) с матрицей неизвестных параметров  $\Theta$  и физически реализуемым регрессором  $\bar{\xi}_f$ . При этом модель объекта управления (1) может быть записана в виде

$$(19) \quad \begin{cases} \dot{x} = Ax + B(u + \Xi\theta + \bar{\epsilon}), & x(0), \\ y = Cx, \end{cases}$$

где  $\Xi = \text{diag}\{\bar{\xi}_{fi}^\top\} \in \mathbb{R}^{m \times q}$  – блочно-диагональный матричный регрессор, имеющий на главной диагонали  $m$  векторных блоков  $\bar{\xi}_{fi}^\top$ ,  $\theta = [\psi_{f1}^\top, \dots, \psi_{fm}^\top]^\top \in \mathbb{R}^q$  – вектор неизвестных параметров, составленный из диагональных блоков  $\psi_{fi}$  матрицы  $\Theta$ .

*Замечание 5.* Альтернативным, но не оптимальным с точки зрения динамического порядка наблюдателя возмущения (а следовательно, и числа настраиваемых параметров в адаптивном регуляторе), является подход, при котором каждая компонента  $\delta_{fi}$  фильтрованного возмущения  $\delta_f$  рассматривается в качестве выхода командного генератора полной размерности (подобный подход использован в [18, 20, 28]). Действительно, для  $\delta_{fi}$  можно записать

$$(20) \quad \begin{aligned} \delta_{fi} &= \sum_{j=1}^m w_{Lij}(s)[\delta_j] = \sum_{j=1}^m h_j^\top w_{Lij}(s)[\xi_j] = \sum_{j=1}^m h_j^\top M_{ij}\xi_j + \epsilon_{ij} = \\ &= \bar{h}_i^\top \xi + \epsilon, \quad i = 1, \dots, m, \end{aligned}$$

где  $w_{Lij}(s)$  – элементы передаточной матрицы  $W_L(s)$ ,  $M_{ij}$  – матрицы подобию для установившихся составляющих,  $\bar{h}_i^\top = [h_1^\top M_{i1}, \dots, h_m^\top M_{im}] \in \mathbb{R}^q$  – вектор неизвестных параметров,  $\xi \in \mathbb{R}^q$  – вектор состояния модели (3). Выражение (20) мотивирует использование наблюдателя возмущения размерности  $q$  в каждом канале системы управления, что ведет к общему динамическому порядку наблюдателя  $m \times q$  и к такому же числу настраиваемых параметров. В то же время предложенная модель (17) может быть получена с помощью наблюдателя  $q$ -го порядка (при статическом компенсаторе  $Q(s)$  – см. пример), содержит  $q$  неизвестных параметров и, следовательно,  $q$  настраиваемых параметров в адаптивном регуляторе.

Анализ модели (19) мотивирует следующий выбор структуры алгоритма управления:

$$(21) \quad u = u_x + u_\delta,$$

где  $u_x$  – стабилизирующая компонента управления, а  $u_\delta$  – компенсирующая возмущение компонента управления. Раздельному синтезу этих компонент посвящены следующие две секции настоящей статьи.

#### 4. Синтез алгоритма адаптации

Выбирая

$$(22) \quad u_\delta = -\Xi\hat{\theta},$$

где  $\hat{\theta}$  – вектор настраиваемых параметров, и подставляя (21) и (22) в (19), получаем модель ошибки управления

$$(23) \quad \begin{cases} \dot{x} = Ax + B(u_x + \Xi\tilde{\theta} + \bar{\epsilon}), & x(0), \\ y = Cx \end{cases}$$

с вектором параметрических ошибок  $\tilde{\theta} = \theta - \hat{\theta}$ .

Так как вектор состояния  $x$  не доступен измерениям, а матрица  $A$  может быть неустойчивой, то для синтеза алгоритма адаптивной настройки параметров  $\hat{\theta}$  сформируем статическую регрессионную модель, определяемую следующим утверждением.

*Лемма 4. Введем в рассмотрение расширенную ошибку*

$$(24) \quad \bar{y} = y - C\hat{x} - \Xi_f\hat{\theta},$$

где вектор  $\hat{x}$  генерируется фильтром (9),  $\Xi_f \in \mathbb{R}^{m \times qm}$  – матричный регрессор вида

$$(25) \quad \Xi_f = \begin{bmatrix} w_{L11}(s) \left[ \bar{\xi}_{f1}^\top(t) \right] & \dots & w_{L1m}(s) \left[ \bar{\xi}_{fm}^\top(t) \right] \\ \vdots & \ddots & \vdots \\ w_{Lm1}(s) \left[ \bar{\xi}_{f1}^\top(t) \right] & \dots & w_{Lmm}(s) \left[ \bar{\xi}_{fm}^\top(t) \right] \end{bmatrix},$$

а  $w_{Lij}(s)$  – элементы асимптотически устойчивой передаточной функции  $W_L(s) = C(sI - A_L)^{-1}B$  ( $i = 1, \dots, m, j = 1, \dots, m$ ). Тогда для сигнала  $\bar{y}$  справедливо следующее равенство:

$$(26) \quad \bar{y} = \Xi_f\tilde{\theta} + \bar{\epsilon}_f,$$

где  $\bar{\epsilon}_f$  экспоненциально затухает.

Доказательство леммы 4 приведено в Приложении.

Модель (26) хорошо известна в современной теории адаптивных систем [36, 37] и идентификации [38] и позволяет использовать широкий класс стандартных алгоритмов адаптации. В том числе:

– *градиентный алгоритм адаптации* [19, 36, 37]

$$(27) \quad \dot{\hat{\theta}} = \gamma \Xi_f^\top \bar{y},$$

– *алгоритм с улучшенной параметрической сходимостью* [19, 39]

$$(28) \quad \dot{\hat{\theta}} = \gamma \left( d(s) \left[ \Xi_f^\top \bar{y} \right] - \Omega \hat{\theta} \right),$$

где  $\gamma > 0$  – коэффициент адаптации,  $\Omega = d(s)[\Xi_f^T \Xi_f]$  – расширенный матричный регрессор,  $d(s)$  – минимально фазовая асимптотически устойчивая положительная передаточная функция с действительными корнями характеристического уравнения и единичным статическим коэффициентом (т.е.  $d(0) = 1$ ). В простейшем случае  $d(s)$  может быть выбран в виде аperiодического звена первого порядка. Свойства сходимости алгоритмов адаптации (27) и (28) определяются следующим утверждением, доказательство которого приведено в Приложении.

*Лемма 5. При выполнении допущений Д.1.1–Д.1.4 и Д.2.1–Д.2.3*

**Л.5.1.** алгоритмы адаптации (27) и (28) обеспечивают ограниченность  $\bar{y}$  и  $\hat{\theta}$ , а также асимптотическое стремление  $|\Xi(t)\tilde{\theta}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ ;

**Л.5.2.** если  $\lambda(t) \notin \mathcal{L}_1$ , где  $\lambda(t)$  – минимальное собственное значение матрицы  $\Omega(t)$ , то алгоритм адаптации (28) дополнительно к свойству Л.5.1 обеспечивает асимптотическую сходимость  $|\tilde{\theta}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ ;

**Л.5.3.** если регрессор  $\xi$  удовлетворяет условию исчезающего возбуждения (в смысле определения 4.3.1 из [36] или определения 3.4 из [19]), то алгоритмы адаптации (27) и (28) дополнительно к свойству Л.5.1 обеспечивают экспоненциальную сходимость  $|\tilde{\theta}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ .

## 5. Синтез стабилизирующей компоненты и свойства замкнутой системы

Так как вектор  $x$  не доступен прямым измерениям, используем наблюдатель состояния вида

$$(29) \quad \dot{\bar{x}} = A\bar{x} + Bu_x + L_y(y - C\bar{x}),$$

где  $\bar{x}$  – вектор состояния наблюдателя с произвольным начальным условием  $\bar{x}(0)$ . Введем в рассмотрение вектор невязки  $\tilde{x} = x - \bar{x}$ . Вычитая (29) из (23), получаем

$$\dot{\tilde{x}} = A_L\tilde{x} + B(\Xi\tilde{\theta} + \bar{\epsilon}).$$

В силу свойства Л.5.1 леммы 5 имеем  $|\Xi(t)\tilde{\theta}(t)| \rightarrow 0$ , следовательно,  $|\tilde{x}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ .

Сформируем стабилизирующее управление вида

$$(30) \quad u_x = -K\bar{x},$$

где матрица обратных связей такая, что матрица  $A_K = A - BK$  является гурвицевой. Подставляя (30) в (23), получаем модель замкнутой системы

$$\dot{x} = A_Kx + B(\Xi\tilde{\theta} - K\tilde{x} + \bar{\epsilon}).$$

Так как  $|\tilde{x}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ , то немедленно получаем, что  $|x(t)| \rightarrow 00$  при  $t \rightarrow \infty$ , что означает выполнение целевого условия (4). Таким образом, доказали следующее утверждение.

*Теорема 1. При выполнении допущений Д.1.1–Д.1.4 и Д.2.1–Д.2.3 закон управления (21), содержащий компенсирующую составляющую (22), стабилизирующую составляющую (30), наблюдатель возмущения (9), (18), алгоритм адаптации (27) или (28) и наблюдатель состояния (29), применительно к объекту управления (1), обеспечивает ограниченность всех сигналов и выполнение целевого условия (4).*

## 6. Пример и результаты моделирования

Рассмотрим неустойчивый объект управления (1) с матрицами

$$A = \begin{bmatrix} 4 & -6 & 0 \\ 4,5 & -7 & 0 \\ 12 & -20 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 2 & 0 \\ 4 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} -2 & 4 & -1 \\ 3 & -5 & 1 \end{bmatrix}$$

и передаточной матрицей

$$W(s) = C(sI - A)^{-1}B = \begin{bmatrix} 0 & \frac{-1}{s-1} \\ 1 & \frac{1}{s-1} \\ \frac{1}{s^2 + 3s - 1} & \frac{1}{s-1} \end{bmatrix}.$$

В данном случае объект не имеет инвариантных нулей и поэтому является минимально-фазовым.

Пусть  $\delta = [\delta_1, \delta_2]^\top = [\sin t, 1]^\top$  – неизмеряемое и заранее неизвестное возмущение, первая компонента которого может рассматриваться в качестве выхода экзосистемы (2) второго порядка, а вторая компонента – экзосистемы первого порядка. В этом случае можно записать

$$\Gamma = \begin{bmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} h_1^\top & 0 \\ 0 & h_2^\top \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Выберем

$$L_y = \begin{bmatrix} 6 & 6 \\ 6 & 6 \\ 10 & 12 \end{bmatrix}.$$

Тогда

$$A_L = \begin{bmatrix} -2 & 0 & 0 \\ -1,5 & -1 & 0 \\ -4 & 0 & -1 \end{bmatrix}, \quad W_L(s) = \begin{bmatrix} 0 & \frac{-1}{s+1} \\ 1 & \frac{1}{s+1} \\ \frac{1}{(s+1)(s+2)} & \frac{1}{s+1} \end{bmatrix}.$$

Наконец, для построения наблюдателя возмущения выберем

$$G = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -6 & -5 & 0 \\ 0 & 0 & -4 \end{bmatrix}, \quad L = \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 6 & 0 \\ 0 & 4 \end{bmatrix}.$$

В методических целях вычислим матрицы  $M_\xi$ ,  $\bar{H}$  и  $M_f$ , являющиеся решениями уравнений (12) и (13) соответственно. Получаем неблочно-диагональную матрицу

$$\bar{H} = CM_\xi = \begin{bmatrix} 0 & 0 & -1 \\ 0,1 & -0,3 & 1 \end{bmatrix}$$

и вырожденную матрицу

$$M_f = \begin{bmatrix} 0 & 0 & -0,33 \\ 0 & 0 & 0 \\ -0,04 & -0,28 & 1 \end{bmatrix}.$$

Таким образом, проиллюстрировано, что в рассматриваемом случае прямое расширение леммы 1 на векторное возмущение невозможно без применения специальных методов.

В соответствии с предложенным в статье подходом, рассчитаем передаточную матрицу последовательного компенсатора

$$(31) \quad Q(s) = \text{adj } W_L(s) = \begin{bmatrix} \frac{1}{s+1} & \frac{1}{s+1} \\ \frac{-1}{(s+1)(s+2)} & 0 \end{bmatrix}$$

и сформируем фильтрованное несмешанное возмущение (15). Тогда можно показать, что для  $\bar{\delta}_f$  справедлива следующая модель:

$$\bar{\delta}_f = \begin{bmatrix} \frac{1}{(s+1)^2(s+2)} & 0 \\ 0 & \frac{1}{(s+1)^2(s+2)} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{(s+1)^2(s+2)} [\delta_1] \\ \frac{1}{(s+1)^2(s+2)} [\delta_2] \end{bmatrix}.$$

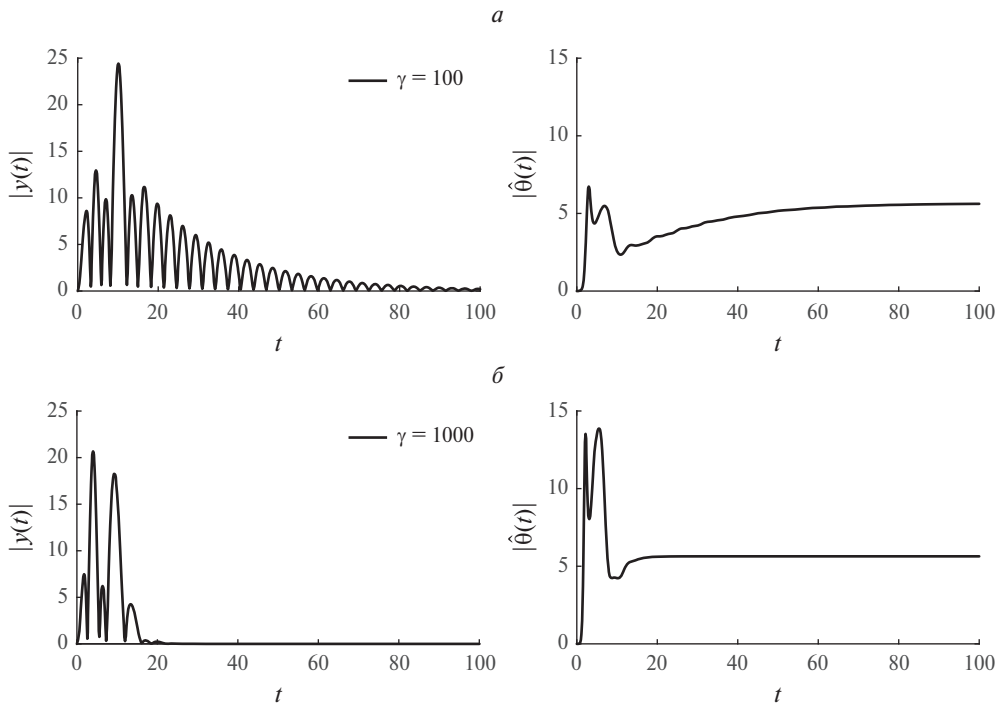
Однако, как было указано в лемме 2, компенсатор (31) не является единственным, приводящим возмущение к несмешанной форме. Действительно, в рассматриваемом примере простой последовательный компенсатор

$$(32) \quad Q(s) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$$

дает следующую модель несмешанного фильтрованного возмущения:

$$\bar{\delta}_f = \begin{bmatrix} \frac{1}{(s+1)(s+2)} & 0 \\ 0 & \frac{1}{s+1} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{(s+1)(s+2)} [\delta_1] \\ \frac{1}{s+1} [\delta_2] \end{bmatrix}.$$

При моделировании будем использовать последовательный компенсатор (32).



Переходные процессы в адаптивной системе, замкнутой алгоритмом адаптации с улучшенной параметрической сходимостью при:  $a - \gamma = 100$ ;  $b - \gamma = 1000$ .

Результаты моделирования замкнутой системы с алгоритмом адаптации с улучшенной сходимостью (28) при  $d(s) = \frac{1}{s+1}$  и двух различных значениях  $\gamma$ , стабилизирующим управлением (30) с

$$K = \begin{bmatrix} 25,0 & -0,91 & -8,38 \\ 36,25 & -4,83 & -10,67 \end{bmatrix},$$

при нулевых начальных условиях в объекте и алгоритме управления представлены на рисунке и демонстрируют достижение цели управления (4) в условиях действия внешнего заранее неизвестного возмущения, а также возможность ускорения процессов настройки регулятора за счет увеличения коэффициента адаптации  $\gamma$ .

## 7. Заключение

В работе представлено решение задачи адаптивной компенсации внешних неизвестных детерминированных возмущений при управлении многоканальным линейным объектом по выходу. Предлагаемое решение основано на адаптивной реализации метода внутренней модели и построении наблюдателя специальной формы, позволяющего получить несмешанную параметризацию



возмущения. Представленный подход дает возможность синтезировать адаптивный регулятор с числом настраиваемых параметров, равным числу неизвестных коэффициентов характеристического полинома модели возмущения, тем самым обеспечивая понижение вычислительной сложности алгоритма по сравнению с известными решениями.

## ПРИЛОЖЕНИЕ

*Доказательство леммы 4.* Перепишем (25) в форме “вход–состояние–выход”

$$(П.1) \quad \begin{cases} \dot{\Xi}_x = A_L \Xi_x + B \Xi, \\ \Xi_f = C \Xi_x. \end{cases}$$

Введем промежуточную переменную

$$(П.2) \quad \epsilon_x = x - \hat{x} - \Xi_x \theta$$

и рассчитаем ее производную по времени в силу (19), (9) и (П.1):

$$\begin{aligned} \dot{\epsilon}_x = x - \dot{\hat{x}} - \dot{\Xi}_x \theta &= Ax + Bu + B \Xi \theta + \bar{\epsilon} - A \hat{x} - Bu - L_y(Cx - C \hat{x}) - \\ &- A_L \Xi_x \theta - B \Xi \theta = A_L \epsilon_x + \bar{\epsilon}. \end{aligned}$$

В силу гурвицевости матрицы  $A_L$  величина  $\epsilon_x(t)$  экспоненциально затухает.

Тогда  $x - \hat{x} = \Xi_x \theta + \epsilon_x$  или

$$(П.3) \quad y - C \hat{x} = \Xi_f \theta + \bar{\epsilon}_f,$$

где  $\bar{\epsilon}_f = C \epsilon_x$ . Подставляя (П.3) в (24), непосредственно получаем (26).

*Доказательство леммы 5.* Выберем функцию Ляпунова

$$(П.4) \quad V = \frac{1}{2\gamma} \tilde{\theta}^\top \tilde{\theta} + \int_t^\infty \bar{\epsilon}_f^\top(\tau) \bar{\epsilon}_f(\tau) d\tau$$

и вычислим ее производную в силу (26), равенства  $\dot{\hat{\theta}} = -\dot{\hat{\theta}}$  и поочередно с учетом алгоритмов (27) и (28):

1. Алгоритм (27).

$$\begin{aligned} \dot{V} &= -\tilde{\theta}^\top \Xi_f^\top \Xi_f \tilde{\theta} - \tilde{\theta}^\top \Xi_f^\top \bar{\epsilon}_f - \bar{\epsilon}_f^\top \bar{\epsilon}_f = \\ &= -\frac{1}{2} |\Xi_f \tilde{\theta}|^2 - \frac{1}{2} |\bar{\epsilon}_f| - \frac{1}{2} \left( |\Xi_f \tilde{\theta}| + |\bar{\epsilon}_f| \right)^2 \leq -\frac{1}{2} |\Xi_f \tilde{\theta}|^2. \end{aligned}$$

Из последнего следует ограниченность  $\hat{\theta}$  и  $|\Xi_f \tilde{\theta}| \in \mathcal{L}_2$ . В силу ограниченности  $\Xi$  и устойчивости фильтра  $W_L(s)$  выход линейной регрессии (26)  $\bar{y}$  и производная  $\dot{\hat{\theta}}$  ограничены. Так как функция  $|\Xi_f(t) \tilde{\theta}(t)| \in \mathcal{L}_2$  и является непрерывной и ограниченной, то  $|\Xi_f(t) \tilde{\theta}(t)|, \dot{\hat{\theta}}(t) \rightarrow 0$  при  $t \rightarrow \infty$ .

Согласно лемме о перестановке, применяемой для многоканальных систем (см. [19, раздел D.2]), имеем

$$(II.5) \quad W_L(s) \left[ \Xi \tilde{\theta} \right] = \Xi_f \tilde{\theta} + W_L(s) \left[ \Xi_f \dot{\tilde{\theta}} \right],$$

где  $\Xi_f$  – выход фильтра (II.1) из Приложения. Из последнего выражения следует, что в силу ограниченности  $\Xi(t)$ , устойчивости  $W_L(s)$  и сходимостей  $|\Xi_f(t)\tilde{\theta}(t)|$ ,  $\dot{\tilde{\theta}}(t) \rightarrow 0$  справедлива сходимость  $W_L(s) \left[ \Xi(t)\tilde{\theta}(t) \right] \rightarrow 0$  при  $t \rightarrow \infty$ . Так как  $W_L(s)$  невырождена и минимально фазовая, а  $\dot{\Xi}$ ,  $\dot{\tilde{\theta}}$  ограничены, то  $|\Xi(t)\tilde{\theta}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ .

2. Алгоритм (28). Проанализируем производную функции  $V$  с учетом неравенства Коши–Буняковского:

$$\begin{aligned} \dot{V} &= -\tilde{\theta}^\top \Omega \tilde{\theta} - \tilde{\theta}^\top d(s) \left[ \Xi_f^\top \bar{\epsilon}_f \right] - \bar{\epsilon}_f^\top \bar{\epsilon}_f = \\ &= -\tilde{\theta}^\top \Omega \tilde{\theta} - \tilde{\theta}^\top(t) \int_0^t h(t-\tau) \Xi_f^\top(\tau) \bar{\epsilon}_f(\tau) d\tau - |\bar{\epsilon}_f|^2 \leq \\ &\leq -\tilde{\theta}^\top \Omega \tilde{\theta} + \left( \int_0^t h(t-\tau) \left( \tilde{\theta}(t) \Xi_f^\top(\tau) \right)^2 d\tau \right)^{\frac{1}{2}} \left( \int_0^t h(t-\tau) \bar{\epsilon}_f^2(\tau) d\tau \right)^{\frac{1}{2}} - |\bar{\epsilon}_f|^2, \end{aligned}$$

где  $h(t-\tau) \geq 0$  – импульсная функция фильтра  $d(s)$ . Принимая во внимание, что

$$\int_0^t h(t-\tau) \left( \tilde{\theta}(t) \Xi_f^\top(\tau) \right)^2 d\tau \leq c_d \tilde{\theta}^\top(t) \Omega(t) \tilde{\theta}(t), \quad \int_0^t h(t-\tau) \bar{\epsilon}_f^2(\tau) d\tau \leq c_d |\bar{\epsilon}_f|^2,$$

где  $c_d = |d(s)|_\infty = 1$ , продолжаем анализ  $\dot{V}$ :

$$(II.6) \quad \dot{V} \leq -\tilde{\theta}^\top \Omega \tilde{\theta} + \left( \tilde{\theta}^\top \Omega \tilde{\theta} \right)^{\frac{1}{2}} |\bar{\epsilon}_f| - |\bar{\epsilon}_f|^2 \leq -\frac{1}{2} \tilde{\theta}^\top \Omega \tilde{\theta}.$$

Из последнего следует ограниченность  $\tilde{\theta}$  и  $|\Omega^{\frac{1}{2}} \tilde{\theta}| \in \mathcal{L}_2$ . Как следствие, в силу устойчивости фильтров  $W_L(s)$  и  $d(s)$  функции  $\Omega(t)$  и  $\tilde{\Omega}$  ограничены, откуда следует  $|\Omega \tilde{\theta}| \in \mathcal{L}_2$  и  $|\Omega(t)\tilde{\theta}(t)| \rightarrow 0$  при  $t \rightarrow \infty$ . Так как  $\dot{\tilde{\theta}}(t) \rightarrow 0$ , то (см. лемму 1 о перестановке в [40])  $d(s) \left[ \Xi_f^\top(t) \Xi_f(t) \tilde{\theta}(t) \right] \rightarrow 0$  при  $t \rightarrow \infty$ . Так как  $\dot{\Xi}_f$  и  $\dot{\tilde{\theta}}$  ограничены, то из сходимости  $d(s) \left[ \Xi_f^\top(t) \Xi_f(t) \tilde{\theta}(t) \right] \rightarrow 0$  следует сходимость  $\Xi_f(t) \tilde{\theta}(t) \rightarrow 0$  при  $t \rightarrow \infty$ . Учитывая рассуждения, приведенные выше для доказательства свойства Л.5.1 для алгоритма (27), из сходимости  $\Xi_f(t) \tilde{\theta}(t) \rightarrow 0$  имеем сходимость  $\Xi(t) \tilde{\theta}(t) \rightarrow 0$  при  $t \rightarrow \infty$ .

Свойство Л.5.1 доказано.

Для доказательства свойства Л.5.2 продолжим анализ производной (П.6) с учетом (П.4):

$$\dot{V} \leq -\frac{1}{2}\tilde{\theta}^\top \Omega \tilde{\theta} \leq -\frac{1}{2}\lambda(t)\tilde{\theta}^\top \tilde{\theta} = -\gamma V + \gamma \int_t^\infty \bar{\epsilon}_f^\top(\tau)\bar{\epsilon}_f(\tau)d\tau.$$

Решая полученное дифференциальное неравенство, имеем

$$V(t) \leq e^{-\gamma \int_0^t \lambda(\tau_1)d\tau_1} V(0) + \gamma \int_0^t e^{-\gamma \int_{\tau_3}^t \lambda(\tau_2)d\tau_2} \int_{\tau_3}^\infty \bar{\epsilon}_f^\top(\tau_1)\bar{\epsilon}_f(\tau_1)d\tau_1 d\tau_3,$$

откуда следует свойство Л.5.2.

Для доказательства свойства Л.5.3 для алгоритма (27) см. [37, раздел 2.8]. Для доказательства свойства Л.5.3 для алгоритма (28) см. [19, раздел 3.2.3].

## СПИСОК ЛИТЕРАТУРЫ

1. *Landau I.D., Airimitoae T.-B., Castellanos-Silva A., et al.* Adaptive and Robust Active Vibration Control: Methodology and Tests. Springer, 2017.
2. *Zapateiro de la Hoz M., Pozo F. (Eds.)* Advances on Analysis and Control of Vibrations-Theory and Applications. InTech., 2012.
3. *Fuller C., Elliott S., Nelson P.* Active Control of Vibration. Academic, Cambridge, 1999.
4. *Njeh M., Cauet S., Coirault P.* Persistent disturbances rejection on internal combustion engine torque in hybrid electric vehicles // In Proc of IEEE Conference on Decision and Control. 2010. P. 6421–6426.
5. *Orivuori J., Zazas I., Daley S.* Active control of frequency varying disturbances in a diesel engine // Control Engineering Practice. 2012. V. 20. P. 1206–1219.
6. *Kugi A., Hass W., Schlacher K., et al.* Active compensation of roll eccentricity in rolling mills // IEEE Trans. Industrial Application. 2000. V. 36. P. 625–632.
7. *Shin K., Jang J., Kang J., and Song S.* Compensation method for tension disturbance due to an unknown roll shape in a web transport system // IEEE Trans. Industrial Application. 2003. V. 39. P. 1422–1428.
8. *Furtmuller C., Colaneri P., Rea L.* Adaptive robust stabilization of continuous casting // Automatica. 2012. V. 48. P. 225–232.
9. *Liu J.-J., Yang Y.-P.* Stability of the frequency adaptive control technique and its application to compact disk drives // Control Engineering Practice. 2005. V. 13. No. 5. P. 629–639.
10. *Wu Z., Li Y., Wang P., Liu M.* Dynamic head-disk interface modeling and adaptive control of a hybrid actuator for optical data storage systems // Int. J. Optomechatronics. 2015. V. 9. P. 62–88.
11. *Burdisso R.A., Fuller C.R., Suarez L.E.* Adaptive feedforward control of structures subjected to seismic excitations // In Proc. of the American Control Conf. 1993. P. 2104–2108.

12. *Messineo S., Celani F., Egeland O.* Crane feedback control in offshore moonpool operations // *Control Engineering Practice*. 2008. V. 16. P. 356–364.
13. *Messineo S., Serrani A.* Offshore crane control based on adaptive external models // *Automatica*. 2009. V. 9. P. 2546–2556.
14. *Pham H., Jung H., Hu T.* State-space approach to modeling and ripple reduction in AC-DC converters // *IEEE Trans. Control Systems Technology*. 2013. V. 21. No. 5. P. 1949–1955.
15. *Maier S., Bals J., Bodson M.* Periodic disturbance rejection of a PMSM with adaptive control algorithms // *Proc. the IEEE Int. Electric Machines and Drives Conf.* 2011. P. 1070–1075.
16. *Francis D.A., Wonham W.M.* The internal model principle for linear multivariable regulators // *Appl. Math. Optim.* 1975. V. 2. P. 170–194.
17. *Davison E.J.* The robust control of a servomechanism problem for linear time-invariant multivariable systems // *IEEE Trans. Autom. Control*. 1976. V. 21. P. 25–34.
18. *Isidori A.* *Lectures in Feedback design for Multivariable systems.* Switzerland: Springer, 2017.
19. *Nikiforov V., Gerasimov D.* *Adaptive Regulation. Reference Tracking and Disturbance Rejection.* Switzerland: Springer, 2022.
20. *Pyrkin A., Isidori A.* Adaptive output regulation of right-invertible MIMO LTI systems, with application to vessel motion control // *Eur. J. Control*. 2019. V. 46. P. 63–79.
21. *Marino R., Tomei P.* Output regulation for linear systems via adaptive internal model // *IEEE Trans. Autom. Control*. 2003. V. 48. No. 12. P. 2199–2202.
22. *Бобцов А.А.* Алгоритм управления по выходу с компенсацией гармонического возмущения со смещением // *АиТ*. 2008. № 8. С. 25–32.  
*Bobtsov A.A.* Output control algorithm with the compensation of biased harmonic disturbances // *Autom. Remote Control*. 2008. V. 69. No. 8. P. 1289–1296.
23. *Бобцов А.А., Пыркин А.А.* Компенсация неизвестного синусоидального возмущения для линейного объекта любой относительной степени // *АиТ*. 2009. № 3. С. 114–122.  
*Bobtsov A.A., Pyrkin A.A.* Compensation of unknown sinusoidal disturbances in linear plants of arbitrary relative degree // *Autom. Remote Control*. 2009. V. 70. No. 3. P. 449–456.
24. *Yilmaz C.T., Basturk H.I.* Output feedback control for unknown LTI systems driven by unknown periodic disturbances // *Automatica*. 2019. V. 99. P. 112–119.
25. *Marino R., Santosuosso G.L.* Regulation of linear systems with unknown exosystems of uncertain order // *IEEE Trans. Autom. Control*. 2007. V. 52(2). P. 352–359.
26. *Никифоров В.О., Парамонов А.В., Герасимов Д.Н.* Алгоритмы адаптивного регулирования в многоканальных линейных системах с запаздыванием по управлению // *АиТ*. 2020. № 6. С. 153–172.  
*Nikiforov V.O., Paramonov A.V., Gerasimov D.N.* Adaptive control algorithms in MIMO linear systems with control delay // *Autom. Remote Control*. 2020. V. 81. No. 6. P. 1091–1106.
27. *Obregon-Pulido G., Castillo-Toledo B., Loukianov A.* A Structurally Stable Globally Adaptive Internal Model Regulator for MIMO Linear Systems // *IEEE Trans. Autom. Control*. 2011. V. 56. No. 11. P. 160–165.

28. *Borisov O., Isidori A., Pyrkin A.* Adaptive output regulation of MIMO LTI systems with unmodeled input dynamics // 62nd IEEE Conference on Decision and Control. 2023. P. 1537–1542.
29. *Jafari S., Ioannou P.* Rejection of unknown periodic disturbances for continuous-time MIMO systems with dynamic uncertainties // Int. J. Adapt. Control Signal Process. 2016. V. 30(12). P. 1674–1688.
30. *Никифоров В.О.* Наблюдатели внешних возмущений. Часть 1. Объекты с известными параметрами // АиТ. 2004. № 10. С. 13–23.
31. *Никифоров В.О.* Адаптивное и робастное управление с компенсацией возмущений. СПб.: Наука, 2003.
32. *Lyubomir T. Gruyitch.* Control of Linear Systems. Observability and Controllability of General Linear Systems. U.S.: Taylor and Francis, 2019.
33. *Hespanha J.P.* Linear Systems Theory. Princeton and Oxford: Princeton University Press, 2018.
34. *Chen C-T.* Linear Systems. Theory and design. NY.: Oxford University Press, 1999.
35. *Gasparyan O.* Linear and Nonlinear Multivariable Feedback Control. John Wiley & Sons, Ltd., 2008.
36. *Ioannou P., Sun J.* Robust Adaptive Control. Prentice-Hall, Englewood Cliffs, NJ, 1996.
37. *Narendra K.S., Annaswamy A.M.* Stable Adaptive Systems. Prentice Hall, 1989.
38. *Льюнг Л.* Идентификация систем: Теория для пользователя. М.: Наука, 1991.
39. *Kreisselmeier G.* Adaptive observers with exponential rate of convergence // IEEE Trans. Autom. Control. 1977. V. 22. No. 1. P. 2–8.
40. *Gerasimov D.N., Nikiforov V.O.* On key properties of the Lion's and Kreisselmeier's adaptation algorithms // Int. J. Adapt. Control Signal Proc. 2021. V. 36. No. 6. P. 1285–1304.

*Статья представлена к публикации членом редколлегии С.А. Красновой.*

Поступила в редакцию 10.09.2024

После доработки 28.11.2024

Принята к публикации 28.01.2025

© 2025 г. А.М. ПАШАЕВ, академик НАН Азербайджана (mail@naa.az),  
А.Д. ИСКЕНДЕРОВ, д-р физ.-мат. наук (asaf.iskander@mail.ru)  
(Национальная академия авиации, Баку, Азербайджан),  
М.А. МУСАЕВА, д-р философ. мат. наук (musayeva08@inbox.ru)  
(Азербайджанский государственный педагогический университет, Баку)

## ЗАДАЧА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ КОЭФФИЦИЕНТАМИ НЕЛИНЕЙНОГО УРАВНЕНИЯ ТИПА ШРЕДИНГЕРА

Изучается разрешимость задачи оптимального управления коэффициентом старшей производной и квантовым потенциалом в нелинейном и нестационарном уравнении типа Шредингера, которое обобщает известное квантовомеханическое уравнение. Рассматривается задача одновременного управления несколькими коэффициентами уравнения состояния по критерию качества, являющаяся невязкой граничных данных решения. Для этой задачи найдены условия корректности постановки и доказана теорема существования решения. Рассматривается также задача с «возмущенным» критерием качества, для которой доказана теорема существования и единственности решения. Определен явный вид первой вариации функционала качества и описан итеративный алгоритм решения изучаемых задач. Результаты являются новыми также для стандартного уравнения Шредингера в квантовой механике.

*Ключевые слова:* разрешимость задачи оптимального управления, уравнение типа Шредингера, управление квантовыми процессами, управление коэффициентами уравнения состояния.

DOI: 10.31857/S0005231025040024, EDN: CAZCPL

### 1. Введение

Уравнение типа Шредингера обобщает известное квантовомеханическое уравнение Шредингера и используется как модель в теориях сверхпроводимости и других областях практики. Это уравнение часто возникает при диагностике наноструктурных материалов, в атомно-молекулярных компьютерных расчетах, когда требуется управлять внутриатомными и внутримолекулярными потенциалами взаимодействия, лазерными импульсами и характеристиками материалов. Это уравнение широко применяется в квантовой обработке информации, в адаптивной оптике и квазиоптике, при изучении бифуркации в нелинейных моделях, магнитных квантовых явлениях и в других областях практики (см. [1–9] и др.). Ранее изучались линейные и билинейные квантовые системы, задачи оптимизации этих систем с одним фактором

управления, с реальным или комплексным потенциалом взаимодействия и др. В [9–11] развит метод конечных разностей численного решения таких задач управления.

В современной практике часто возникает необходимость управления многими коэффициентами, наиболее влиятельными квантовыми характеристиками нелинейных и нестационарных процессов [1, 4–10]. Такие задачи управления наиболее трудны для теоретического анализа и численного решения, так как они не только нелинейные, но и в них оператор состояния задается неявно, они часто принадлежат к классу некорректно поставленных задач [9, 10], отсутствуют прямые методы их наблюдения и формулировка критериев качества. Форма задания наблюдения особенно важна для образования критериев качества в квантовых процессах управления, так как она должна соответствовать квантовому характеру изучаемых процессов. В этом смысле граничное наблюдение удобно для измерений, а также для обработки результатов. Изучение процессов управления коэффициентами старших производных уравнения состояния особенно важно [1–3, 6–11]. Ниже изучается задача одновременного оптимального управления коэффициентом старшей производной и квантовым потенциалом в нелинейном нестационарном уравнении типа Шредингера. Рассматривается также задача оптимизации с «возмущенным» функционалом качества, доказаны теоремы существования и единственности решения рассматриваемых задач, описаны итеративные алгоритмы их решения. Основные результаты являются новыми также для обычного уравнения Шредингера в квантовой механике.

## 2. Постановка задачи

Пусть  $l > 0$ ,  $T > 0$  – заданные числа и  $0 \leq x \leq l$ ,  $0 \leq t \leq T$ ,  $\Omega_t = (0, l) \times (0, t)$ ,  $\Omega = \Omega_T$ ,  $\psi(x, t)$  является комплексной волновой функцией с пространственной координатой  $x$  и временем  $t$ . Используемые ниже функциональные пространства введены, например, в [8, 11] и др. Соответственно этому  $L_p(0, l)$  лебегово пространство измеримых на  $(0, l)$  функций, которые интегрируются со степенью  $p \geq 1$ ,  $C^k[0, T; B)$  – банахово пространство  $k \geq 0$  раз непрерывно дифференцируемых на  $[0, T]$  функций, значения которых принадлежат банаховому пространству  $B$ . Пусть  $W_p^k(0, l)$  и  $W_p^{k, m}(\Omega)$  пространства Соболева функций с обобщенными производными порядка  $k \geq 0$  по  $x$  и порядка  $m \geq 0$  по  $t$ , которые интегрируются со степенью  $p \geq 1$ . Ниже через  $W_\infty^1(0, l) = \{w : w \in L_\infty(0, l), \frac{dw}{dx} \in L_\infty(0, l)\}$  обозначим банахово пространство с указанными свойствами. Символ  $\forall^o$  означает, что указанные свойства имеют место почти для всех значений переменной, а символ  $\forall$  означает, что указанные свойства имеют место для всех значений переменной. Положительные постоянные, не зависящие от оцениваемых величин, обозначаются через  $c_j, j = 0, 1, 2, \dots$ , ниже всюду  $a_i, b_i, s_i, i = 0, 1, 2, \dots$  – заданные положительные числа.

Рассмотрим управляемый процесс, описываемый следующей начально-краевой задачей для уравнения типа Шредингера:

$$(1) \quad i\rho^2 \frac{\partial \psi}{\partial t} + \frac{\partial}{\partial x} \left( v_0(x) \frac{\partial \psi}{\partial x} \right) - v_1(x)\psi + a_1 |\psi|^2 \psi = f(x, t), \quad (x, t) \in \Omega,$$

$$(2) \quad \psi(x, 0) = \varphi(x), \quad x \in (0, l),$$

$$(3) \quad \frac{\partial \psi(0, t)}{\partial x} = \frac{\partial \psi(l, t)}{\partial x} = 0, \quad t \in (0, T),$$

где  $\rho > 0$ ,  $a_1$  – заданные действительные числа,  $\varphi(x)$ ,  $f(x, t)$  заданные комплексные измеримые функции, удовлетворяющие условиям:  $\varphi \in W_2^2(0, l)$ ,  $\frac{d\varphi(0)}{dx} = \frac{d\varphi(l)}{dx} = 0$ ,  $f \in W_2^{0,1}(\Omega)$ , коэффициенты уравнения  $v_0(x), v_1(x)$  вещественные функции управления.

Предположим, что дополнительно задано граничное наблюдение следующего вида:

$$(4) \quad \psi(0, t) = y_0(t), \quad \psi(l, t) = y_1(t), \quad 0 \leq t \leq T,$$

где  $y_0 = y_0(t)$ ,  $y_1 = y_1(t)$  – заданные комплексные функции из пространства  $L_2(0, T)$ . Пусть вектор-функция  $v = v(x) = (v_0(x), v_1(x))$  является элементом следующего множества допустимых управлений:  $V \equiv \{v = (v_0, v_1) : v_0 \in W_2^1(0, l), v_1 \in L_2(0, l), 0 < b_0 \leq v_0(x) \leq b_1, \left| \frac{dv_0(x)}{dx} \right| \leq b_2, 0 < b_3 \leq v_1(x) \leq b_4, \forall^o x \in (0, l)\}$ , где  $b_j > 0, j = 0, 1, \dots, 4$  – заданные числа. При каждом элементе  $v \in V$  волновую функцию  $\psi = \psi(x, t) \equiv \psi(x, t; v)$ , принадлежащую пространству  $B \equiv C^0([0, T]; W_2^1(0, l)) \cap C^1([0, T]; L_2(0, l))$ , назовем решением прямой задачи, когда она удовлетворяет уравнению (1) для любого  $t \in [0, T]$  и почти всех  $x \in (0, l)$ , условиям (2), (3) для почти всех  $x \in (0, l)$  и  $t \in (0, T)$ . Прямая задача является начально-краевой задачей для уравнения типа Шредингера (1). Эта задача изучалась в [1–11] и др., где установлено, что при принятых выше условиях при каждом  $v \in V$  она имеет единственное решение в пространстве  $B$  и для него верна следующая априорная оценка:

$$(5) \quad \|\psi(\cdot, t)\|_{W_2^1(0, l)} + \left\| \frac{\partial \psi(\cdot, t)}{\partial t} \right\|_{L_2(0, l)} \leq c_0 \left( \|\varphi\|_{W_2^1(0, l)} + \|f\|_{L_2(0, T; W_2^1(0, l))} \right).$$

Теперь рассмотрим задачу оптимального управления функцией  $v = v(x) = (v_0(x), v_1(x))$  в уравнении (1) с начально-краевыми условиями (2), (3), где требуется минимизировать функционал качества

$$(6) \quad J_0(v) = \beta_0 \|\psi(0, \cdot) - y_0\|_{L_2(0, T)}^2 + \beta_1 \|\psi(l, \cdot) - y_1\|_{L_2(0, T)}^2$$

на множестве  $V$ , здесь числа  $\beta_0 > 0$ ,  $\beta_1 > 0$  заданы и  $\beta_0 + \beta_1 = 1$ . Заметим, что решения задач с управлениями в коэффициентах уравнения состояния часто являются неустойчивыми (см. [9–12] и др.) и для них целесообразно



изучить также задачу с «возмущенным» критерием качества. Ниже изучается задача минимизации функционала

$$(7) \quad J_\alpha(v) = \beta_0 \|\psi(0, \cdot) - y_0\|_{L^2(0,T)}^2 + \beta_1 \|\psi(l, \cdot) - y_1\|_{L^2(0,T)}^2 + \alpha \|v - \omega\|_H^2$$

на множестве  $V$  при условиях (1)–(3), где  $\alpha \geq 0$  – заданное число,  $H \equiv W_2^1(0, l) \times L_2(0, l)$  – пространство управлений, элемент  $\omega \in H$  – заданная вектор-функция. Ради краткости изложения ниже задачи (1)–(3), (6) и (1)–(3), (7) назовем задачами (6) и (7) соответственно.

### 3. Разрешимость задачи оптимального управления

*Теорема 1. Задача (7) при любом  $\alpha \geq 0$  имеет хотя бы одно решение.*

Доказательство дано в Приложении.

Пусть  $H_\infty \equiv W_\infty^2(0, l) \times W_\infty^1(0, l)$ ,  $H_1 \equiv W_2^4(0, l) \times W_2^1(0, l)$ , и задачу оптимального управления изучим на следующем множестве:  $V_1 \equiv \left\{ v = (v_0, v_1); v_0 \in W_2^2(0, l), v_1 \in W_2^1(0, l), 0 < s_0 \leq v_0(x) \leq s_1, \left| \frac{dv_0(x)}{dx} \right| \leq s_2, \left| \frac{d^2 v_0(x)}{dx^2} \right| \leq s_3, 0 < s_4 \leq v_1(x) \leq s_5, \left| \frac{dv_1(x)}{dx} \right| \leq s_6, \forall x \in (0, l) \right\}$ , где  $s_i > 0$ ,  $i = 0, 1, 2, 3, 4, 5, 6$  – заданные числа. Предположим, что функции  $\varphi(x)$ ,  $f(x, t)$  удовлетворяют условиям:  $\frac{d^3 \varphi}{dx^3} \in L_2(0, l)$ ,  $\frac{\partial f}{\partial x} \in W_2^{0,1}(\Omega)$ . При этих условиях в [9, 10] доказано, что для каждого элемента  $v = v(x)$  из множества  $V_1$  решение прямой задачи (1)–(3) существует, единственно при каждом  $t \in [0, T]$ , принадлежит пространству  $B_1 = C^0([0, T]; W_2^4(0, l)) \cap C^1([0, T]; L_2(0, l))$  и верна априорная оценка:

$$(8) \quad \begin{aligned} & \|\psi(\cdot, t)\|_{W_2^1(0,l)} + \left\| \frac{\partial \psi(\cdot, t)}{\partial t} \right\|_{L_2(0,l)} + \left\| \frac{\partial^2 \psi(\cdot, t)}{\partial t^2} \right\|_{L_2(0,l)} \leq \\ & \leq c_1 \left( \|\varphi\|_{W_2^1(0,l)} + \|f\|_{W_2^{1,1}(\Omega)} + \left\| \frac{\partial f}{\partial x} \right\|_{W_2^{1,1}(\Omega)} \right), \quad \forall t \in [0, T]. \end{aligned}$$

Теперь на множестве  $V_1 \subseteq H_1$  рассмотрим аналог задачи (7) о минимизации функционала

$$(9) \quad J_\alpha(v) = \beta_0 \|\psi(0, \cdot) - y_0\|_{L_2(0,T)}^2 + \beta_1 \|\psi(l, \cdot) - y_1\|_{L_2(0,T)}^2 + \alpha \|v - \omega\|_{H_1}^2$$

при условиях (1)–(3), где  $\alpha \geq 0$  – заданное число,  $H_1 = W_2^2(0, l) \times W_2^1(0, l)$  – пространство управлений,  $\omega \in H_1$  – заданная вектор-функция. Легко проверить, что  $V_1 \subseteq V$ . Если же  $\alpha = 0$  и  $V = V_1$ ,  $H = H_1$ , то задачи (7) и (9) совпадают с задачей (6). При  $V = V_1$  и  $H = H_1$  задачи (7) и (9) совпадают. Примеры, аналогичные приведенным в [9, 10], показывают, что при  $\alpha = 0$  решение задачи (7) или (9) неустойчиво и неединственное. Однако при  $\alpha > 0$  верна

*Теорема 2. В пространстве  $H_1$  существует всюду плотное подмножество  $K$ , что при всех  $\alpha > 0$  и при  $\omega \in K$  задача (9) имеет единственное решение.*

Доказательство дано в Приложении.

#### 4. Заключение

Указанные теоремы дают основания для решения рассмотренных выше задач. Опыт решения задач оптимального управления показывает, что для этой цели наиболее эффективными методами являются итеративные методы численного решения (см. [2, 9, 10, 13, 16, 17]). Пакеты программ, такие как Maple, Matlab, ANSYS и др. с визуализацией результатов, являются более предпочтительными и дают более удобный аппарат для численного решения. Рассмотрим итерационный процесс решения задачи (7) или (9), основанный на следующей схеме метода условного градиента численного решения:

$$v^{(k+1)}(x) = v^{(k)}(x) + \lambda_k \left( w^{(k)}(x) - v^{(k)}(x) \right), \quad k = 0, 1, 2, \dots,$$

где  $v^{(0)}(x)$  – начальный шаг итераций, который может быть любым элементом множества  $V$  (или  $V_1$ ), параметр алгоритма  $\lambda_k \in (0, 1)$  выбирается из условия  $J_0(v^{(k+1)}) < J_0(v^{(k)})$ , элемент  $w^{(k)}(x)$  определяется из минимума линейного функционала:

$$\delta J_0(v^{(k)}, w - v^{(k)}) + 2a\langle v^{(k)} + v^{(k-1)}, w - v^{(k)} \rangle \rightarrow \infty,$$

где  $k = 1, 2, \dots$  на множестве  $V$  (или  $V_1$ ),  $\delta J_0(v)$  – первая вариация функционала  $J_0(v)$ .

Пусть  $h \in H$  – приращение управления  $v \in V$  такое, что  $v + h \in V$ . Первая вариация функционала  $J_0(v)$  имеет вид

$$\delta J_0(v, h) = \int_{\Omega} \left[ \operatorname{Re} \left( \frac{\partial \psi(x, t)}{\partial x} \frac{\partial \varphi(x, t)}{\partial x} \right) h_0(x, t) + \right. \\ \left. + \operatorname{Re} (\psi(x, t) \varphi(x, t)) h_1(x, t) \right] dx dt,$$

где  $\psi(x, t)$  и  $\varphi(x, t)$  являются решениями прямой и сопряженной задач соответственно. Действуя по указанной выше схеме, найдем последовательность управлений  $(v^{(k)})$ . В [9, 10, 13, 14] и др. дан ряд достаточных условий для данных выше изучаемых задач, которые обеспечивают сходимость вышеприведенной последовательности управлений. Там же изучаются широкие классы итерационных процессов, основанных на вариантах методов градиента, Ньютона и др., для решения экстремальных задач, проведены численный анализ и регуляризация решения в случае неустойчивости задачи.

*Доказательство теоремы 1.* Из ограниченности снизу функционала  $J_\alpha(v)$  следует существование последовательности  $(v^k(x)) \in V$  минимизации решения задачи (7). Обозначим  $\psi_k = \psi_k(x, t) \equiv \psi(x, t; v^k)$ ,  $k = 1, 2, \dots$ ,  $H_\infty \equiv W_\infty^1(0, l) \times L_\infty(0, l)$ . Поскольку множество  $V$  является замкнутым, ограниченным и выпуклым подмножеством в пространстве  $H$ , из последовательности  $(v^k)$  можно выделить подпоследовательность, которую для простоты изложения снова обозначим через  $(v^k)$  так, что  $v_0^k \rightarrow v_0$  (\*) слабо в  $L_\infty(0, l)$ ,  $\frac{dv_0^k}{dx} \rightarrow \frac{dv_0}{dx}$  (\*) слабо в  $L_\infty(0, l)$  и  $v_1^k \rightarrow v_1$  (\*) слабо в  $L_\infty(0, l)$  при  $k \rightarrow \infty$ . Кроме того, из определения множества  $V$  следует, что оно является (\*) слабокompактным в пространстве  $H$ . Из указанных только что предельных соотношений и вложения пространства  $W_\infty^k(0, l)$  в  $L_\infty(0, l)$  получаем, что  $v_0^k \rightarrow v_0$  сильно в  $L_\infty(0, l)$  при  $k \rightarrow \infty$ .

Из априорной оценки (5) следует, что последовательность  $\{\psi_k(x, t)\}$  равномерно ограничена по норме  $B$ . Тогда из последовательности  $\{\psi_k(x, t)\}$  можно выделить подпоследовательность, которую для простоты изложения снова обозначим через  $\{\psi_k(x, t)\}$ , что  $\psi_k(\cdot, t) \rightarrow \psi(\cdot, t)$  либо в  $W_2^2(0, l)$  и  $\frac{\partial \psi_k(\cdot, t)}{\partial t} \rightarrow \frac{\partial \psi(\cdot, t)}{\partial t}$  либо в  $L_2(0, l)$  для  $t \in [0, T]$  при  $k \rightarrow \infty$ . Из компактности вложения пространства  $B$  в  $C^0([0, T]; W_0^2(0, l))$  (см.: [15, с. 214]) следует, что  $\|\psi_k(\cdot, t) - \pi(\cdot, t)\|_{W_0^2(0, l)} \rightarrow 0$  при  $k \rightarrow \infty$  равномерно по  $t \in [0, T]$ . Нетрудно проверить, что каждый элемент последовательности удовлетворяет тождеству

$$(П.1) \quad \int_0^l \left\{ \frac{i\rho^2 \partial \psi_k(x, t)}{\partial t} + \frac{\partial}{\partial x} \left( v_0^k(x) \frac{\partial \psi_k(x, t)}{\partial x} \right) - v_1^k(x) \psi_k(x, t) \right\} + \\ + a_1 \|\psi_k(x, t)\|^2 \|\psi(x, t) - \overline{f(x, t)}\| \eta(x) dx = 0, \quad k = 1, 2, \dots$$

для всех  $t \in [0, T]$  и произвольной функции  $\eta \in L_2(0, l)$ . Кроме того,  $\psi_k(x, t) = 1, 2, \dots$ , удовлетворяют начальное и краевые условия (2) и (3). Если в тождестве (П.1) перейти к пределу по  $k = 1, 2, \dots$  для каждой  $t \in [0, T]$  и любой функции  $\eta \in L_2(0, l)$  получим, что предельная функция  $\psi_k(x, t)$  будет удовлетворять этому тождеству. Таким же образом проверяется, что предельная функция  $\psi_k(x, t)$  удовлетворяет уравнению (1) для каждого  $t \in [0, T]$  и почти для всех  $x \in (0, l)$ . При  $t = 0$  получим, что  $\|\psi_k(\cdot, 0) - \psi(\cdot, 0)\|_{L_2(0, l)} \rightarrow 0$ , когда  $k \rightarrow \infty$ . Отсюда и из предельного перехода при  $k \rightarrow \infty$  в

$$\|\psi_k(\cdot, 0) - \varphi\|_{L_2(0, l)} \leq \|\psi_k(\cdot, 0) - \psi(\cdot, 0)\|_{L_2(0, l)} + \|\psi(\cdot, 0) - \varphi\|_{L_2(0, l)},$$

получим, что функция  $\psi(x, t)$  удовлетворяет начальному условию (2) почти для всех  $x \in (0, l)$ . Наконец, докажем, что предельная функция  $\psi_k(x, t)$  удовлетворяет граничным условиям (3). Действительно, для функций из пространства  $B$  выполняются предельные соотношения  $\frac{\partial \psi_k(s, \cdot)}{\partial x} \rightarrow \frac{\partial \psi(s, \cdot)}{\partial x}$  слабо

в  $L_2(0, T)$  при  $k \rightarrow \infty$ , где  $s = 0$  и  $s = l$ . Тогда используя это соотношение, после предельного перехода  $k \rightarrow \infty$  в равенстве

$$\int_0^T \frac{\partial \psi(s, t)}{\partial x} \eta(t) dt = \int_0^T \left( \frac{\partial \psi(s, t)}{\partial x} - \frac{\partial \psi_k(s, t)}{\partial x} \right) \eta(t) dt + \\ + \int_0^T \frac{\partial \psi_k(s, t)}{\partial x} \eta(t) dt, \quad s = 0, l$$

получим выражение

$$\int_0^T \frac{\partial \psi(s, t)}{\partial x} \eta(t) dt = 0, \quad s = 0, l$$

для любой  $\eta(t)$  функции из  $L_2(0, T)$ . Отсюда следует, что почти при всех  $t \in (0, T)$  функция  $\psi(x, t)$  удовлетворяет граничным условиям (3). Таким образом, предел  $\psi(x, t)$  последовательности  $(\psi_k(x, t))$  является решением прямой задачи (1)–(3) при каждом управлении  $v$  из  $V$ . Кроме того, предельная функция является слабым пределом слабосходящейся последовательности  $\{\psi_k(x, t)\}$  для каждого  $t \in [0, T]$ . Из компактности вложения пространства  $B$  в  $C([0, l]; L_2(0, T))$  имеем, что  $\|\psi_k(x, t) - \psi(x, t)\|_{L_2(0, T)} \rightarrow 0$  при  $k \rightarrow \infty$  для каждого  $x \in [0, l]$ . Отсюда при  $x = 0, x = l$  из слабой полунепрерывности снизу нормы в  $L_2(0, T)$  и не отрицательности чисел  $\beta_0, \beta_l, \alpha$  получим справедливость неравенства:  $J_{\alpha*} \leq J_{\alpha}(v) \leq J_{\alpha*}$ . Это означает, что элемент  $v = v(x) \in V$  является решением задачи (7) при любом  $\alpha \geq 0$ . Теорема 1 доказана.

*Доказательство теоремы 2.* Проверим непрерывность функционала  $J_0(v)$  на множестве  $V_1$ . Пусть  $\delta\psi = \delta\psi(x, t) = \psi(x, t; v + \delta v) - \psi(x, t; v)$ , где  $\delta v \in H_1$  приращение элемента  $v \in V_1$ , что  $v + \delta v \in V_1$ ,  $\psi(x, t; v)$  решение прямой задачи (1)–(3) для элемента  $v \in V_1$ . Из (1)–(3) следует, что функция  $\delta\psi = \delta\psi(x, t)$  удовлетворяет уравнению

$$(P.2) \quad i\rho^2 \frac{\partial \delta\psi}{\partial t} + \frac{\partial}{\partial x} \left( (v_0(x) + \delta v_0(x)) \frac{\partial \delta\psi}{\partial x} \right) - (v_1(x) + \delta v_1(x)) \delta\psi + \\ + a_1 \left( |\psi_{\delta}|^2 + |\psi|^2 \right) \delta\psi + a_1 \psi_{\delta} \psi \delta \bar{\psi} = \\ = - \frac{\partial}{\partial x} \left( \delta v_0(x) \frac{\partial \psi}{\partial x} \right) + \delta v_1(x) \psi, \quad (x, t) \in \Omega.$$

с однородными начально-краевыми условиями. Для оценки решения этой начально-краевой задачи обе части уравнения (P.2) умножим на функцию  $\delta\psi(x, t)$  и результат проинтегрируем по области  $\Omega_t$ . После вычета из полу-

ченного равенства его комплексного сопряжения с учетом однородности начального условия имеем:

$$\begin{aligned} & \|\rho\delta\psi(\cdot, t)\|_{L_2(0,l)}^2 = 2a_1 \int \operatorname{Im}(\psi\psi_\delta, (\delta\bar{\psi})^2) dx d\tau + \\ & + 2 \int \operatorname{Im} \left[ \left( -\frac{\partial}{\partial x} \left( \delta v_0(x) \frac{\partial \psi}{\partial x} \right) + \delta v_1(x)\psi \right) \delta\bar{\psi} \right] dx dt, \quad \forall t \in [0, T]. \end{aligned}$$

Применяя к этому тождеству неравенство Коши–Шварца, получим:

$$\begin{aligned} & \|\rho\delta\psi(\cdot, t)\|_{L_2(0,l)}^2 \leq 2|a_1| \int_{\Omega_t} |\psi_\delta| |\psi| \delta\psi^2 dx d\tau + \\ (П.3) \quad & + \int_{\Omega_t} \left[ -\frac{\partial}{\partial x} \left( \delta v_0(x) \frac{\partial \psi}{\partial x} \right) + \delta v_1(x)\psi \right]^2 dx d\tau + \\ & + \iint_{\Omega_t} |\delta\psi|^2 dx d\tau, \quad \forall t \in [0, T]. \end{aligned}$$

Известно, что если функция  $\varphi(\cdot, t) \in W_2^1(0, l)$  не равна нулю на концах отрезка  $[0, l]$ , тогда для нее выполняется следующее неравенство (см. [15, с. 80, замечание 2.1]):

$$\|\varphi(\cdot, t)\|_{L_2(0,l)} \leq c \left\| \frac{\partial \varphi(\cdot, t)}{\partial x} \right\|_{L_2(0,l)}^{\frac{1}{2}} \|\varphi(\cdot, t)\|_{L_2(0,l)}^2 + d \|\varphi\|_{L_2(0,l)},$$

где  $c > 0$ ,  $d > 0$  – некоторые постоянные. Из этого неравенства и априорной оценки (8) решения прямой задачи (1)–(3) при каждом  $V \in V_1$  следует, что

$$(П.4) \quad \|\psi(\cdot, t)\|_{L_2(0,l)}^2 \leq c_4; \quad \|\psi^2(\cdot, t)\|_{L_2(0,l)}^2 \leq c_4.$$

С учетом этих неравенств в уравнении (П.2), имеем:

$$\begin{aligned} & \|\delta\psi(\cdot, t)\|_{L_2(0,l)}^2 \leq (2|a_1|c_4^2 + 1) \int_0^t \|\delta\psi(\cdot, \tau)\|_{L_2(0,l)}^2 d\tau + \\ & + \int_{\Omega} \left| -\frac{\partial}{\partial x} \left( \delta v_0(x) \frac{\partial \psi}{\partial x} \right) + \delta v_1(x)\psi \right|^2 dx d\tau, \quad \forall t \in [0, T]. \end{aligned}$$

Оценивая левую часть этого неравенства, имеем:

$$\int_{\Omega_t} \left| -\frac{\partial}{\partial x} \left( \delta v_0(x) \frac{\partial \psi}{\partial x} \right) + \delta v_1(x)\psi \right|^2 dx d\tau \leq c_5 \int_0^t \|\psi(\cdot, \tau)\|_{W_2^2(0,l)}^2 d\tau \|\delta v\|_{B^1}^2.$$

Учитывая это неравенство в уравнении (П.2), получим:

$$\|\delta\psi(\cdot, t)\|_{L_2(0,l)}^2 \leq c_6 \int_0^t \|\delta\psi(\cdot, \tau)\|_{L_2(0,l)}^2 d\tau + c_7 \|\delta v\|_H^2, \quad \forall t \in [0, T].$$

Применяя к этому неравенству лемму Гроноуллы, имеем:

$$(П.5) \quad \|\delta\psi(\cdot, t)\|_{L_2(0,l)}^2 \leq c_8 \|\delta v\|_{B_1}^2, \quad \forall t \in [0, T].$$

Теперь оценим функцию  $\frac{\partial \delta\psi}{\partial x}(x, t)$ . Для этой цели обе части уравнения (П.2) умножим на  $L(\delta\bar{\psi}) = -\frac{\partial}{\partial x} \left( (v_0(x) + \delta v_0(x)) \frac{\partial \delta\bar{\psi}}{\partial x} \right)$  и интегрируем полученное выражение по области  $\Omega_t$ :

$$\begin{aligned} & \int_{\Omega_t} \left( i\rho^2 \frac{\partial \delta\psi}{\partial t} L(\delta\psi) - |L(\delta\psi)|^2 - (v_0(x) + \delta v_0(x)) \delta\psi L(\delta\psi) + \right. \\ & \quad \left. + a_1 (|\psi_\delta|^2 + |\psi|^2) \delta\psi L(\delta\bar{\psi}) + a_1 \psi_\delta \psi \delta\bar{\psi} L(\delta\bar{\psi}) \right) dx d\tau = \\ & = \int_{\Omega} \left( \frac{\partial}{\partial x} \left( \delta v_0(x) \frac{\partial \psi}{\partial x} \right) - \delta v_1(x) \psi \right) L(\delta\bar{\psi}) dx d\tau, \quad \forall t \in [0, T]. \end{aligned}$$

Интегрируем по частям относительно  $x$  в обеих частях этого уравнения и из полученного равенства вычтем его комплексное сопряженное, тогда получим:

$$\begin{aligned} & \int_{\partial_t} \rho^2 \frac{\partial}{\partial t} (v_o(x) + \delta v_o(x)) \left| \frac{\partial \delta\psi}{\partial x} \right|^2 dx d\tau = \\ & = 2 \int_{\Omega_t} \left[ \operatorname{Im} (v_o(x) + \delta v_o(x)) \left( \frac{dv_1(x)}{dx} + \frac{d\delta v_1(x)}{dx} \right) \delta\psi \frac{\partial \delta\bar{\psi}}{\partial x} \right] dx d\tau - \\ & - 2a_1 \int_{\Omega_t} \left[ \operatorname{Im} (v_o(x) + \delta v_o(x)) \frac{\partial}{\partial x} (|\psi_\delta|^2 + |\psi|^2) \delta\psi \frac{\partial \delta\bar{\psi}}{\partial x} \right] dx d\tau - \\ & - 2a_1 \int_{\Omega_t} \left[ \operatorname{Im} (v_o(x) + \delta v_o(x)) \frac{\partial}{\partial x} (\psi_\delta \psi) \delta\bar{\psi} \frac{\partial \delta\bar{\psi}}{\partial x} \right] dx d\tau - \\ & - 2a_1 \int_{\Omega_t} \left[ \operatorname{Im} (v_o(x) + \delta v_o(x)) \psi_\delta \psi \left( \frac{\partial \delta\bar{\psi}}{\partial x} \right)^2 \right] dx d\tau - \\ & - 2 \int_{\Omega_t} \left[ \operatorname{Im} (v_o(x) + \delta v_o(x)) \frac{\partial}{\partial x} \left[ \frac{\partial}{\partial x} \left( \delta v_o(x) \frac{\partial \psi}{\partial x} \right) - \delta v_1(x) \psi \right] \frac{\partial \delta\bar{\psi}}{\partial x} \right] dx d\tau, \quad \forall t \in [0, T]. \end{aligned}$$

Если здесь учтем, что  $v + \delta v \in V_1$ ,  $\delta\psi(x, 0) = 0$ ,  $x \in (0, l)$ , тогда с учетом неравенства Коши–Шварца и оценок (П.4), имеем

$$\begin{aligned}
 s_0 \left\| \frac{\partial \delta\psi(\cdot, t)}{\partial x} \right\|_{L_2(0, l)}^2 &\leq c_9 \int_0^t \left\| \frac{\partial \delta\psi(\cdot, \tau)}{\partial x} \right\|_{L_2(0, l)}^2 d\tau + c_{10} \int_{\Omega_t} |\delta\psi|^2 dx d\tau + \\
 &+ 4s_1 \int_{\Omega_t} |\delta v_0(x)|^2 \left| \frac{\partial^3 \psi}{\partial x^3} \right|^2 dx d\tau + 8s_1 \int_{\Omega_t} \left| \frac{d\delta v_0(x)}{dx} \right|^2 \left| \frac{\partial^2 \psi}{\partial x^2} \right|^2 dx d\tau + \\
 \text{(П.6)} \quad &+ 4s_1 \int_{\Omega_t} \left| \frac{d\delta v_1(x)}{dx} \right|^2 |\psi|^2 dx d\tau + \\
 &+ 4s_1 \int_{\Omega_t} \left( \left| \frac{d^2 \delta v_0(x)}{dx^2} \right| + |\delta v_1(x)| \right)^2 \left| \frac{\partial \psi}{\partial x} \right|^2 dx d\tau + \\
 &+ c_8 |a_1| (2 + s_1) \int_{\Omega_t} \left( \left| \frac{\partial \psi_\delta}{\partial x} \right| + \left| \frac{\partial \psi}{\partial x} \right| \right)^2 |\delta\psi|^2 dx d\tau, \quad \forall t \in [0, T],
 \end{aligned}$$

где  $c_9 = s_1(|a_1|c_4 + |a_1|c_4^2 + 1 + s_5)$ ,  $c_{10} = s(|a_1|c_4^2 + s_5)$ ,  $\forall t \in [0, T]$ . Теперь оценим последнее слагаемое в правой части этого неравенства. Известно, что для произвольной функции  $\varphi(\cdot, t) \in W_\lambda(0, l)$ ,  $\forall t \in [0, T]$  верно следующее неравенство [15, с. 79–80]:

$$\text{(П.7)} \quad \left\| \frac{\partial \varphi(\cdot, t)}{\partial x} \right\|_{L_\infty(0, l)} \leq c_0 \left\| \frac{\partial \varphi(\cdot, t)}{\partial x} \right\|_{L_2(0, l)} \|\varphi(\cdot, t)\|_{L_2(0, l)}, \quad \forall t \in [0, T],$$

где  $c_0 > 0$  – некоторая постоянная. Если в этом неравенстве вместо функции  $\varphi(x, t)$  возьмем  $\frac{\partial \psi_\delta(x, t)}{\partial x}$  и  $\frac{\partial \psi(x, t)}{\partial x}$  воспользуемся (П.5), (П.6), тогда для этих функций имеем:

$$\text{(П.8)} \quad \left. \frac{\partial \psi_\delta(\cdot, t)}{\partial x} \right|_{L_\infty(0, l)} \leq c_{11}, \quad \left. \frac{\partial \psi(x, t)}{\partial x} \right|_{L_\infty(0, l)} \leq c_{11}.$$

Используя эти оценки в (7) и учитывая неравенство (П.5), получим:

$$\left\| \frac{\partial \delta\psi(\cdot, t)}{\partial x} \right\|_{L_2(0, l)}^2 \leq c_{12} \int_0^t \left\| \frac{\partial \psi(\cdot, \tau)}{\partial x} \right\|_{L_2(0, l)}^2 d\tau + c_{13} \|\delta v\|_{B_1}^2.$$

Применяя сюда известную лемму Гронуолла, получим оценку

$$\left\| \frac{\partial \psi(x, t)}{\partial x} \right\|_{L_2(0, l)}^2 \leq c_{15} \|\delta v\|_{B_1}^2, \quad \forall t \in [0, T].$$

Объединяя эту оценку с оценкой (П.5), имеем:

$$(П.9) \quad \left\| \delta\psi(\cdot, t) \right\|_{W_2^1(0, l)}^2 \leq c_{15} \|\delta v\|_{B_1}^2, \quad \forall t \in [0, T].$$

Преобразуем приращение функционала  $J_0(v)$  на произвольном элементе  $v \in V_1$ :

$$(П.10) \quad \begin{aligned} \Delta J_0(v) &= J_0(v + \delta v) - J_0(v) = \\ &= 2\beta_0 \int_0^T \mathcal{R}\epsilon[(\psi(0, t) - y_0(t)) \delta\psi(0, t)] dt + \\ &+ 2\beta_1 \int_0^t \mathcal{R}\epsilon[(\psi(l, t) - y_1(t)) \delta\psi(l, t)] dt + \\ &+ \beta_0 \left\| \delta\psi(x, t) \right\|_{L_2(0, T)}^2 + \beta_1 \left\| \delta\psi(x, t) \right\|_{L_2(0, T)}^2. \end{aligned}$$

Из теоремы вложения для следа функций из пространства  $W_2^{1,0}(\Omega)$ , имеем:

$$\begin{aligned} \|\psi(0, \cdot)\|_{L_2(0, T)}^2 + \|\psi(l, \cdot)\|_{L_2(0, T)}^2 &\leq c_{16} \|\psi\|_{W_2^{1,0}(\Omega)}^2, \\ \|\delta\psi(0, \cdot)\|_{L_2(0, T)}^2 + \|\delta\psi(l, \cdot)\|_{L_2(0, T)}^2 &\leq c_{16} \|\delta\psi\|_{W_2^{1,0}(\Omega)}^2. \end{aligned}$$

Из этих неравенств следует, что:

$$(П.11) \quad \|\psi(0, \cdot)\|_{L_2(0, T)}^2 + \|\psi(l, \cdot)\|_{L_2(0, T)}^2 \leq c_{17},$$

$$(П.12) \quad \|\delta\psi(0, \cdot)\|_{L_2(0, T)}^2 + \|\delta\psi(l, \cdot)\|_{L_2(0, T)}^2 \leq c_{17}.$$

Применяя к выражению (П.10) неравенство Коши–Шварца, используя (П.11), (П.12), а также учитывая условие  $y_0, y_1 \in L_2(0, l)$ , получим оценку:

$$|\Delta J_0(v)| \leq c_{19} (\|\delta v\|_{B^1} + \|\delta v\|_{B_1}^2), \quad v \in V_1.$$

Это неравенство показывает непрерывность функционала  $J_0(v)$  на множестве  $V_1$ . Функционал  $J_0(v)$  положителен:  $J_0(v) \geq 0$  для  $v \in V_1$ . Кроме того, множество  $V_1$  является замкнутым, ограниченным и выпуклым множеством в гильбертовом пространстве  $H_1$ . Гильбертовы пространства являются равномерно выпуклыми. Если функционал  $I_0(v)$  полунепрерывен снизу и ограничен снизу на замкнутом ограниченном множестве  $U \subset X$  равномерно выпуклого банахова пространства  $X$ , тогда, как доказано в [16], существует такое всюду плотное подмножество  $K \subset X$ , что для любых  $w \in K$  и при любом  $\alpha > 0$  функционал  $I_0(v) + \alpha \|v - w\|_X^2$  достигает минимального значения на единственном элементе множества  $K \cap U$ . Для задачи (9) все условия и утверждения этой теоремы удовлетворяются. Поэтому утверждение теоремы верно. Теорема 2 доказана.



## СПИСОК ЛИТЕРАТУРЫ

1. *Воронцов И.М., Шмальсдугеи В.И.* Основы адаптивной оптики. М.: Наука, 1985.
2. «Interatomic Potentials for Atomistic Simulations» // Materials Research Society Bulletin. 1996. V. 21. No. 2. P. 3–97.
3. *Балакин А.А., Балакина М.А., Пермитин Г.В., Смирнов А.И.* Скалярное уравнение для волновых пучков в магнитной плазме // Физика плазмы. 2007. Т. 33. № 4. С. 334–345.
4. *Бутковский А.Г., Самойленко Ю.И.* Управление квантово-механическими процессами. М.: Наука, 1994.
5. *Саябаев В.Ж.* Задача Коши для линейного дифференциального уравнения с вырождением и усреднение аппроксимирующих ее регуляризацией // Современная математика. Фундаментальные направления. Т. 43. 2012. С. 3–172.
6. *Кротов В.Ф., Булатов А.В., Батрушка О.В.* Оптимизация линейных систем с управляемыми коэффициентами // АиТ. 2011. № 6. С. 64–78.
7. *Baudoin L., Kavian O., Fuel J.-P.* Regularity for Schrodinger equation with singular potentials and application to bilinear optimal control // J. Differ. Equat. 2005. V. 21. No. 6. P. 188–222.
8. *Искендеров А.Д., Ягубов Г.Я.* Оптимальное управление нелинейными квантово-механическими системами // АиТ. 1989. № 12. С. 27–38.
9. *Искендеров А.Д., Ягубов Г.Я., Мусаева М.А.* Идентификация квантовых потенциалов. Баку: Чашмоглу, 2012.
10. *Мусаева М.А.* Вариационные методы определения квантовых потенциалов. Баку: Елм-Техсил, 2018.
11. *Мусаева М.А.* Вариационный метод определения комплексных коэффициентов нелинейного и нестационарного уравнения типа Шредингера // Журн. выч. мат. и мат. физики. 2020. Т. 60. № 11. С. 1985–1997.
12. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. М.: Наука, 1986.
13. *Тихонов А.Н., Леонов В.С., Ягола А.Г.* Нелинейные некорректные задачи. М.: КУРС, 2017.
14. *Васильев Ф.П.* Методы оптимизации. М.: МЦНМО, 2011.
15. *Ладженская О.А., Уральцева Н.Н.* Линейные и квазилинейные уравнения эллиптического типа. М.: Наука, 1964.
16. *Goebel M.* On existence of optimal control // Math. Nachr. 1978. V. 93. No. 1. P. 67–73.
17. *Лапин А.В.* Итерационные методы решения сеточных вариационных неравенств. Казань: КГУ, 2008.

*Статья представлена к публикации членом редколлегии А.В. Арутюновым.*

Поступила в редакцию 31.05.2024

После доработки 30.11.2024

Принята к публикации 09.12.2024

# Стохастические системы

© 2025 г. Ю.Г. БУЛЫЧЕВ, д-р техн. наук (profbulychhev@yandex.ru)  
(АО «Концерн Радиоэлектронные технологии», Москва)

## ДЕКОМПОЗИЦИОННО-АВТОКОМПЕНСАЦИОННОЕ РАСПОЗНАВАНИЕ СИГНАЛОВ НА БАЗЕ ПРИНЦИПОВ НЕПРЕРЫВНОСТИ, ИНВАРИАНТНОСТИ, РАЗМНОЖЕНИЯ И РАНЖИРОВАНИЯ С УЧЕТОМ РЕГУЛЯРНЫХ И НЕРЕГУЛЯРНЫХ ПОМЕХ

С учетом принципов непрерывности, инвариантности, размножения и ранжирования развивается новый метод оптимального распознавания сигналов в условиях существенной априорной неопределенности применительно к информационно-измерительным системам реального времени. Предполагается, что в уравнении наблюдения помимо случайного шума (с неизвестным законом распределения, но заданной корреляционной матрицей) могут присутствовать регулярная помеха (допускает аналитическое конечно-спектральное представление) и нерегулярная помеха (для ее описания не удастся использовать какой-либо эффективный вариант вероятностного описания). Относительно последней возможно лишь введение некоторых числовых характеристик и ограничений, подтверждающихся практикой эксплуатации конкретной системы. Метод инвариантен к указанным помехам, не требует традиционного расширения пространства состояний и обеспечивает декомпозицию вычислительной процедуры. Анализируются случайные и методические погрешности, а также достигаемый вычислительный эффект. Приводится иллюстративный пример.

*Ключевые слова:* существенная априорная неопределенность, регулярная помеха, нерегулярная помеха, коэффициент «засорения» выборки, принцип непрерывности, принцип инвариантности, принцип размножения, принцип ранжирования, спектральные коэффициенты, критерий оптимальности, декомпозиция, алгоритмы распознавания.

DOI: 10.31857/S0005231025040036, EDN: CATNGU

### 1. Введение

Для современной информационно-измерительной системы (далее системы) одной из важнейших является задача распознавания сигналов, включающая обнаружение, различение, разрешение, сглаживание, оценивание параметров и др. [1–18]. Наиболее актуальны системы реального времени применительно к условиям существенной неопределенности и жестких требований к качеству распознавания не только в среднем (по ансамблю реализаций), но и единичном случае (по одной фиксированной выборке измерений). Речь

идет о системе, для которой риски (потери) от неправильного решения в таком случае достаточно велики.

Существенная неопределенность возникает, например, когда реальные измерения подвержены воздействию не только случайного шума с неизвестным законом распределения, но также регулярных и нерегулярных помех (соответственно РП и НП). При этом сведения о шуме ограничиваются лишь корреляционной матрицей, для РП – заданным конечным набором базисных функций, а для НП – некоторыми числовыми характеристиками и ограничениями, характерными для конкретной системы. Например, в качестве одной из таких характеристик может использоваться коэффициент «засорения» выборки.

Для преодоления указанной неопределенности зачастую используется метод наименьших квадратов или одна из его модификаций применительно к простому или расширенному варианту [5, 7–9, 14–18]. Данный метод обеспечивает построение наилучших в среднем оценок с учетом известной теоремы Гаусса–Маркова (см., например, [5, с. 34]). Реже используются методы максимума апостериорной плотности вероятности, максимального правдоподобия, линейной и нелинейной фильтрации и др. [1–7, 11, 14]. Данные методы, как правило, требуют достаточно большого объема априорной статистической информации и трудно реализуемы в реальном времени (особенно при расширении пространства состояний). Кроме того, данные методы зачастую являются неустойчивыми (в вычислительном плане) и обладают плохой сходимостью при итерационных расчетах, например, при неудачном выборе начальных условий или работе с целевыми функциями «овражного» типа. Чаще всего они применяются на этапах, связанных с математическим моделированием при обосновании потенциальных возможностей системы. В указанных выше единичных случаях (когда речь идет об одной фиксированной выборке) результаты оценивания (особенно в условиях НП) могут приводить к значительным потерям.

В [19, 20] развит метод обобщенного инвариантно-несмещенного оценивания (МОИНО) любых линейных числовых характеристик полезных сигналов (например, спектральных коэффициентов, производных различного порядка, интегралов, сглаженных значений и т.д.) с учетом РП (иногда их называют сигналоподобными, систематическими, сингулярными или динамическими) без расширения пространства состояний. МОИНО обеспечивает автокомпенсацию РП, сглаживание случайного шума и построение оптимальных оценок с минимальным следом корреляционной матрицы ошибок оценивания, при этом не требуется априорная информация о законе распределения, а достаточно знать лишь корреляционную матрицу для данного шума. Метод обеспечивает максимально возможную декомпозицию вычислительной процедуры, что приводит к обращению матриц существенно меньшей размерности по сравнению с расширенным методом наименьших квадратов (РМНК).

Однако возможности МОИНО существенно ограничены, если в измерениях наряду с РП присутствует и НП (например, в виде отдельных одиночных импульсов или пачек импульсов различной формы с неизвестными параметрами). НП может возникать как на всем интервале наблюдения, так и на отдельных, заранее не известных участках данного интервала и плохо поддается формализации (например, в рамках конечно-аналитического спектрального анализа). Попытка введения дополнительных базисных функций для описания НП приводит, как правило, к резкому возрастанию размерности и существенным ошибкам. Расширение спектрального состава совокупной помехи (РП и НП) с учетом условий существенной неопределенности и реальных условий наблюдения требует обращения матриц слишком высокой размерности с плохой обусловленностью, что сильно искажает результаты обработки измерений в рамках МОИНО.

Ниже развивается новый метод распознавания сигналов, который учитывает в себе все достоинства МОИНО и не требует расширения пространства состояний, что позволяет реализовывать формируемые на его основе алгоритмы в реальном времени применительно к системе в условиях неопределенности. Метод позволяет добиться компромисса для оценивания в среднем (с учетом условий несмещенности, инвариантности к РП и минимума следа корреляционной матрицы ошибок) и оценивания в частном (по фиксированной выборке) с учетом условия минимума влияния НП на конечную результирующую оценку.

Чтобы избежать достаточно громоздких выкладок, основной материал статьи будет ориентирован на задачу оценивания линейных параметров сигналов, однако при этом последуют соответствующие обобщения и практические рекомендации для распространения полученных результатов и на другие задачи, связанные с распознаванием сигналов, в том числе и с нелинейными параметрами (по аналогии с [20]).

## 2. Математическая постановка задачи распознавания для случая существенной априорной неопределенности

На заданном временном отрезке  $[0, T]$  введем основную сетку узлов  $C_0 = \{t_n, n = \overline{1, N_0}\}$ , для которой сформируем следующее уравнение наблюдения:

$$(1) \quad \mathbf{H}_0 = \mathbf{S}_0 + \mathbf{\Theta}_0 + \mathbf{D}_0 + \mathbf{\Xi}_0,$$

где  $\mathbf{H}_0 = [h_n, n = \overline{1, N_0}]^T$ ,  $\mathbf{S}_0 = [s_n(\mathbf{A}), n = \overline{1, N_0}]^T$ ,  $\mathbf{\Theta}_0 = [\theta_n(\mathbf{B}_\theta), n = \overline{1, N_0}]^T$ ,  $\mathbf{D}_0 = [q_n d_n(\mathbf{B}_d), n = \overline{1, N_0}]^T$  и  $\mathbf{\Xi}_0 = [\xi_n, n = \overline{1, N_0}]^T$  – векторы отсчетов входного наблюдения  $h(t)$ , полезного сигнала  $s(t, \mathbf{A})$ , РП  $\theta(t, \mathbf{B}_\theta)$ , НП  $d(t, \mathbf{B}_d)$  и шума  $\xi(t)$  соответственно,  $h_n = h(t_n)$ ,  $s_n(\mathbf{A}) = s(t_n, \mathbf{A})$ ,  $\theta_n(\mathbf{B}_\theta) = \theta(t_n, \mathbf{B}_\theta)$ ,  $d_n(\mathbf{B}_d) = d(t_n, \mathbf{B}_d)$ ,  $\xi_n = \xi(t_n)$ ;  $\mathbf{A} = [a_m, m = \overline{1, M_s}]^T$ ,  $\mathbf{B}_\theta = [b_{\theta r}, r = \overline{1, M_\theta}]^T$  и  $\mathbf{B}_d = [b_{dr}, r = \overline{1, M_d}]^T$  – неизвестные векторные спектральные коэффициенты линейных разложений сигнала, РП и НП,  $q_n \in \{0, 1\}$  – индикатор нулевых

и ненулевых отсчетов НП,  $\sum_{n=1}^{N_0} q_n = M_d$ ,  $M_d$  – параметр, соответствующий реальному числу ненулевых отсчетов НП в уравнении наблюдения (1).

Обозначим через  $\{\overline{t_1}, \overline{t_2}, \dots, \overline{t_{M_d}}\}$  множество узлов, связанных с ненулевыми отсчетами НП, где  $\overline{t_m} \in \{t_1, t_2, \dots, t_{N_0}\}$ ,  $\overline{t_{m+1}} > \overline{t_m}$ . Для описания НП воспользуемся следующим подходом. Пусть  $K_{\max} \in \{0, 1, \dots\}$  – максимально возможное количество узлов, связанных с НП, т.е.  $M_d \leq K_{\max}$ . Потребуем выполнения условия

$$(2) \quad N_{\min} + K_{\max} \leq N_0,$$

где  $N_{\min}$  – минимальное число отсчетов входного наблюдения, которых достаточно для качественного решения задачи распознавания полезного сигнала при отсутствии РП и НП.

В дальнейшем будут рассматриваться различные временные сетки узлов длиной  $N$ , для которых

$$(3) \quad N_{\min} \leq N \leq N_0.$$

При таком подходе к описанию НП их ненулевые отсчеты могут принимать произвольные значения (в том числе и аномальные), быть рассредоточенными (одиночными) или сосредоточенными (пакетными). В физических каналах таким помехам могут соответствовать, например, импульсы различной формы, длительности и интенсивности. Эти помехи зачастую связаны с различными переходными процессами, коммутациями, интерференцией, наводками, естественными и искусственными помехами и т.д., они могут маскироваться сигналом и шумом. Ненулевые отсчеты НП могут располагаться на отрезке  $[0, T]$  самым произвольным образом, и для их описания не существует какой-либо универсальной и удовлетворительной модели. Единственным способом учета НП является введение некоторых количественных ограничений (типа (2), (3)), которые соответствуют практике эксплуатации конкретной системы.

Для произвольных значений  $\mathbf{A}$ ,  $\mathbf{B}_\theta$  и  $\mathbf{B}_d$  используем следующие линейные конечномерные комбинации (модели сигнала и РП, широко распространенные на практике):

$$(4) \quad s(t, \mathbf{A}) = \mathbf{A}^T \Psi(t),$$

$$(5) \quad \theta(t, \mathbf{B}_\theta) = \mathbf{B}_\theta^T \Omega_\theta(t),$$

$$(6) \quad d(t, \mathbf{B}_d) = \mathbf{B}_d^T \Omega_d(t),$$

где  $\Psi(t) = [\psi_m(t), m = \overline{1, M_s}]^T$ ,  $\Omega_\theta(t) = [\omega_{\theta r}(t), r = \overline{1, M_\theta}]^T$  и  $\Omega_d(t) = [\omega_{dr}(t), r = \overline{1, M_d}]^T$  – заданные базисные функции сигнала, РП и НП,  $\omega_{dr}(\overline{t_r}) = 1$ ,  $\omega_{dr}(t) = 0$  для всех  $t \neq \overline{t_r}$ .

Полагаем, что расширенный функциональный базис  $\{\Psi(t), \Omega_\theta(t), \Omega_d(t)\}$  линейно независим на сетке  $C_0$  (по аналогии с [19, 20]). Шум  $\Xi_0$  характеризуется нулевым математическим ожиданием и корреляционной матрицей  $\mathbf{K}_{\Xi_0}$ .

Сразу оговоримся, что достаточно редкие аномальные выбросы шума  $\xi(t)$  в дальнейшем ассоциируются с НП, т.е. объединяются с соответствующими ненулевыми координатами вектора  $\mathbf{D}_0$ . Таким образом, в любой выборке  $\Xi_0$  все координаты можно заключить в стробы, размер которых определяется по одному из известных правил, например, по правилу «трех сигм».

Применительно к задаче распознавания будем рассматривать две основные подзадачи.

*Подзадача 1* (классическая задача оценивания параметров полезного сигнала) состоит в построении оптимальной оценки  $\mathbf{A}^*$  вектора  $\mathbf{A}$  на основе алгоритма, обеспечивающего автокомпенсацию РП и НП. *Подзадача 2* связана с построением оценок  $\mathbf{A}^*$  и  $\mathbf{B}^*$  для векторов  $\mathbf{A}$  и  $\mathbf{B}$  соответственно, инвариантных к НП. При этом оценивание параметров сигнала должно обеспечивать автокомпенсацию РП и, наоборот, оценивание параметров РП должно обеспечивать автокомпенсацию сигнала.

На базе (1)–(6) требуется разработать оптимальный метод распознавания (для двух основных подзадач 1 и 2) в условиях существенной априорной неопределенности (из статистической информации привлекается лишь корреляционная матрица  $\mathbf{K}_{\Xi_0}$ ), не требующий расширения пространства состояний. Смысл оптимальности для обеих подзадач состоит в том, что оценки должны быть несмещенными, обладать минимальным следом корреляционных матриц ошибок оценивания и обеспечивать существенный вычислительный эффект по сравнению с РМНК и МОИНО в плане декомпозиции и уменьшении объема вычислений, а также выигрыш по точности за счет обращения матриц меньшей размерности.

### **3. Общий подход к решению задачи распознавания на основе принципов непрерывности, размножения и ранжирования**

В основе предлагаемого подхода лежит принцип непрерывной зависимости качества распознавания от параметров временной сетки. В частности, последовательное уменьшение объема сетки достаточно большого объема (путем исключения тех или иных узлов) приводит к плавной эволюции точности результирующей оценки. Это относится и к непрерывному изменению положения узлов сетки на заданном временном интервале.

Предположим, что на базе основной сетки  $C_0$  и различных вариантов удаления некоторых ее узлов можно сформировать набор различных редуцированных сеток  $\{C_j, j = \overline{1, J-1}\}$  (где  $J \geq 2$ ,  $C_j = \{t_{jn}, n = \overline{1, N_j}\}$ ,  $N_j < N_0$ ,  $t_{jn} \in C_0$ ,  $t_{jn} \neq t_{jk} \forall n \neq k$ ,  $n, k \in \overline{1, N_j}$ ,  $t_{j,n+1} > t_{jn}$ ).

Все семейство сеток  $C = \{C_0, C_1, \dots, C_{J-1}\}$ , включая  $C_0$ , представим в виде

$$(7) \quad C = C^\# \cup C^\wedge \cup C^\&$$

Под символом  $C^\#$  понимается множество сеток, которые не содержат узлов, связанных с ненулевыми отсчетами НП. В качестве символа  $C^\wedge$  рассматривается множество сеток, которые могут содержать узлы, связанные с нормальными отсчетами НП. Речь идет об отсчетах НП, которые в совокупности слабо сказываются на результатах оценивания. Под символом  $C^\&$  понимается множество сеток, которые могут содержать узлы, связанные как с нормальными, так и аномальными отсчетами НП, которые в совокупности обесценивают результаты оценивания. Множества  $C^\#$  и  $C^\wedge$  будем называть допустимыми, а множество  $C^\&$  назовем недопустимым.

Основная идея разрабатываемого метода состоит в том, что, оперируя с  $C^\#$  и  $C^\wedge$ , можно получать оценки, инвариантные или практически инвариантные к РП и НП (принцип инвариантности). Для этого каждой сетке  $C_j$ ,  $j = \{0, 1, \dots, J-1\}$  поставим в соответствие уравнение наблюдения

$$(8) \quad \mathbf{H}_j = \mathbf{S}_j + \mathbf{\Theta}_j + \mathbf{D}_j + \mathbf{\Xi}_j$$

и оптимальные (согласно МОИНО) оценки

$$(9) \quad \begin{cases} \mathbf{A}_j^* = \mathbf{P}_j^{\mathbf{A}} \mathbf{H}_j, & s_j^*(t, \mathbf{A}_j^*) = (\mathbf{A}_j^*)^T \mathbf{\Psi}(t), \\ \mathbf{B}_{\theta_j}^* = \mathbf{P}_{\theta_j}^{\mathbf{B}} \mathbf{H}_j, & \theta_j^*(t, \mathbf{B}_{\theta_j}^*) = (\mathbf{B}_{\theta_j}^*)^T \mathbf{\Omega}_{\theta_j}(t), \end{cases}$$

где  $\mathbf{H}_j$  – вектор отсчетов входного наблюдения  $\mathbf{H}_0$ , соответствующий редуцированной сетке  $C_j$ ,  $\mathbf{P}_j^{\mathbf{A}}$  и  $\mathbf{P}_{\theta_j}^{\mathbf{B}}$  – матрицы линейного декомпозиционного оценивания:

$$(10) \quad \begin{cases} \mathbf{P}_j^{\mathbf{A}} = \left[ \mathbf{\Lambda}_{\theta_j}^{\mathbf{A}} \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Psi}_j \left( \mathbf{\Psi}_j^T \mathbf{\Lambda}_{\theta_j}^{\mathbf{A}} \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Psi}_j \right)^{-1} \right]^T, \\ \mathbf{P}_{\theta_j}^{\mathbf{B}} = \left[ \mathbf{\Lambda}_{\theta_j}^{\mathbf{B}} \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Omega}_{\theta_j} \left( \mathbf{\Omega}_{\theta_j}^T \mathbf{\Lambda}_{\theta_j}^{\mathbf{B}} \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Omega}_{\theta_j} \right)^{-1} \right]^T, \end{cases}$$

$$(11) \quad \begin{cases} \mathbf{\Lambda}_{\theta_j}^{\mathbf{A}} = \mathbf{E}_{N_j} - \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Omega}_{\theta_j} \left( \mathbf{\Omega}_{\theta_j}^T \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Omega}_{\theta_j} \right)^{-1} \mathbf{\Omega}_{\theta_j}^T, \\ \mathbf{\Lambda}_{\theta_j}^{\mathbf{B}} = \mathbf{E}_{N_j} - \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Psi}_j \left( \mathbf{\Psi}_j^T \mathbf{K}_{\Xi_j}^{-1} \mathbf{\Psi}_j \right)^{-1} \mathbf{\Psi}_j^T. \end{cases}$$

В (9)–(11) под  $\mathbf{\Psi}_j = [\psi_{jnm}, n = \overline{1, N_j}, m = \overline{1, M_s}]$  и  $\mathbf{\Omega}_{\theta_j} = [q_{jn} \omega_{\theta_j nr}, n = \overline{1, N_j}, r = \overline{1, M_{\theta_j}}]$  понимаются базисные матрицы сигнала  $\mathbf{S}_j$  и помехи  $\mathbf{\Theta}_j$  соответственно,  $\psi_{jnm} = \psi_m(t_{jn})$ ,  $\omega_{\theta_j nr} = \omega_{\theta_r}(t_{jn})$ ,  $\mathbf{E}_{N_j}$  – единичная матрица, соответствующая размерности вектора  $\mathbf{H}_j$ .

Если  $C_j \in C^\#$ , то указанные оценки отвечают соответствующим условиям минимума, несмещенности и инвариантности [19, 20], т.е. являются оптимальными. Если  $C_j \in C^\wedge$ , то получим квазиоптимальные оценки, а в случае  $C_j \in C^\&$  ошибки оценивания могут стать недопустимо большими.

В возможности формирования множества частных оценок, соответствующих семейству сеток  $C = \{C_0, C_1, \dots, C_{J-1}\}$ , удовлетворяющих условиям (2), (3), и состоит указанный выше принцип размножения.

Остановимся на подзадаче 1. Применительно к каждой паре сеток  $C_j$  и  $C_m$  сформируем скалярные невязки

$$\Delta_{\mathbf{A}_{jm}}(t) = s(t, \mathbf{A}_j^*) - s(t, \mathbf{A}_m^*) = [(\mathbf{A}_j^*) - (\mathbf{A}_m^*)]^T \Psi(t), \quad j, m = \overline{0, J-1}, \quad j > m,$$

а также их нормы  $\Delta_{\mathbf{A}_{jm}} = \|\Delta_{\mathbf{A}_{jm}}(t)\|$ ,  $j, m = \overline{0, J-1}$ , при этом можно использовать любую из норм соответствующего функционального пространства. На основе полученных норм строится вариационный ряд  $\Delta_{\mathbf{A}^{[v]}}$ ,  $v = \overline{1, (J^2 - J)/2}$  (монотонно возрастающая последовательность скалярных невязок), который полностью характеризует качество всех используемых сеток. Если в паре  $C_j$  и  $C_m$  хотя бы одна из сеток принадлежит множеству  $C^{\&}$ , то соответствующий ей элемент  $\Delta_{\mathbf{A}^{[v]}}$  будет находиться в «хвосте» вариационного ряда. Начальным элементам данного ряда будут отвечать пары  $C_j$  и  $C_m$ , отвечающие множествам  $C^{\#}$  и  $C^{\wedge}$ . Таким образом, при выполнении принятых ранее ограничений происходит гарантированная кластеризация некоторых элементов  $\Delta_{\mathbf{A}^{[v]}}$  вариационного ряда в достаточно малой окрестности  $(0, \delta_{\mathbf{A}})$ , где  $\delta_{\mathbf{A}} > 0$  – параметр усечения вариационного ряда. Константа  $\delta_{\mathbf{A}}$  для заданной сетки  $C_0$  выбирается заранее применительно к конкретной системе с учетом аналитических выражений для случайных и методических ошибок оценивания на основе МОИНО, приведенных в [19, 20], в рамках планирования измерительного эксперимента [5].

Все пары сеток, для которых не выполняется условие

$$(12) \quad \Delta_{\mathbf{A}^{[v]}} < \delta_{\mathbf{A}},$$

отбрасываются. Именно в этом состоит указанный ранее принцип ранжирования.

Введем ряд важных определений.

*Определение 1. Пара сеток  $C_j$  и  $C_m$  называется допустимой, если выполнено условие (12).*

*Определение 2. Произвольный отсчет НП называется аномальным, если соответствующий ему узел, присутствующий в любой сетке множества  $C$ , приводит к нарушению условия (12). В противном случае отсчет НП называется нормальным.*

*Определение 3. Группа отсчетов (из двух и более) называется аномальной (даже если каждый отсчет сам по себе является нормальным), когда соответствующие ей узлы, присутствующие в любой сетке множества  $C$ , приводят к нарушению условия (12). В противном случае группа отсчетов НП называется нормальной.*

*Определение 4. Под коэффициентом «засорения» выборки  $\mathbf{H}_0$  понимается число  $k_d = 100 (K_{\max}/N_0)$ , которое выражается в процентах.*

Поскольку все сетки множества  $C^{\&}$  приводят к нарушению условия (12), то они автоматически отсеиваются. Строгий критерий аномальности отсчетов входного наблюдения  $\mathbf{H}_0$  звучит так: отсчет является аномальным, если



он ни разу не встречается в парах сеток, удовлетворяющих условию (12). Данный критерий можно несколько смягчить и говорить о редких присутствиях «подозреваемого» (на аномальность) отсчета в указанных парах. Это особенно актуально для тех случаев, когда возникает неопределенность с тем, к какому классу (нормальных или аномальных) отнести тот или иной отсчет. Но в любом случае критерий позволяет отсеивать как аномальные, так и «подозреваемые» отсчеты входного наблюдения. Однако возможен ложный отсев некоторых нормальных отсчетов.

Критерий аномальности группы отсчетов звучит так: группа является аномальной, если она неоднократно встречается в отсеянных (недопустимых) парах сеток и ни разу в допустимых парах сеток (прошедших отсев).

В общем случае принцип ранжирования может приводить к ситуациям, когда количество хороших сеток, удовлетворяющих условию (12), может возрасти, т.е. появляется некоторая избыточность частных оценок. Следовательно, возникает законный вопрос, связанный с учетом этой избыточности с целью построения более надежной результирующей оценки.

Для формирования результирующей оценки  $\mathbf{A}^*$  вектора  $\mathbf{A}$  поступим следующим образом. Все сетки, которые вошли в допустимые пары, обозначим как  $C_j^*$ ,  $j = \overline{1, J^*}$  (в дальнейшем эти сетки будем называть конкурирующими). Для каждой фиксированной сетки  $C_j^*$  и всех других конкурирующих сеток  $C_m^*$  сформируем суммарную скалярную невязку

$$(13) \quad \Delta_{\mathbf{A}j}^* = \sum_{\substack{m=1 \\ m \neq j}}^{J^*} \Delta_{\mathbf{A}jm}.$$

Такая невязка показывает эффективность сетки  $C_j^*$  по сравнению со всеми другими конкурирующими сетками  $C_m^*$  (где  $m \neq j$ ,  $m \in \overline{1, J^*}$ ), входящими в допустимые пары.

Очевидно, что в условиях принятых моделей и ограничений критерий выбора оптимального номера  $j^* \in \{1, \dots, J^*\}$  оптимальной сетки  $C_{j^*}^* \in \{C_1^*, \dots, C_{J^*}^*\}$  выглядит так:

$$(14) \quad j^* = \arg \min_j \Delta_{\mathbf{A}j}^*.$$

Реализация данного критерия на практике не вызывает каких-либо затруднений, если построена матрица скалярных невязок для всех допустимых пар временных сеток. Среди отсчетов входного наблюдения, соответствующих оптимальной сетке  $C_{j^*}^*$ , отсутствуют аномальные отсчеты и аномальные группы отсчетов НП, а оценки

$$(15) \quad \begin{cases} \mathbf{A}_{j^*}^* = \mathbf{P}_{j^*}^{\mathbf{A}} \mathbf{H}_{j^*}, \\ s(t, \mathbf{A}_{j^*}^*) = (\mathbf{A}_{j^*}^*)^T \Psi(t) \end{cases}$$

будут оптимальными (по МОИНО) применительно к сетке  $C_{j^*}^*$ .

Остановимся на подзадаче 2. Она легко сводится к подзадаче 1, если ввести единый вектор оцениваемых параметров  $\mathbf{Z} = [\mathbf{A}^T, \mathbf{B}_\theta^T]^T$ . Тогда каждой сетке  $C_j$  с учетом (9) можно поставить в соответствие оптимальную (по МОИНО) оценку  $\mathbf{Z}_j^* = \left[ \left( \mathbf{A}_j^* \right)^T, \left( \mathbf{B}_{\theta j}^* \right)^T \right]^T$ . При этом нахождение оценки  $\mathbf{A}_j^*$  требует инвариантности к РП (т.е. параметр  $\Theta_0$  в (1) рассматривается как мешающий фактор). В свою очередь, нахождение оценки  $\mathbf{B}_{\theta j}^*$  требует инвариантности к сигналу  $\mathbf{S}_0$  (т.е. сигнал по отношению к РП теперь рассматривается как мешающий фактор). Для нахождения оценок  $\mathbf{A}_j^*$  и  $\mathbf{B}_{\theta j}^*$  используются матрицы  $\mathbf{P}_j^A$  и  $\mathbf{P}_{\theta j}^B$  соответственно, формируемые в соответствии с (10) и (11). Очевидно, что для решения подзадачи 2 достаточно в (12)–(14) заменить величины  $\Delta_{\mathbf{A}[v]}$ ,  $\delta_{\mathbf{A}}$ ,  $\Delta_{\mathbf{A}jm}$ ,  $\Delta_{\mathbf{A}j}^*$  на  $\Delta_{\mathbf{Z}[v]}$ ,  $\delta_{\mathbf{Z}}$ ,  $\Delta_{\mathbf{Z}jm}$ ,  $\Delta_{\mathbf{Z}j}^*$  соответственно.

Применительно к сетке  $C_{j^*}^*$  получаем решение задачи оптимальной идентификации РП

$$(16) \quad \begin{cases} \mathbf{B}_{\theta j^*}^* = \mathbf{P}_{\theta j^*}^B \mathbf{H}_{j^*}, \\ \theta(t, \mathbf{B}_{\theta j^*}^*) = (\mathbf{B}_{\theta j^*}^*)^T \boldsymbol{\Omega}_\theta(t). \end{cases}$$

С учетом (15) и (16) окончательная оценка для подзадачи 2 принимает вид  $\mathbf{Z}_{j^*}^* = \left[ \left( \mathbf{A}_{j^*}^* \right)^T, \left( \mathbf{B}_{\theta j^*}^* \right)^T \right]^T$ .

Для рассматриваемого метода принципиальным является отбор конкурирующих сеток  $C_j^*$ ,  $j = \overline{1, J^*}$  удовлетворяющих условию (12), и выбор из них оптимальной сетки  $C_{j^*}^*$ . Очевидно, что это напрямую связано с исходным множеством сеток  $\mathbf{C} = \{C_0, C_1, \dots, C_{J-1}\}$ , удовлетворяющих условиям (2) и (3). Современный прогресс в области построения ЭВМ параллельного действия (особенно на новых принципах [21–24]) позволяет надеяться, что предложенный принцип размножения сеток и частных оценок не станет препятствием для перспективных систем реального времени. Однако далеко не для всех существующих систем возможен такой подход, который на выборках большого объема может потребовать огромного числа каналов параллельной обработки данных и существенных вычислительных затрат. Поэтому наряду с оптимальным вариантом необходимо рассмотреть и квазиоптимальные подходы к формированию редуцированных сеток.

#### 4. Несколько вариантов построения редуцированных временных сеток

Принцип размножения сеток состоит в формировании в заданной временной области на базе некоторого правила  $\chi$  множества редуцированных сеток

$$(17) \quad \chi : C_0 \rightarrow \{C_j\}_{j=0}^{J-1}$$

таких, что при выполнении некоторых ограничений на НП среди формируемых на их основе частных оценок  $\left\{ \mathbf{A}_j^* \right\}_{j=0}^{J-1}$  и  $\left\{ \mathbf{Z}_j^* \right\}_{j=0}^{J-1}$  гарантированно будут не только плохие, но и хорошие оценки, для которых выполняется условие (12). Очевидно, что выбор правила  $\chi$  в общем случае неоднозначен и зависит от конкретной системы и условий наблюдения.

*Правило 1.* Пусть система оперирует с выборкой малого объема. Используя комбинаторику, сформируем на основе сетки  $C_0$  все возможные редуцированные сетки с числом узлов не менее  $N_{\min}$ . В этом случае количество таких сеток равно (метод полного перебора)

$$(18) \quad J = \sum_{m=0}^{N_0 - N_{\min}} C_{N_0}^{N_{\min} + m},$$

где  $C_{N_0}^{N_{\min} + m}$  – число сочетаний из  $N_0$  по  $N_{\min} + m$ .

Для фиксированного значения  $M_d$  количество сеток, не связанных с нулевыми отсчетами НП, определяется как

$$(19) \quad \bar{J} = \sum_{m=0}^{\bar{N}} C_{N_0 - M_d}^{N_{\min} + m},$$

где  $\bar{N} = N_0 - N_{\min} - M_d$ .

В частном случае, когда  $M_d = K_{\max}$ , можно сформировать минимальное число таких сеток  $\bar{J}_{\min} = \sum_{m=0}^{N_{\min}} C_{N_0 - K_{\max}}^{N_{\min} + m}$ , где  $\bar{N}_{\min} = N_0 - N_{\min} - K_{\max}$ .

Например, для случая  $N_0 = 5$ ,  $N_{\min} = 3$ ,  $K_{\max} = 2$  и  $M_d = 1$  имеем  $J = 16$ ,  $\bar{J}_{\min} = 1$  и  $\bar{J} = 5$ . Если отсчет НП (для случая  $M_d = 1$ ) является аномальным, то для пяти сеток условие ограниченности (12) выполнится, а для одиннадцати – нет. При этом среди хороших сеток четыре будут иметь длину  $N_j = 3$ ,  $j = \overline{1, 4}$ , а пятая сетка – длину  $N_5 = 4$ .

*Правило 2.* Для выборок большого объема более подходящим является правило  $\chi$ , основанное на представлении сетки  $C_0$  в виде  $L$  примыкающих друг к другу элементарных сеток  $\bar{C}_l$  (метод полного покрытия):

$$(20) \quad C_0 = \bigcup_{l=1}^L \bar{C}_l, \quad \bar{C}_l \cap \bar{C}_d = \emptyset, \quad \forall l \neq d \quad l, d \in \overline{1, L},$$

где  $\bar{C}_l = \{t_{l,m}, m = \overline{1, M_l}\}$  – элементарная сетка узлов  $t_{l,m} \in C_0$ ,  $l_m \in \overline{1, N_0}$ ,  $\sum_{l=1}^L M_l = N_0$ .

При формировании  $\bar{C}_l$  следует помнить, что ей соответствуют только такие узлы из сетки  $C_0$ , которые являются соседними, т.е. следуют друг за другом ( $t_{l,m+1} = t_{l+m}, m = \overline{0, M_l - 1}$ ). Проще всего рассматривать элементарные сетки  $\bar{C}_l$  одинаковой длины, т.е. принять  $M_l = M$  и  $L \cdot M = N_0$ .

Из  $L$  элементарных сеток, используя комбинаторику, также можно сформировать искомое семейство различных редуцированных сеток  $\{C_j, j = \overline{1, J-1}\}$ , число узлов в которых должно быть не меньше  $N_{\min}$ . В отличие от правила 1 теперь «подозрение» на возможную аномальность связываем с отдельными элементарными сетками. Если задаться максимально возможным количеством элементарных «подозрительных» сеток, то несложно найти общее число редуцированных сеток, которые не содержат данных областей, а также другие характеристики правила  $\chi$ .

*Правило 3.* Наряду с (20) можно использовать метод частичного покрытия

$$(21) \quad C_0 \supset \bigcup_{l=1}^L \bar{C}_l.$$

Этот метод может привести к минимизации вычислений, но сопряжен с возможными ошибками в принятии решений.

*Правило 4.* Чтобы избежать таких ошибок, допускается пошаговый вариант реализации правила  $\chi$  с использованием различных наборов элементарных сеток на каждом шаге. В этом случае предлагаются следующие пошаговые методы полного и частичного покрытия элементарными сетками длиной  $M_{li}$ :

$$(22) \quad \begin{cases} C_{0i} = \bigcup_{l=1}^{L_i} \bar{C}_{li}, & i = \overline{1, I}, \\ C_{0i} \supset \bigcup_{l=1}^{L_i} \bar{C}_{li}, & i = \overline{1, I}, \end{cases}$$

где  $i$  – номер шага,  $L_i > L$ .

Согласно (22) вместо  $C$  на каждом  $i$ -м шаге строится семейство сеток  $C_i$  и для всех его элементов проверяется выполнение условия (12). Если это условие не выполняется, то переходим к следующему шагу. Если на  $I$ -м шаге хотя бы для одной редуцированной сетки, входящей в состав  $C_I$ , данное условие выполнится, то данную сетку будем считать оптимальной. Очевидно, что пошаговый метод хорошо подходит для сетки  $C_0$  большого объема, поскольку требует незначительных вычислительных затрат (по сравнению с методом полного перебора), но может проигрывать по оперативности при больших значениях  $I$ . Однако при выполнении условий (2) и (3) такая ситуация возникает достаточно редко.

*Правило 5.* Еще более простой способ формирования редуцированных сеток возможен тогда, когда в составе системы присутствует анализатор входного наблюдения, позволяющий выявить такие области на отрезке  $[0, T]$ , которые «подозрительны» на предмет наличия в них НП. Исключая из  $C_0$  узлы, соответствующие данным областям, можно сформировать искомое семейство редуцированных сеток, удовлетворяющих условиям (2) и (3).

Выбор того или иного правила формирования редуцированных сеток целиком зависит от типа рассматриваемой системы и предъявляемых к ней требований. Соотношения типа (17)–(22) являются достаточно удобными для количественного обоснования формируемых редуцированных сеток.

## 5. Алгоритмы распознавания в условиях неопределенности

Алгоритм распознавания для подзадачи 1.

**Шаг 1.1.** На базе  $C_0$  строим набор сеток  $\{C_j, j = \overline{0, J-1}\}$ .

**Шаг 1.2.** Для каждой сетки  $C_j$  находим матрицу оптимального оценивания  $P_j^A$ .

**Шаг 1.3.** Для каждой сетки  $C_j$  находим оценку  $A_j^* = P_j^A H_j$ .

**Шаг 1.4.** Для каждой пары  $C_j$  и  $C_m$  формируем скалярную невязку  $\Delta_{Ajm}$ .

**Шаг 1.5.** На основе всех невязок  $\Delta_{Ajm}$  строим вариационный ряд  $\Delta_{A[v]}$ ,  $v = \overline{1, (J^2 - J)/2}$ .

**Шаг 1.6.** На базе вариационного ряда с учетом условия (12) формируем набор конкурирующих сеток  $C_j^*$ ,  $j = \overline{1, J^*}$ .

**Шаг 1.7.** Для конкурирующих сеток вычисляем суммарные невязки  $\Delta_{Aj}^*$ ,  $j = \overline{1, J^*}$ .

**Шаг 1.8.** На базе критерия (14) находим оптимальный номер  $j^* \in \{1, \dots, J^*\}$  оптимальной сетки  $C_{j^*}^*$ .

**Шаг 1.9.** Используя матрицу  $P_{j^*}^A$ , вычисляем применительно к сетке  $C_{j^*}^*$  оценку  $A_{j^*}^*$  для вектора  $A$  и оценку  $s(t, A_{j^*}^*) = (A_{j^*}^*)^T \Psi(t)$  для полезного сигнала  $s(t, A)$ .

Алгоритм распознавания для подзадачи 2.

**Шаг 2.1.** На базе  $C_0$  строим набор сеток  $\{C_j, j = \overline{0, J-1}\}$ .

**Шаг 2.2.** Для каждой сетки  $C_j$  находим матрицы оптимального оценивания  $P_j^A$  и  $P_{\theta j}^B$ .

**Шаг 2.3.** Для каждой сетки  $C_j$  находим оценки  $A_j^*$  и  $B_{\theta j}^*$ .

**Шаг 2.4.** Формируем единый вектор  $Z_j^* = \left[ (A_j^*)^T, (B_{\theta j}^*)^T \right]^T$ .

**Шаг 2.5.** Для каждой пары  $C_j$  и  $C_m$  формируем скалярную невязку  $\Delta_{Zjm}$ .

**Шаг 2.6.** На основе всех невязок  $\Delta_{Zjm}$  строим вариационный ряд  $\Delta_{Zj}^*$ ,  $j = \overline{1, J^*}$ .

**Шаг 2.7.** На базе вариационного ряда с учетом условия (12) формируем набор конкурирующих сеток  $C_j^*$ ,  $j = \overline{1, J^*}$ .

**Шаг 2.8.** Для конкурирующих сеток вычисляем суммарные невязки  $\Delta_{Zj}^*$ ,  $j = \overline{1, J^*}$ .

**Шаг 2.9.** На базе критерия (14) находим оптимальный номер  $j^* \in \{1, \dots, J^*\}$  оптимальной сетки  $C_{j^*}^*$ .

**Шаг 2.10.** Используя матрицы  $\mathbf{P}_{j^*}^{\mathbf{A}}$  и  $\mathbf{P}_{\theta j^*}^{\mathbf{B}}$ , вычисляем применительно к сетке  $\mathbf{C}_{j^*}^*$  оценку  $\mathbf{A}_{j^*}^*$  для вектора  $\mathbf{A}$  и оценку  $s(t, \mathbf{A}_{j^*}^*) = (\mathbf{A}_{j^*}^*)^T \Psi(t)$  для полезного сигнала  $s(t, \mathbf{A})$ , а также оценку  $\mathbf{B}_{\theta j^*}^*$  для вектора  $\mathbf{B}_\theta$  и оценку  $\theta(t, \mathbf{B}_{\theta j^*}^*) = (\mathbf{B}_{\theta j^*}^*)^T \Omega_\theta(t)$  для РП  $\theta(t, \mathbf{B}_\theta)$  соответственно.

## 6. К сравнительному анализу метода

Корреляционные матрицы ошибок оценивания применительно к сетке  $\mathbf{C}_{j^*}^*$  находятся по правилу

$$(23) \quad \begin{cases} \mathbf{K}_{j^*}^{\mathbf{A}} = \mathbf{P}_{j^*}^{\mathbf{A}} \mathbf{K}_{\Xi_{j^*}} (\mathbf{P}_{j^*}^{\mathbf{A}})^T, \\ \mathbf{K}_{\theta j^*}^{\mathbf{B}} = \mathbf{P}_{\theta j^*}^{\mathbf{B}} \mathbf{K}_{\Xi_{j^*}} (\mathbf{P}_{\theta j^*}^{\mathbf{B}})^T. \end{cases}$$

Выражение (23) позволяет в каждом конкретном случае оценить потенциальные возможности метода с учетом требований, предъявляемых к системе.

Согласно используемому критерию оптимальности данные матрицы имеют минимальные следы, т.е.  $Sp \mathbf{K}_{j^*}^{\mathbf{A}} \rightarrow \min$  и  $Sp \mathbf{K}_{\theta j^*}^{\mathbf{B}} \rightarrow \min$  в классе всех линейных оценок. Методическая ошибка, обусловленная конечностью представлений (4) и (5), может быть учтена следующим образом. Пусть уравнение наблюдений для оптимальной сетки  $\mathbf{C}_{j^*}^*$  имеет вид (с учетом компенсации НП)

$$(24) \quad \mathbf{H}_{j^*} = (\mathbf{S}_{j^*} + \Delta \mathbf{S}_{j^*}) + (\Theta_{j^*} + \Delta \Theta_{j^*}) + \Xi_{j^*},$$

где  $\Delta \mathbf{S}_{j^*}$  и  $\Delta \Theta_{j^*}$  – добавки к сигналу и РП, обусловленные учетом «хвостов» используемых функциональных рядов.

В этом случае для оценки  $\mathbf{A}_{j^*}^*$  вектора  $\mathbf{A}$  (вычисленной без учета этих добавок) имеем

$$(25) \quad \mathbf{A}_{j^*}^* = \mathbf{P}_{j^*}^{\mathbf{A}} (\mathbf{S}_{j^*} + \Delta \mathbf{S}_{j^*}) + \mathbf{P}_{j^*}^{\mathbf{A}} (\Theta_{j^*} + \Delta \Theta_{j^*}) + \mathbf{P}_{j^*}^{\mathbf{A}} \Xi_{j^*},$$

где матрица оптимального оценивания  $\mathbf{P}_{j^*}^{\mathbf{A}}$  вычислена с учетом конечных представлений (4) и (5) путем отбрасывания указанных «хвостов».

Соответственно для истинного значения  $\mathbf{A}$  справедливо представление (полагая  $\Xi_{j^*} = \mathbf{0}$ )

$$(26) \quad \mathbf{A} = (\mathbf{P}_{j^*}^{\mathbf{A}} + \Delta \mathbf{P}_{j^*}^{\mathbf{A}}) (\mathbf{S}_{j^*} + \Delta \mathbf{S}_{j^*}) + (\mathbf{P}_{j^*}^{\mathbf{A}} + \Delta \mathbf{P}_{j^*}^{\mathbf{A}}) (\Theta_{j^*} + \Delta \Theta_{j^*}).$$

В качестве среднего значения методической ошибки можно принять величину

$$(27) \quad \overline{\Delta \mathbf{A}_{j^*}} = \mathbf{M} \{ \mathbf{A} - \mathbf{A}_{j^*}^* \} = \Delta \mathbf{P}_{j^*}^{\mathbf{A}} (\mathbf{S}_{j^*} + \Delta \mathbf{S}_{j^*}) + \Delta \mathbf{P}_{j^*}^{\mathbf{A}} (\Theta_{j^*} + \Delta \Theta_{j^*}),$$

где  $\mathbf{M}\{\cdot\}$  – символ математического ожидания с учетом, что  $\mathbf{M}\{\Xi_{j^*}\} = \mathbf{0}$ .

Используя (23)–(27), можно подобрать необходимые параметры развиваемого метода, обеспечивающие минимизацию результирующей ошибки оценивания в каждом конкретном случае. Необходимые и достаточные условия существования и единственности решения задачи оценивания в рамках развиваемого метода требуют невырожденности и соблюдения некоторых ограничений на ранги ряда матриц (по аналогии с [19, 20]). Выполнение заданных условий на практике обеспечивается рациональным выбором используемых функциональных базисов, числа степеней свободы в моделях сигнала и РП, а также заданием соответствующих условий наблюдения. Все эти вопросы относятся к планированию вычислительного эксперимента и далее не рассматриваются, поскольку требуют отдельных исследований в каждом конкретном случае.

В дальнейшем для сравнительного анализа рассмотрим ряд методов:  $M_{(1)}$  (это РМНК),  $M_{(2)}$  (это МОИНО) и  $M_{(3)}$  (это разработанный метод). Каждому возможному варианту НП можно поставить в соответствие одну из гипотез  $\Gamma_l$ ,  $l = \overline{1, L}$ . Очевидно, что  $L = J$  и для фиксированного  $l$ , по аналогии с (1) и (8), справедливо модельное наблюдение

$$\mathbf{H}_{0l} = \mathbf{S}_0 + \mathbf{\Theta}_0 + \mathbf{D}_{0l} + \mathbf{\Xi}_0.$$

Задача распознавания на базе  $M_{(1)}$  решается с использованием критерия минимума квадратичной формы  $\chi(\mathbf{Z}_{dl})$ :

$$(28) \quad \begin{aligned} (l^*, \mathbf{Z}_{dl}^*) &= \arg \min_{l, \mathbf{Z}_{dl}} \chi(\mathbf{Z}_{dl}) = \\ &= \arg \min_{l, \mathbf{Z}_{dl}} [\Delta(l, \mathbf{Z}_{dl})]^\top (\mathbf{K}_{\Xi_0})^{-1} [\Delta(l, \mathbf{Z}_{dl})], \quad i \in \{1, 2, 3\}, \end{aligned}$$

где  $\Delta(l, \mathbf{Z}_{dl}) = \mathbf{H}_0 - \mathbf{H}_{0l}$ ,  $\mathbf{H}_{0l} = \mathbf{H}_{0l}(l, \mathbf{Z}_{dl})$ ,  $\mathbf{Z}_{dl} = [\mathbf{Z}^\top, \mathbf{B}_{dl}^\top]^\top = [\mathbf{A}^\top, \mathbf{B}_\theta^\top, \mathbf{B}_{dl}^\top]^\top$ .

Одним из показателей эффективности  $M_{(i)}$  является размерность обрабатываемых матриц. Несложно убедиться, что  $M_{(1)}$  требует для фиксированного  $l$  обращения матрицы размером  $\rho_{(1)l} = M_s + M_\theta + M_{dl}$ , где  $M_{dl}$  – количество ненулевых координат в векторе  $\mathbf{D}_{0l}$ . В свою очередь  $M_{(2)}$  предполагает формирование объединенной базисной матрицы РП и НП, что для фиксированного  $l$  приводит к обращению двух матриц размером  $\rho_{(2)s} = M_s$  и  $\rho_{(2)\theta dl} = M_\theta + M_{dl}$  соответственно. Независимо от номера рассматриваемой гипотезы  $M_{(3)}$  требует обращения двух матриц размером  $\rho_{(3)s} = \rho_{(2)s} = M_s$  и  $\rho_{(3)\theta} = M_\theta$ . С учетом этого для выбора номера наиболее предпочтительного метода можно воспользоваться критерием:

$$i^* = \arg \min_i \rho_{(i)},$$

где  $i, i^* \in \{1, 2, 3\}$ ,  $\rho_{(1)} = \max_l \{M_s + M_\theta + M_{dl}\}$ ,  $\rho_{(2)} = \max_l \{M_s, M_\theta + M_{dl}\}$ ,  $\rho_{(3)} = \max_l \{M_s, M_\theta\}$ .

Видим, что во всех случаях  $\rho_{(3)} \leq \rho_{(1)}$  и  $\rho_{(3)} \leq \rho_{(2)}$ , а при отсутствии НП (но наличии РП) имеем  $\rho_{(3)} < \rho_{(1)}$  и  $\rho_{(3)} = \rho_{(2)}$ . Когда РП и НП отсутствуют, то  $\rho_{(1)} = \rho_{(2)} = \rho_{(3)}$ .

Анализ показывает, что при высокой размерности РП и НП, выполнении условия  $M_\theta + M_{dl} > M_s$ , а также плохой обусловленности обрабатываемых матриц  $M_{(3)}$  в плане вычислительной устойчивости гораздо предпочтительнее  $M_{(1)}$  и  $M_{(2)}$  (в этом случае всегда  $\rho_{(3)} < \rho_{(1)}$  и  $\rho_{(3)} < \rho_{(2)}$ ). Для данных условий наряду с  $\rho_{(1)}$ ,  $\rho_{(2)}$  и  $\rho_{(3)}$  необходимо использовать числа обусловленности  $\mu_{(1)}$ ,  $\mu_{(2)}$  и  $\mu_{(3)}$ , являющиеся характеристиками устойчивости сравниваемых методов.

Для сравнительной оценки вычислительной сложности сравниваемых методов будем использовать следующие характеристики:  $V_{(i)}^\Sigma$  – суммарное количество узлов исходной сетки  $S_0$ , используемое для проверки всех гипотез,  $Q_{(i)}^\Sigma$  и  $T_{(i)}^\Sigma$  – соответственно суммарное количество операций (сложения и умножения) и время, необходимые для реализации  $M_{(i)}$ . Для сравнительной оценки точности используются характеристики:  $\Delta s_{(i)}$  и  $\Delta \theta_{(i)}$  – результирующие ошибки оценивания сигнала и РП соответственно. При этом  $\Delta s_{(i)} = \max_t |s(t, \mathbf{A}) - s(t, \mathbf{A}_{(i)}^*)|$  и  $\Delta \theta_{(i)} = \max_t |\theta(t, B_\theta) - \theta(t, B_{\theta(i)}^*)|$ .

Опуская промежуточные выкладки, приводим окончательные результаты:

$$V_{(1)}^\Sigma = V_{(2)}^\Sigma = LN_0, \quad V_{(3)}^\Sigma = \sum_{m=0}^{N_0 - N_{\min}} C_{N_0}^{N_{\min} + m} (N_0 - m)$$

– для характеристики  $V_{(i)}^\Sigma$ ,

$$Q_{(1)}^\Sigma = \sum_{m=0}^{N_0 - N_{\min}} \left\{ C_{N_0}^{N_{\min} + m} \left[ 2N_0^3 + 6N_0^2 \gamma_m + N_0 (4\gamma_m^2 - 3\gamma_m) + 2\gamma_m^3 - \gamma_m^2 \right] \right\},$$

$$\gamma_m = M_s + M_\theta + m,$$

$$Q_{(2)}^\Sigma = \sum_{m=0}^{N_0 - N_{\min}} \left\{ C_{N_0}^{N_{\min} + m} \left[ 2(N_0^3 + N_0^2 \varphi_m) + N_0 (4\beta_m - \varphi_m) + 2(\alpha_m - \varphi_m) \right] \right\},$$

$$\beta_m = \eta_m^2 + M_s^2, \quad \alpha_m = \eta_m^3 + M_s^3, \quad \varphi_m = 2\eta_m + 3M_s, \quad \eta_m = M_\theta + m,$$

$$Q_{(3)}^\Sigma = \sum_{m=0}^{N_0 - N_{\min}} \left\{ C_{N_0}^{N_{\min} + m} \left[ 2(N_0 - m)^3 + 2(N_0 - m)^2 \varphi_0 + \right. \right. \\ \left. \left. + (N_0 - m)(4\beta_0 - \varphi_0) + 2\alpha_0 - \beta_0 \right] \right\},$$

$$\alpha_0 = (M_\theta + M_d)^3 + M_s^3, \quad \beta_0 = M_\theta^2 + M_s^2, \quad \varphi_0 = 2M_\theta + 3M_s, \quad \eta_0 = M_\theta$$

– для характеристики  $Q_{(i)}^\Sigma$ .



Анализ данных выражений показывает, что за счет обработки данных на редуцированных сетках меньшего объема (по сравнению с исходной сеткой  $C_0$ ) можно добиться существенного вычислительного выигрыша  $M_{(3)}$  по сравнению с  $M_{(1)}$  и  $M_{(2)}$ . Этот выигрыш также возрастает с увеличением размерности НП.

Расчет характеристик  $T_{(i)}^\Sigma$  и  $\Delta\theta_{(i)}^\Sigma$  зависит как от исходных данных задачи, так и от возможностей используемой вычислительной среды.

Для проверки семейства гипотез можно организовать  $L$  каналов параллельной обработки данных для выполнения большого объема векторно-матричных вычислений. Однако следует помнить, что возможна существенная оптимизация таких вычислений, поскольку все используемые в различных каналах векторы и матрицы получены соответствующей редукцией векторов и матриц, относящихся к базовой сетке  $C_0$ , т.е. большинство операций в различных каналах повторяется.

Предложенный метод несложно распространить и на другие задачи (например, обнаружения и различения), связанные с распознаванием сигналов (в том числе и с нелинейными параметрами [20]) и решаемые в рамках теории гипотез. В этом случае количество каналов параллельной обработки данных может существенно возрасти, поскольку теперь построение редуцированных сеток надо проводить непосредственно для каждой гипотезы и для каждого узла области определения нелинейного параметра. В этом случае наиболее эффективными для борьбы с НП являются методы полного (20) и частично (21) покрытия, например, пошагового (22).

## 7. Иллюстративный пример

Применим развитый метод к решению задачи однопозиционной пассивной дальнометрии для излучающей цели с частично известными параметрами движения только по энергетическим измерениям (пассивно-энергетический способ [25, 26]). Пусть цель движется в направлении на дальномер, при этом наклонная дальность изменяется по закону  $R(t) = R_0 + R_0^{(1)}t + 2^{-1}R_0^{(2)}t^2$ , где  $R_0$ ,  $R_0^{(1)}$  и  $R_0^{(2)}$  – соответственно начальная дальность, радиальная скорость и радиальное ускорение для момента времени  $t = 0$ . В качестве измеряемого сигнала (энергетического параметра) примем  $s(t) = 1 - q_0^{-1}(t)$ , где  $q_0(t) = [p_0^{-1}p(t)]^{1/2}$ ,  $p(t)$  – мощность электромагнитной волны на входе дальномера,  $p_0 = p(0)$ .

Воспользуемся известным уравнением пассивной локации для стационарного канала:

$$p = \zeta_0 R^{-2},$$

где  $p = p(t)$ ,  $R = R(t)$ ,  $\zeta_0 = \text{const}$  – обобщенный коэффициент, связывающий мощность и дальность.

Такая упрощенная модель встречается на практике при выполнении ряда ограничений на условия наблюдений за излучающей целью [25, 26]. В рас-

смаатриваемом здесь случае она позволит более наглядно продемонстрировать эффективность разработанного метода, не прибегая к сложным алгоритмам дальнометрии для нестационарного случая.

Воспользуемся следующей формулой для искомой дальности:

$$(29) \quad R(t) = s^{-1}(t)D_0(t)[1 - s(t)], \quad t > 0,$$

где  $D_0(t) = \left( R_0^{(1)}t + 2^{-1}R_0^{(2)}t^2 \right)$  – расстояние, пройденное целью за время  $t$ .

Формула (29) является основой для расчета дальности на основе оценок параметра  $s(t) = 1 - q_0^{-1}(t)$ . Несложно заметить, что  $s(t)$  можно представить в виде

$$s(t) = - \left( R_0^{-1}R_0^{(1)}t + 2^{-1}R_0^{-1}R_0^{(2)}t^2 \right), \quad t > 0,$$

следовательно, с учетом (4) имеем  $s(t) = s(t, a_1, a_2) = a_1\psi_1(t) + a_2\psi_2(t)$ , т.е.  $M_s = 2$ ,  $a_1 = -R_0^{-1}R_0^{(1)}$ ,  $a_2 = -2^{-1}R_0^{-1}R_0^{(2)}$ ,  $\psi_1(t) = t$  и  $\psi_2(t) = t^2$ . Таким образом, после оценивания коэффициентов  $a_1$  и  $a_2$  (по измерениям энергетического параметра  $s(t, a_1, a_2)$ ) можно найти искомую дальность по формуле (29). Далее для оценивания будем использовать  $M_{(3)}$ , а для сравнительного анализа также  $M_{(1)}$  и  $M_{(2)}$ . В дальнейшем все временные параметры измеряются в секундах, расстояния – в метрах, радиальная скорость – в метр/секунда, радиальное ускорение – в метр/секунда<sup>2</sup>, коэффициент «засорения» выборки – в процентах, энергетический параметр – безразмерная величина.

Для эксперимента примем  $R_0 = 3 \times 10^4$ ,  $R_0^{(1)} = 2 \times 10^3$ ,  $R_0^{(2)} = 2 \times 10^2$ ,  $T = 10$ ,  $N_0 = 10$ ,  $t_n - t_{n-1} = \Delta t = 1$ ,  $t_1 = 1$ ,  $t_{10} = 10$ ,  $M_d = K_{\max} = 2$ ,  $N_{\min} = 6$ ,  $k_d = 20$ ,  $M_\theta = 0$  (т.е. влияние РП не учитывалось). Шумы измерений моделировались с использованием датчика случайных чисел согласно нормальному закону распределения с диагональной матрицей  $\mathbf{K}_{\Xi_0}$  с ненулевым элементом  $\sigma_0^2 = 10^{-10}$ . Для случаев, когда  $N \geq N_{\min}$ , и при отсутствии НП усредненная (по ансамблю из 500 реализаций наблюдения) среднеквадратическая ошибка оценивания дальности любым методом должна удовлетворять условию  $\overline{\Delta R} \leq 5 \times 10^2$ .

Вычисления выполнялись в среде MATLAB (версия R20119b) на компьютере с 4-ядерным процессором, тактовой частотой 2,6 ГГц и ОЗУ 8 Гб DDR3.

В табл. 1 представлены точные значения энергетического параметра для сетки  $C_0$ .

**Таблица 1**

|       |       |       |       |       |       |       |       |       |       |    |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----|
| $t_n$ | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10 |
| $s_n$ | 0,067 | 0,147 | 0,230 | 0,320 | 0,417 | 0,520 | 0,630 | 0,747 | 0,870 | 1  |

Для принятых исходных данных были вычислены промежуточные значения:  $L = J = 386$ ,  $\gamma_m = 2 + m$ ,  $\beta_m = m^2 + 4$ ,  $\alpha_m = m^3 + 8$ ,  $\varphi_m = 2m + 6$ ,  $\eta_m = m$ ,  $0 \leq m \leq 4$ .

Окончательные результаты вычислительного эксперимента (с учетом округлений) сведены в табл. 2.

**Таблица 2**

| Характеристики<br>$M_{(i)}$ | $\rho_{(i)}$ | $\mu_{(i)}$ | $V_{(i)}^{\Sigma}$ | $Q_{(i)}^{\Sigma}$ | $T_{(i)}^{\Sigma}$   | $\overline{\Delta R_{(i)}}$ |
|-----------------------------|--------------|-------------|--------------------|--------------------|----------------------|-----------------------------|
| $M_{(1)}$                   | 6            | 544         | $3,9 \times 10^3$  | $2,5 \times 10^6$  | $1,3 \times 10^{-3}$ | $1,3 \times 10^3$           |
| $M_{(2)}$                   | 4            | 148         | $3,9 \times 10^3$  | $1,9 \times 10^6$  | $9,9 \times 10^{-4}$ | $9,3 \times 10^2$           |
| $M_{(3)}$                   | 2            | 103         | $2,6 \times 10^3$  | $4,7 \times 10^5$  | $2,4 \times 10^{-4}$ | $8,2 \times 10^2$           |

Данные этой таблицы наглядно показывают, что РМНК и МОИНО по всем характеристикам проигрывают разработанному методу даже при невысокой размерности обрабатываемых матриц. Используя [26], можно рассмотреть нестационарный случай (когда  $\zeta_0 \neq \text{const}$ ) с учетом не только НП, но и РП. В таких условиях расширяется вектор оцениваемых параметров, что требует наращивания объема исходной сетки  $S_0$ . Достижимый эффект при этом существенно возрастает с увеличением размерности задачи и указанного объема.

Рассмотренный в данном примере способ дальнометрии в совокупности с предложенным алгоритмом оценивания энергетического параметра (на базе разработанного метода) может использоваться как самостоятельно, так и в комплексе с другими известными способами пассивной локации. Первый случай может быть актуален, когда в дальнометре единственным источником достоверной информации является энергетический канал (например, либо при однопозиционном энергетическом варианте построения дальнометра, либо при «деградации» структуры двух- или многопозиционной системы: выход из строя отдельных позиций, нарушение нормальной работы угломерных каналов, линий связи и т.д.). Способность автономного решения задачи дальнометрии с учетом не только регулярных, но и нерегулярных помех может найти широкое применение в различных областях применительно к пассивным системам локации.

Второй случай позволяет рассматривать пассивно-энергетический метод как альтернативный при построении комплексного алгоритма функционирования интеллектуальной системы, способной надежно функционировать в самых различных, и даже неблагоприятных, условиях. Основу такого алгоритма помимо предлагаемого метода могут составить угломерно-энергетический, и широко применяющиеся на практике пассивные методы: разностно-дальномерный, триангуляционный и другие, а также их различные модификации.

Поскольку энергетические измерения традиционно не относятся к классу надежных измерений, то предлагаемый метод распознавания сигналов в условиях РП и НП может позволить повысить эффективность функционирования энергетических каналов, обеспечивающих требуемую точность амплитудных и мощностных измерений.

## 8. Заключение

Развитый метод распознавания сигналов, реализующий принципы непрерывности, инвариантности, размножения и ранжирования с использованием семейства редуцированных сеток, может эффективно сочетаться с алгоритмами ортогональных разложений и решения некорректных задач [27–29]. Возможность декомпозиции вычислительных процедур, снижения размерности обрабатываемых матриц и уменьшения объема вычислений позволяет более эффективно решать целый круг прикладных задач, связанных с обработкой измерений в различных областях. Полученные алгоритмы распознавания сигналов в условиях РП и НП несложно реализовать в специализированных многоканальных ЭВМ, ориентированных на системы, предназначенные для функционирования в реальном времени.

Приведены компактные аналитические выражения, позволяющие заранее, под конкретную прикладную задачу, подобрать необходимые модели сигналов и помех, а также количественные значения их параметров, при которых разработанный метод обеспечит достижение своих потенциальных возможностей. Все вычислительные процедуры в каждом канале сводятся к простейшим математическим операциям над векторами и матрицами, допускается возможность комбинирования данного метода с традиционными подходами к решению прикладных задач, связанных с оптимальной и квазиоптимальной обработкой измерений.

Достижения в области параллельных вычислений позволяют надеяться, что любые задачи, связанные с распознаванием сигналов, в скором времени могут быть решены с использованием предлагаемого метода.

## СПИСОК ЛИТЕРАТУРЫ

1. *Репин В.Г., Тартаковский Г.П.* Статистический синтез при априорной неопределенности и адаптация информационных систем. М.: Сов. радио, 1977.
2. *Сосулин Ю.Г., Костров В.В., Паршин Ю.Н.* Оценочно-корреляционная обработка сигналов и компенсация помех. М.: Радиотехника, 2014.
3. *Сергиенко А.Б.* Цифровая обработка сигналов. СПб.: БХВ-Петербург, 2011.
4. *Богданович В.А., Вострецов А.Г.* Теория устойчивого обнаружения, различения и оценивания сигналов. М.: Физматлит, 2004.
5. *Жданюк Б.Ф.* Основы статистической обработки траекторных измерений. М.: Сов. радио, 1978.
6. *Ярлыков М.С.* Применение марковской теории нелинейной фильтрации в радиотехнике. М.: Сов. радио, 1980.
7. *Фомин А.Ф., Новоселов О.Н., Плющев А.В.* Отбраковка аномальных результатов измерений. М.: Энергоатомиздат, 1985.
8. *Гаджиев Ч.М.* Подход к отбраковке аномальных измерений, робастный к систематическим погрешностям // *Автометрия* 2003. Т. 39. № 4. С. 39–46.
9. *Методы цифровой обработки сигналов для решения прикладных задач / Под ред. В.И. Марчука.* М.: Радиотехника, 2012.

10. Бульчев Ю.Г., Васильев В.В., Джуган Р.В. и др. Информационно-измерительное обеспечение натуральных испытаний сложных технических комплексов. М.: Машиностроение – Полет, 2016.
11. Денисов В.П., Дубинин Д.В., Мещеряков А.А. Исключение аномально больших ошибок пеленгования в процессе устранения неоднозначности измерений в фазовых пеленгаторах, реализующих метод максимального правдоподобия // Радиотехника и электроника. 2016. Т. 61. № 10. С. 957–963.
12. Шэнь К., Шахтарин Б.И., Неусыпин Б.И., Нгуен Д. Алгоритмические методы коррекции навигационной информации с использованием спутниковой радионавигационной системы в условиях аномальных измерений // Радиотехника и электроника. 2019. Т. 64. № 1. С. 31–37.
13. Иванов А.В., Шишкин В.Ю., Бойков Д.В. и др. Адаптивные алгоритмы обработки информации в навигационных комплексах подвижных наземных объектов // Радиотехника и электроника. 2021. Т. 66. № 8. С. 760–771.
14. Калянов А.А., Лукин О.В., Цыганова Ю.В. Об алгоритме адаптивной фильтрации параметров движения объекта // Автоматизация процессов управления. 2023. № 1 (71). С. 75–87.
15. Gu T., Tu Y., Tang D., Luo T. A Robust Moving Total Least-Squares Fitting Method for Measurement Data // IEEE Transact. Instrument. Measurement. 2020. V. 69. No. 10. P. 7566–7573.
16. Gu T., Luo Z., Guo T., Luo T. A New Reconstruction Method for Measurement Data with Multiple Outliers // IEEE Transact. Instrument. Measurement. 2022. V. 71. P. 1–9.
17. Ji C., Song C., Li S., et. al. An Online Combined Compensation Method of Geomagnetic Measurement Error // IEEE Sensor. J. 2022. V. 22. No. 14. P. 14026–14037.
18. Gao G., Gao B., Gao S., et. al. Hypothesis Test-Constrained Robust Kalman Filter for INS/GNSS Integration with Abnormal Measurement // IEEE Transact. Vehicular Tech. 2023. V. 72. No. 2. P. 1662–1673.
19. Бульчев Ю.Г., Елисеев А.В. Вычислительная схема инвариантно-несмещенного оценивания значений линейных операторов заданного класса // Журн. вычисл. матем. и мат. физ. 2008. Т. 48. № 4. С. 580–592.
20. Бульчев Ю.Г. Распознавание сигналов без расширения пространства состояний по результатам наблюдений, содержащих сингулярную помеху, с учетом нелинейности // АИТ. 2024. № 2. С. 81–102.
21. Ежова Н.А., Соколинский Л.Б. Обзор моделей параллельных вычислений // Вестник ЮУрГУ. Серия «Вычислительная математика и информатика». 2019. Т. 8. № 3. С. 58–91.
22. Иванов А.И., Шпилевая С.Г. О квантовых параллельных вычислениях // Вестник Балт. ун-та им. Канта. Серия «Физико-математические и технические науки». 2021. № 2. С. 95–99.
23. Sutti C. Lokal and global optimization by parallel algorithms for MIMD systems // Ann. Oper. Res. 1984. V. 1. P. 151–164.
24. Price W.L. Global optimization algorithms for a CAD workstation // J. Optim. Theory Appl. 1987. V. 55. P. 133–146.
25. Мельников Ю.П., Попов С.В. Радиотехническая разведка. Методы оценки эффективности местоопределения источников излучения. М.: Радитехника, 2008.

26. Бульчев Ю.Г., Ивакина С.С., Мозоль А.А., Насенков И.Г. Анализ модификации энергетического метода пассивной дальнометрии // Автометрия. 2016. Т. 52. № 1. С. 37–44.
27. Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов. М.: Наука, 1986.
28. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1986.
29. Бакушинский А.Б., Гончарский А.В. Некорректные задачи. Численные методы и приложения. М.: Изд-во МГУ, 1989.

*Статья представлена к публикации членом редколлегии О.Н. Граничиным.*

Поступила в редакцию 27.06.2024

После доработки 13.02.2025

Принята к публикации 17.02.2025

# Стохастические системы

© 2025 г. А.В. ЛЕБЕДЕВ, д-р физ.-мат. наук (avlebed@yandex.ru)  
(Московский государственный университет им. М.В. Ломоносова)

## МНОГОМЕРНЫЕ НЕПРЕРЫВНЫЕ РАСПРЕДЕЛЕНИЯ И КОПУЛЫ, ПОРОЖДАЮЩИЕ НЕТРАНЗИТИВНЫЕ НАБОРЫ ЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Работа продолжает авторский цикл по изучению феномена нетранзитивности отношения стохастического предшествования в теории вероятностей. Исходя из парадокса Кондорсе, построены примеры трехмерных непрерывных распределений и копул, порождающих нетранзитивные наборы зависимых случайных величин. Доказаны предельные теоремы для многомерных смесей.

*Ключевые слова:* нетранзитивность, стохастическое предшествование, копулы, многомерные смеси, предельные теоремы.

DOI: 10.31857/S0005231025040043, EDN: CARPIK

### 1. Введение

В теории и практике различные отношения превосходства между объектами часто обладают свойством транзитивности: если  $A$  превосходит  $B$  и  $B$  превосходит  $C$ , то  $A$  превосходит  $C$ . Однако бывает и по-другому. Например, в классической игре «камень, ножницы, бумага» «камень» побеждает «ножницы», «ножницы» побеждают «бумагу», но при этом «бумага» побеждает «камень».

Различным аспектам и многочисленным примерам нетранзитивности отношений превосходства в природе, технике и обществе посвящен ряд статей А.Н. Поддьякова, в частности [1, 2].

Из недавних зарубежных работ отметим интересные исследования нетранзитивности в биологии, из жизни муравьев [3] и почвенных бактерий [4].

Далее будем говорить о нетранзитивности отношения стохастического предшествования (stochastic precedence) между случайными величинами, а именно, когда одна величина с большей вероятностью меньше другой, чем больше. В [5, 6] это отношение применялось в задачах статистического анализа. Методам сравнения вероятностных распределений на основе статистических данных (гистограмм) посвящена недавняя работа [7], где рассматривались различные стохастические порядки.

К первому упоминанию проблемы нетранзитивности с вероятностной точки зрения можно отнести работу С. Трыбулы совместно с Г. Штейнгаузом,

где в качестве приложения была рассмотрена прочность материалов [8, 9]. Пусть в лаборатории сравнивают попарно на прочность железные бруски с трех разных заводов. Тогда теоретически может сложиться ситуация, что бруски с первого завода «хуже» брусков со второго завода (чаще оказываются менее прочными), бруски со второго «хуже» брусков с третьего, а бруски с третьего «хуже» брусков с первого.

В недавней работе [10] подобное явление было реально обнаружено в медицинской статистике при сравнении продолжительности жизни в некоторых группах людей.

Тема нетранзитивности приобрела популярность на примере *нетранзитивных костей* (nontransitive dice). Имеются в виду наборы игральных костей, на грани которых нанесены числа таким образом, чтобы создать нетранзитивные отношения соответствующих случайных величин. Нетранзитивные кости были популяризованы М. Гарднером [11, 12], им посвящена обширная литература.

Автор настоящей работы продолжает начатые в [13–17] исследования нетранзитивности (в этих работах можно ознакомиться более подробно с проблематикой, историей вопроса, литературой и приложениями). При этом, если ранее изучался случай независимых случайных величин, теперь речь о зависимых случайных величинах.

Классическим примером в этом смысле может служить знаменитый парадокс Кондорсе при голосовании. Изложим его в простейшей формулировке. Пусть имеются три кандидата на некоторый пост и три избирателя, которые имеют предпочтения кандидатов (в баллах) в соответствии со следующими вектор-строками:

$$(1) \quad \begin{array}{l} (1, 2, 3) \\ (3, 1, 2) \\ (2, 3, 1). \end{array}$$

Тогда при выборе между первым и вторым кандидатом побеждает второй. Действительно, у первого и третьего избирателей оценка второго кандидата выше, чем первого ( $1 < 2$  и  $2 < 3$  соответственно), а у второго избирателя ниже ( $3 > 1$ ), так что второй кандидат побеждает двумя голосами из трех. Аналогично, при выборе между вторым и третьим кандидатом побеждает третий, но между первым и третьим – первый.

Подобные ситуации часто возникают при экспертном оценивании, например, качества товаров. При этом трудности связаны с тем, что многие характеристики не допускают прямого количественного измерения (техническими средствами) и оцениваются интуитивно в условных единицах (баллах). В случае продуктов питания важную роль играют органолептические показатели: вкус, цвет, запах и т.п. Необходимо исключать нетранзитивные подмножества в экспертных измерениях, для этого используются различные методы [18–20].



Считается, что увеличение числа экспертов (опрашиваемых) решает проблему, но порой это не так. Например, при массовом опросе о выборе проекта освоения территории (зеленая зона, парк семейного отдыха, деловой центр) была выявлена нетранзитивность в ответах [21]. Предполагается, что люди оценивают проекты с учетом определенных факторов (экологических, социальных, финансовых), но веса (значимости) этих факторов для разных людей различны. Если совместное распределение весов подобно тем, что изучаются в настоящей статье, это могло бы объяснить результаты. При этом более подробный анализ данных компьютерными методами статистики (кластерный, факторный анализ и т.п.) мог бы показать деление людей на группы по взглядам, даже если эти группы никак формально не позиционируются и не осознаются участниками.

Отметим также работы [22, 23].

Свойства методов принятия решений в многокритериальных задачах индивидуального выбора разносторонне изучались в [24]. В условиях нетранзитивности при сравнении какого-то текущего варианта выбора с другими всегда может найтись «лучший» вариант, что при формальном отношении ведет к бесконечному марковскому блужданию [25]. При современном развитии искусственного интеллекта, автоматических систем принятия решения, роботов необходимо учитывать феномен нетранзитивности. В подобной ситуации человек может осознать происходящее и принять волевое решение, а машина может заикнуться.

## 2. Определения и известные результаты

Дадим более строгие определения.

*Определение 1. Отношение  $\prec$  называется нетранзитивным, если для каких-либо объектов  $A, B, C$  из соотношений  $A \prec B, B \prec C$  НЕ следует  $A \prec C$ , а напротив, может быть  $C \prec A$ .*

*Определение 2. Набор объектов  $A, B, C$  называется нетранзитивным, если на нем реализуется нетранзитивность отношения превосходства, т.е. выполнено:  $A \prec B \prec C \prec A$  (либо в обратном порядке).*

*Определение 3.  $X$  стохастически предшествует  $Y$  ( $X \prec Y$ ), если*

$$\mathbf{E} \operatorname{sign}(Y - X) > 0$$

*или*

$$\mathbf{P}(X < Y) > \mathbf{P}(X > Y).$$

Если  $\mathbf{P}(X = Y) = 0$  (например, случайные величины независимы и непрерывны или с непересекающимися областями значений), то  $X \prec Y$  эквивалентно

$$\mathbf{P}(X < Y) > \frac{1}{2}.$$

Заметим, что если случайные величины независимы и непрерывны, а кроме того, одинаково распределены, то всегда верно  $\mathbf{P}(X < Y) = 1/2$ . Для непрерывных случайных величин это равенство может нарушаться как от разной распределенности, так и от зависимости, эти случаи требуют отдельного рассмотрения.

Итак, пусть заданы случайные величины  $X_1$ ,  $X_2$  и  $X_3$  такие, что

$$\mathbf{P}(X_1 < X_2) > \frac{1}{2}, \quad \mathbf{P}(X_2 < X_3) > \frac{1}{2}, \quad \mathbf{P}(X_3 < X_1) > \frac{1}{2},$$

тогда  $X_1 \prec X_2$ ,  $X_2 \prec X_3$ , но  $X_3 \prec X_1$ . Таким образом, ситуация нетранзитивности возникает, когда

$$P_{X_1 X_2 X_3} = \min\{\mathbf{P}(X_1 < X_2), \mathbf{P}(X_2 < X_3), \mathbf{P}(X_3 < X_1)\} > \frac{1}{2},$$

причем интерес представляет не только сам факт, но и сила нетранзитивности, мерой которой может служить величина  $P_{X_1 X_2 X_3}$ .

В [9] показано, что для независимых случайных величин

$$(2) \quad \max_{X_1, X_2, X_3} P_{X_1 X_2 X_3} = \frac{\sqrt{5} - 1}{2} \approx 0,618.$$

В [26, 27] изучалась комбинация из  $n \geq 3$  независимых случайных величин, впоследствии в [28] было показано (и позднее в [29] передоказано геометрическим методом), что максимум вероятностей

$$P_{X_1 \dots X_n} = \min\{\mathbf{P}(X_1 < X_2), \dots, \mathbf{P}(X_{n-1} < X_n), \mathbf{P}(X_n < X_1)\}$$

составляет следующую величину:

$$\max_{X_1, \dots, X_n} P_{X_1 \dots X_n} = 1 - \left(4 \cos^2 \frac{\pi}{n+2}\right)^{-1}, \quad n \geq 3,$$

что совпадает с (2) при  $n = 3$ , поскольку

$$\cos \frac{\pi}{5} = \frac{1 + \sqrt{5}}{4}.$$

Для зависимых величин в [27] было доказано:

$$\max_{X_1, \dots, X_n} P_{X_1 \dots X_n} = \frac{n-1}{n}, \quad n \geq 3,$$

в частности, при  $n = 3$  получаем  $P_{X_1 X_2 X_3} = 2/3$ .

Применительно к трем зависимым случайным величинам игровая интерпретация может быть следующей: разыгрывается случайный вектор  $(X_1, X_2, X_3)$ , перед этим два игрока последовательно выбирают каждый свой

номер компоненты вектора. С чьим номером компонента окажется больше, тот и выиграл. Нетранзитивность означает, что какой бы номер ни выбрал первый игрок, второй игрок может выбрать себе номер так, чтобы обеспечить вероятность выигрыша  $P_{X_1 X_2 X_3} > 1/2$ , т.е. получает преимущество.

Рассмотрим вероятностно-статистическую модификацию парадокса Кондорсе, соответствующую сколь угодно большому числу избирателей, каждый из которых может иметь одно из трех мнений о кандидатах, описываемых (1), равномерно.

Пусть набор зависимых случайных величин  $(X, Y, Z)$  принимает значения из (1) с вероятностями по  $1/3$ . Тогда

$$\mathbf{P}(X < Y) = \mathbf{P}(Y < Z) = \mathbf{P}(Z < X) = \frac{2}{3}.$$

При этом каждая из величин равномерно распределена на множестве  $\{1, 2, 3\}$ .

Возникает вопрос, как перейти от дискретного распределения к непрерывному. Далее рассмотрим различные примеры.

Напомним также понятие копулы [30].

*Определение 4. Копулой ( $m$ -мерной)  $C$  называется функция многомерного распределения на  $[0, 1]^m$  с равномерными частными (маргинальными) распределениями.*

*Копулой распределения  $F$  в  $R^m$  называется такая копула  $C$ , которая удовлетворяет выражению*

$$F(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m)),$$

где  $F_1, \dots, F_m$  – частные функции распределения.

Такое представление существует по теореме Склера и единственно в случае непрерывных частных распределений. Далее будем предполагать непрерывность.

Если заданы  $X_1, \dots, X_m$  и  $U_i = F_i(X_i)$ ,  $1 \leq i \leq n$ , то все  $U_i$  равномерно распределены на  $[0, 1]$ , с совместным распределением  $C$ , т.е.

$$\mathbf{P}(U_1 \leq u_1, \dots, U_m \leq u_m) = C(u_1, \dots, u_m).$$

В непрерывном случае, если  $X_i$  одинаково распределены, то

$$\mathbf{P}(X_i < X_j) = \mathbf{P}(U_i < U_j), \quad i \neq j.$$

### 3. Трехмерные непрерывные копулы

Рассмотрим сначала некоторые естественные обобщения примера Кондорсе с равномерными частными распределениями на  $[0, 1]$  (полученные трехмерные распределения могут использоваться в качестве копул).

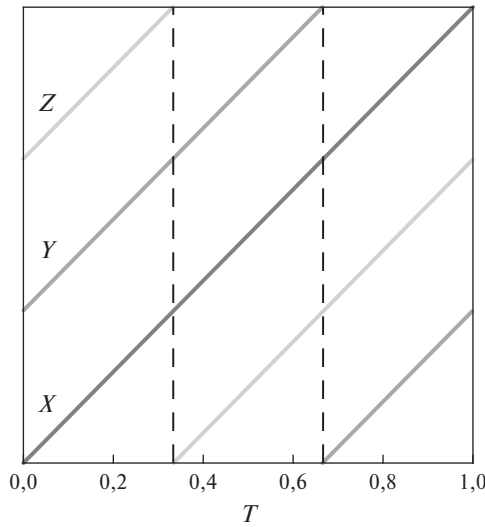


Рис. 1.  $X$ ,  $Y$  и  $Z$  в зависимости от  $T$  в примере 1.

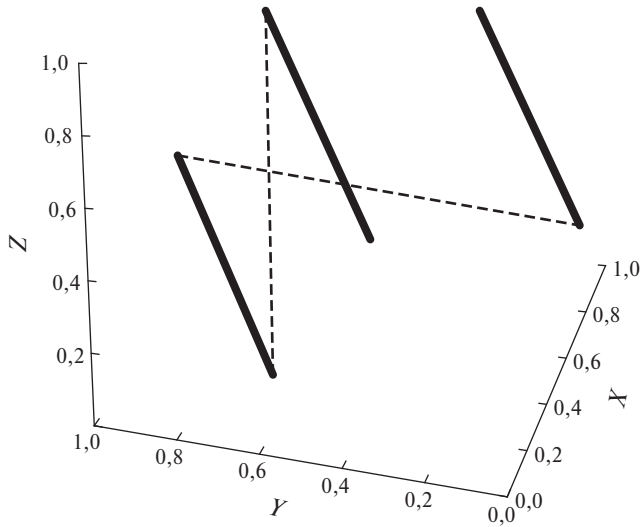


Рис. 2. Распределение в единичном кубе в примере 1.

*Пример 1.* Пусть

$$X = T, \quad Y = \left\{ T + \frac{1}{3} \right\}, \quad Z = \left\{ T + \frac{2}{3} \right\},$$

где  $\{.\}$  – дробная часть числа,  $T$  равномерно распределена на  $[0, 1]$ , тогда  $X$ ,  $Y$ ,  $Z$  равномерно распределены на  $[0, 1]$  и

$$(3) \quad \mathbf{P}(X < Y) = \mathbf{P}(Y < Z) = \mathbf{P}(Z < X) = \frac{2}{3}.$$

Графики  $X$ ,  $Y$  и  $Z$  в зависимости от  $T$  представлены на рис. 1. Вероятности при этом соответствуют долям отрезка  $[0, 1]$ , на которых выполняются соответствующие неравенства.

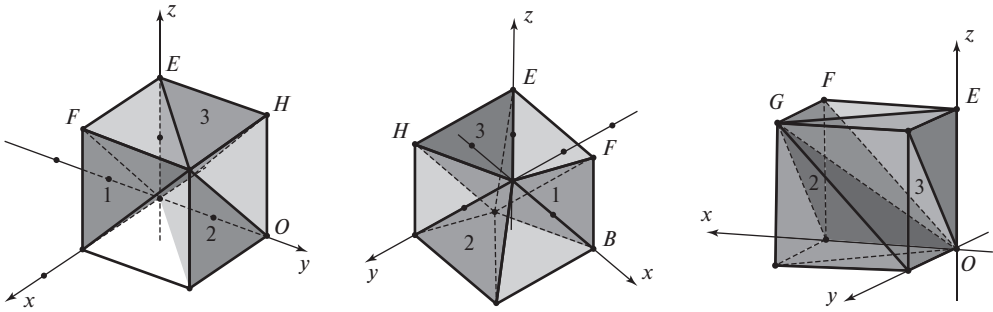


Рис. 3. Распределение в единичном кубе в примере 2.

В этом случае распределение вектора оказывается непрерывным, но не абсолютно непрерывным (имеющим плотность), а сингулярным (сосредоточенным на многообразии меньшей размерности – трех отрезках) в единичном кубе (см. рис. 2).

Перейдем к рассмотрению абсолютно непрерывных распределений. Обозначим плотность совместного распределения случайных величин  $X, Y, Z$  через  $p(x, y, z)$ , где  $x, y, z \in [0, 1]$ .

*Пример 2.* Пусть  $p(x, y, z) = 2$ , если  $x < y < z$  или  $z < x < y$  или  $y < z < x$ , и  $p(x, y, z) = 0$  иначе. Тогда также  $X, Y, Z$  равномерно распределены на  $[0, 1]$  и верно (3). Указанную плотность можно представить в виде:

$$(4) \quad p(x, y, z) = 1 - \text{sign}\{(y - x)(z - y)(x - z)\}.$$

Данное распределение сосредоточено в трех пирамидах внутри единичного куба (см. рис. 3).

В этом случае трехмерное распределение имеет плотность, но эта плотность разрывна. Однако вид (4) наводит на мысль перейти к полиномиальной плотности.

*Пример 3.* Пусть

$$(5) \quad p(x, y, z) = 1 - K(y - x)(z - y)(x - z).$$

Можно проверить, что в этом случае частные распределения остаются равномерными на  $[0, 1]$ . Вопрос заключается в выборе константы  $K$  так, чтобы плотность оставалась неотрицательной.

Имеем

$$\max_{[0,1]^3} \{(y - x)(z - y)(x - z)\} = \frac{1}{4},$$

причем этот максимум достигается в точках  $(1/2, 1, 0)$ ,  $(0, 1/2, 1)$ ,  $(1, 0, 1/2)$ .

Отсюда максимально допустимое  $K = 4$ , положим

$$p(x, y, z) = 1 - 4(y - x)(z - y)(x - z),$$

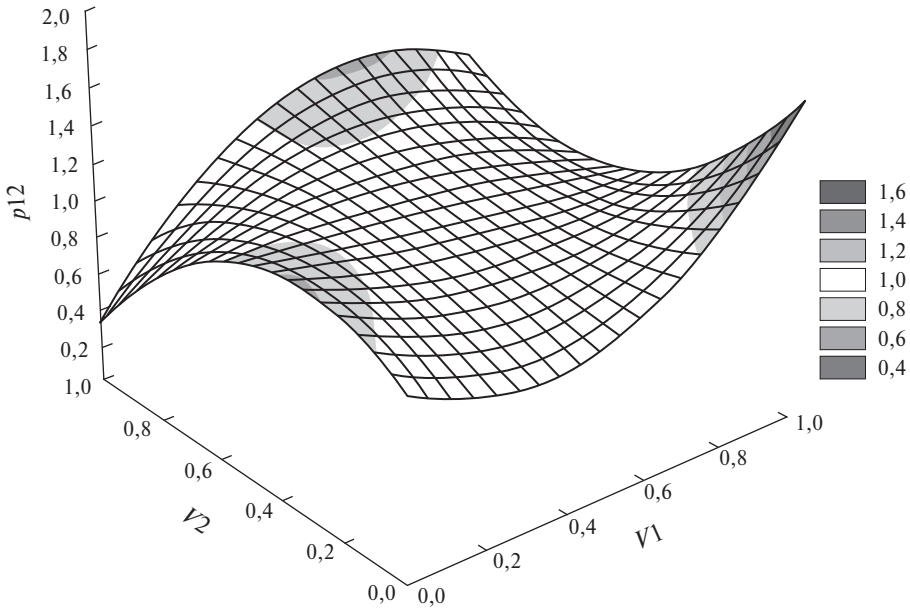


Рис. 4. Двумерная плотность распределения в примере 3.

откуда интегрированием получаем совместную функцию распределения

$$F(x, y, z) = xyz \left( 1 + \frac{2}{3}(xy(y - x) + yz(z - y) + xz(x - z)) \right).$$

Совместная плотность  $X, Y$  принимает вид (см. рис. 4)

$$p_{12}(x, y) = 1 + \frac{2}{3}(y - x)(2 - 3x - 3y + 6xy),$$

откуда получаем

$$\mathbf{P}(X < Y) = \mathbf{P}(Y < Z) = \mathbf{P}(Z < X) = \frac{47}{90} = 0,522\dots$$

*Пример 4.* Рассмотрим плотность, промежуточную между (4) и (5):

$$(6) \quad p(x, y, z) = 1 - \sqrt[k]{4(y - x)(z - y)(x - z)},$$

где  $k$  – нечетное натуральное число,  $k \geq 3$ . Понятно, что при  $k \rightarrow \infty$  будет сходимость к ситуации в примере 2. Оценивая вероятность  $P_{XYZ}$  методом Монте-Карло по  $10^6$  точек, получили следующую таблицу:

| $k$ | $P_{XYZ}$ |
|-----|-----------|
| 3   | 0,569     |
| 5   | 0,595     |
| 7   | 0,610     |
| 9   | 0,621     |
| 11  | 0,628     |

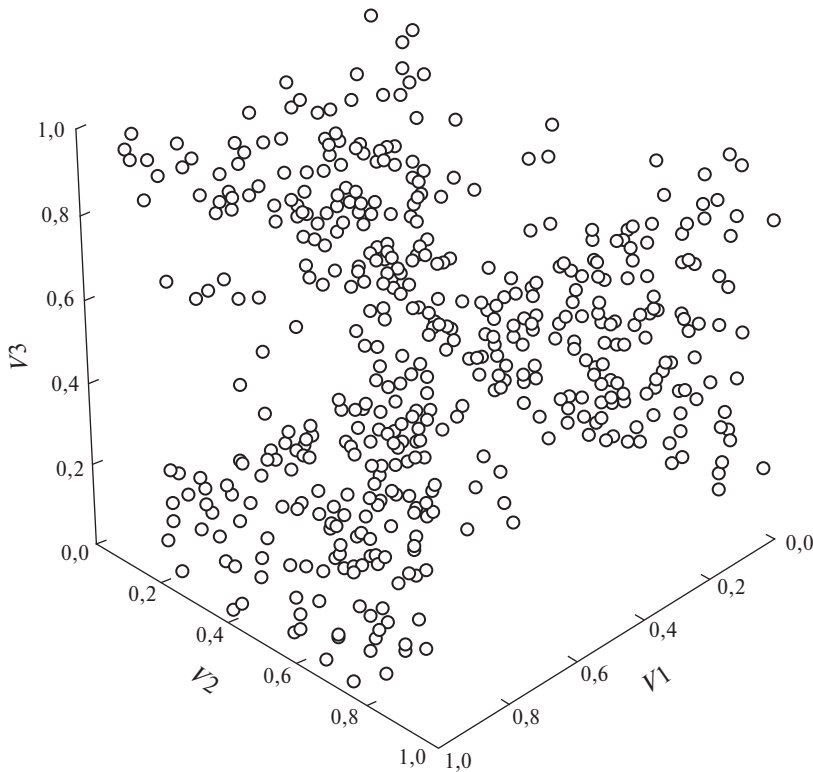


Рис. 5. Результат моделирования в примере 4.

На рис. 5 представлен результат моделирования при  $k = 11$  для  $10^3$  точек. Видна сходимость к трем пирамидам из примера 2.

#### 4. Многомерные смеси

Рассмотрим теперь другой подход, с помощью зашумления дискретных данных, как в [31] применительно к нетранзитивным костям Эфрона (независимым случайным величинам).

*Пример 5.* Пусть случайный вектор  $(X_1, X_2, X_3)$  формируется следующим образом: к значениям (1), принимаемым с вероятностями по  $1/3$ , по всем координатам прибавляются независимые нормальные величины с нулевыми средними и дисперсиями  $\sigma^2 > 0$ .

*Утверждение 1.* В примере 5 верно

$$\begin{aligned}
 P_{X_1 X_2 X_3} &= \mathbf{P}(X_1 < X_2) = \mathbf{P}(X_2 < X_3) = \mathbf{P}(X_3 < X_1) = \\
 &= \frac{1}{3} \left( 2\Phi\left(\frac{1}{\sigma\sqrt{2}}\right) + \Phi\left(-\frac{\sqrt{2}}{\sigma}\right) \right) > \frac{1}{2},
 \end{aligned}$$

где  $P_{X_1 X_2 X_3} \rightarrow 2/3$  при  $\sigma \rightarrow 0$  и  $P_{X_1 X_2 X_3} \rightarrow 1/2$  при  $\sigma \rightarrow \infty$ .

Доказательство представлено в Приложении.

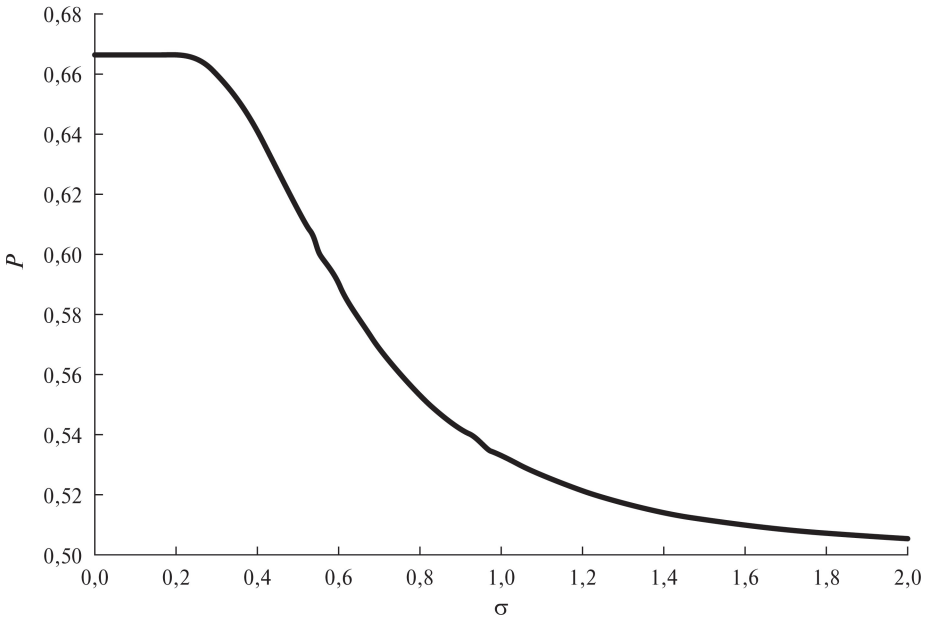


Рис. 6.  $P_{X_1 X_2 X_3}$  в зависимости от  $\sigma$  в примере 5.

Нетранзитивность здесь наблюдается при *любом*  $\sigma^2 > 0$ .

График вероятности  $P_{X_1 X_2 X_3}$  в зависимости от  $\sigma$  представлен на рис. 6.

Зададимся вопросом: при  $\sigma \rightarrow 0$  распределение сходится к дискретному (трехточечному), а что при этом происходит с копулой?

*Утверждение 2. В примере 5 при  $\sigma \rightarrow 0$  распределение копулы сходится к равномерному на объединении кубов*

$$\begin{aligned} & [0, 1/3] \times [1/3, 2/3] \times [2/3, 1], \\ & [2/3, 1] \times [0, 1/3] \times [1/3, 2/3], \\ & [1/3, 2/3] \times [2/3, 1] \times [0, 1/3], \end{aligned}$$

*с плотностью  $p(x_1, x_2, x_3) = 9$  на этом множестве и 0 иначе.*

Доказательство представлено в Приложении.

Введем обозначение для функции стандартного равномерного распределения:

$$(7) \quad R(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1, \end{cases}$$

тогда предельная копула из утверждения 2 представима в виде

$$\begin{aligned} C_0(u_1, u_2, u_3) = & \frac{1}{3} \left( R(3u_1)R(3u_2 - 1)R(3u_3 - 2) + \right. \\ & \left. + R(3u_1 - 2)R(3u_2)R(3u_3 - 1) + R(3u_1 - 1)R(3u_2 - 2)R(3u_3) \right). \end{aligned}$$



Для предельной копулы верно

$$\mathbf{P}(U_1 < U_2) = \mathbf{P}(U_2 < U_3) = \mathbf{P}(U_3 < U_1) = 2/3.$$

Перейдем теперь к более общему случаю.

Напомним, что (дискретная) *многомерная смесь* – это распределение случайного вектора, который с некоторыми вероятностями (весами) принимает значения случайных векторов с заданными распределениями (причем выбор вектора делается независимо от значений).

Если берется смесь гауссовских векторов, то ее называют *гауссовской* смесью и, таким образом, пример 5 был примером гауссовской смеси.

Пусть случайный вектор  $(X_1, X_2, X_3)$  имеет распределение смеси с весами  $p_i$ ,  $i = 1, 2, 3$ , случайных векторов вида

$$(a_{i1}, a_{i2}, a_{i3}) + \lambda(\eta_{i1}, \eta_{i2}, \eta_{i3}),$$

где  $A = (a_{ij})_{i,j=1}^n$  – некоторая матрица,  $(\eta_{i1}, \eta_{i2}, \eta_{i3})$ ,  $i = 1, 2, 3$ , – случайные векторы,  $\lambda > 0$ .

Пусть  $0 < p_i < 1/2$ ,  $i = 1, 2, 3$ .

*Теорема 1.* Если  $\eta_{i1} - \eta_{i2}$ ,  $\eta_{i2} - \eta_{i3}$ ,  $\eta_{i3} - \eta_{i1}$  непрерывны при всех  $i = 1, 2, 3$  и  $a_{ik}$  по строкам  $A$  идут в том же порядке, что в (1), тогда существует  $\lambda_1 > 0$  такое, что при  $\lambda < \lambda_1$  имеет место нетранзитивность, т.е.  $P_{X_1 X_2 X_3} > 1/2$ , и

$$\lim_{\lambda \rightarrow 0} P_{X_1 X_2 X_3} = 1 - \max\{p_1, p_2, p_3\}.$$

Доказательство представлено в Приложении.

*Теорема 2.* Если  $\eta_{ij} = \eta_j$ , где  $\eta_j$  независимы и имеют одинаковое непрерывное распределение,  $i, j = 1, 2, 3$ , и  $a_{ik}$  по столбцам  $A$  идут в том же порядке, что в (1), то существует  $\lambda_2 > 0$  такое, что при  $\lambda < \lambda_2$  имеет место нетранзитивность  $(U_1, U_2, U_3)$ , т.е.  $P_{U_1 U_2 U_3} > 1/2$ , и при  $\lambda \rightarrow 0$  распределение копулы сходится к смеси равномерных распределений с весами  $p_i$ ,  $i = 1, 2, 3$  на параллелепипедах

$$\begin{aligned} & [0, p_1] \times [p_1, p_1 + p_2] \times [p_1 + p_2, 1], \\ & [p_1 + p_2, 1] \times [0, p_1] \times [p_1, p_1 + p_2], \\ & [p_1, p_1 + p_2] \times [p_1 + p_2, 1] \times [0, p_1], \end{aligned}$$

так что

$$\lim_{\lambda \rightarrow 0} P_{U_1 U_2 U_3} = 1 - \max\{p_1, p_2, p_3\}.$$

Доказательство представлено в Приложении.

В этом случае предельную копулу, с учетом обозначения (7), можно представить в виде

$$\begin{aligned}
 C_0(u_1, u_2, u_3) &= \\
 &= p_1 R\left(\frac{u_1}{p_1}\right) R\left(\frac{u_2 - p_1}{p_2}\right) R\left(\frac{u_3 - (p_1 + p_2)}{p_3}\right) + \\
 &+ p_2 R\left(\frac{u_1 - (p_1 + p_2)}{p_3}\right) R\left(\frac{u_2}{p_1}\right) R\left(\frac{u_3 - p_1}{p_2}\right) + \\
 &+ p_3 R\left(\frac{u_1 - p_1}{p_2}\right) R\left(\frac{u_2 - (p_1 + p_2)}{p_3}\right) R\left(\frac{u_3}{p_1}\right).
 \end{aligned}$$

## 5. Заключение

В работе получены новые результаты в изучении нетранзитивности отношения стохастического предшествования в теории вероятностей. Исходя из парадокса Кондорсе, построены примеры трехмерных непрерывных распределений и копул, порождающих нетранзитивные наборы зависимых случайных величин. Рассмотрены случаи сингулярных распределений, распределений с разрывной плотностью, полиномиальной плотностью, а также случай многомерных смесей. Доказаны предельные теоремы для многомерных смесей. Подобные многомерные непрерывные распределения могут описывать и объяснять проявления нетранзитивности в психологии, экономике, биологии и др. Этот феномен должен учитываться при построении систем автоматического принятия решений, искусственного интеллекта, роботов и т.п.

## ПРИЛОЖЕНИЕ

*Доказательство утверждения 1.* Используем представление

$$X_i = X_i^0 + \varepsilon_i, \quad i = 1, 2, 3,$$

где вектор  $(X_1^0, X_2^0, X_3^0)$  принимает значения (1) равновероятно и случайные величины  $\varepsilon_i \sim N(0, \sigma^2)$  независимы.

Найдем вероятность  $\mathbf{P}(X_1 < X_2)$ , остальные находятся аналогично. Имеем

$$\mathbf{P}(X_1 < X_2) = \mathbf{P}(X_1^0 + \varepsilon_1 < X_2^0 + \varepsilon_2) = \mathbf{P}(\varepsilon_1 - \varepsilon_2 < X_2^0 - X_1^0).$$

Обозначим  $\varepsilon = \varepsilon_1 - \varepsilon_2$ , тогда  $\varepsilon \sim N(0, 2\sigma^2)$ . Получаем

$$\begin{aligned}
 \mathbf{P}(X_1 < X_2) &= \frac{1}{3}\mathbf{P}(\varepsilon < 2 - 1) + \frac{1}{3}\mathbf{P}(\varepsilon < 1 - 3) + \frac{1}{3}\mathbf{P}(\varepsilon < 3 - 2) = \\
 &= \frac{2}{3}\mathbf{P}(\varepsilon < 1) + \frac{1}{3}\mathbf{P}(\varepsilon < -2) = \frac{1}{3}\left(2\Phi\left(\frac{1}{\sigma\sqrt{2}}\right) + \Phi\left(-\frac{\sqrt{2}}{\sigma}\right)\right).
 \end{aligned}$$

*Доказательство утверждения 2.* Обозначим через  $\nu$  номер вектора из (1), выбранный в качестве значения  $(X_1^0, X_2^0, X_3^0)$ . При условии  $\nu = 1$  имеем

$$\begin{aligned} U_1 &= \frac{1}{3} \left( \Phi \left( \frac{1 + \varepsilon_1 - 1}{\sigma} \right) + \Phi \left( \frac{1 + \varepsilon_1 - 2}{\sigma} \right) + \Phi \left( \frac{1 + \varepsilon_1 - 3}{\sigma} \right) \right), \\ U_2 &= \frac{1}{3} \left( \Phi \left( \frac{2 + \varepsilon_2 - 1}{\sigma} \right) + \Phi \left( \frac{2 + \varepsilon_2 - 2}{\sigma} \right) + \Phi \left( \frac{2 + \varepsilon_2 - 3}{\sigma} \right) \right), \\ U_3 &= \frac{1}{3} \left( \Phi \left( \frac{3 + \varepsilon_3 - 1}{\sigma} \right) + \Phi \left( \frac{3 + \varepsilon_3 - 2}{\sigma} \right) + \Phi \left( \frac{3 + \varepsilon_3 - 3}{\sigma} \right) \right). \end{aligned}$$

Заметим, что  $\varepsilon_i/\sigma \sim N(0, 1)$ . Определим случайные величины  $U_i^0 = \Phi(\varepsilon_i/\sigma)$ , тогда они имеют равномерное распределение на  $[0, 1]$  и независимы. При этом

$$\Phi \left( \frac{\varepsilon_i + c}{\sigma} \right) \xrightarrow{P} \begin{cases} 0, & c < 0 \\ U_i^0, & c = 0, \\ 1, & c > 0, \end{cases} \quad \sigma \rightarrow 0.$$

Отсюда

$$(U_1, U_2, U_3) \xrightarrow{P} \left( \frac{1}{3}U_1^0, \frac{1}{3} + \frac{1}{3}U_2^0, \frac{2}{3} + \frac{1}{3}U_3^0 \right), \quad \sigma \rightarrow 0,$$

т.е. в пределе получаем равномерное распределение на кубе  $[0, 1/3] \times [1/3, 2/3] \times [2/3, 1]$ . Аналогично рассматриваем случаи  $\nu = 2, 3$ .

*Доказательство теоремы 1.* Используем представление

$$(X_1, X_2, X_3) = (a_{\nu 1}, a_{\nu 2}, a_{\nu 3}) + \lambda(\eta_{\nu 1}, \eta_{\nu 2}, \eta_{\nu 3}),$$

где  $\nu$  принимает значения 1, 2, 3 с вероятностями  $p_1, p_2, p_3$  независимо от  $\eta_{ij}$ ,  $i, j = 1, 2, 3$ .

Обозначим распределения  $\eta_{i1} - \eta_{i2}$  через  $G_i$ ,  $i = 1, 2, 3$ .

Получаем

$$\mathbf{P}(X_1 < X_2) = \sum_{i=1}^3 p_i G_i \left( \frac{a_{i2} - a_{i1}}{\lambda} \right) \rightarrow p_1 + p_3 = 1 - p_2 > \frac{1}{2}, \quad \lambda \rightarrow 0,$$

поскольку  $a_{12} > a_{11}$ ,  $a_{22} < a_{21}$ ,  $a_{32} > a_{31}$ . Аналогично,

$$\mathbf{P}(X_2 < X_3) \rightarrow 1 - p_3, \quad \mathbf{P}(X_3 < X_1) \rightarrow 1 - p_1, \quad \sigma \rightarrow 0$$

и

$$P_{X_1 X_2 X_3} \rightarrow \min\{1 - p_1, 1 - p_2, 1 - p_3\} = 1 - \max\{p_1, p_2, p_3\} > \frac{1}{2}, \quad \sigma \rightarrow 0.$$

*Доказательство теоремы 2.* Будем доказывать аналогично утверждению 2. Используем представление

$$(X_1, X_2, X_3) = (a_{\nu 1}, a_{\nu 2}, a_{\nu 3}) + \lambda(\eta_1, \eta_2, \eta_3),$$

где случайные величины  $\eta_j$ ,  $j = 1, 2, 3$  непрерывны, независимы и одинаково распределены, обозначим их распределение через  $G$ .

При условии  $\nu = 1$  имеем

$$U_j = \sum_{i=1}^3 p_i G \left( \frac{a_{1j} - a_{ij}}{\lambda} + \eta_j \right), \quad j = 1, 2, 3,$$

откуда

$$(U_1, U_2, U_3) \xrightarrow{d} (p_1 U_0^1, p_1 + p_2 U_2^0, p_1 + p_2 + p_3 U_3^0), \quad \lambda \rightarrow 0,$$

т.е. в пределе получаем равномерное распределение на параллелепипеде  $[0, p_1] \times [p_1, p_1 + p_2] \times [p_1 + p_2, 1]$ . Аналогично рассматриваем случаи  $\nu = 2, 3$ .

#### СПИСОК ЛИТЕРАТУРЫ

1. *Poddiakov A., Valsiner J.* Intransitivity cycles and their transformations: How dynamically adapting systems function? / *Qualitative Mathematics for the Social Sciences: Mathematical Models for Research on Cultural Dynamics*. Abingdon, NY: Routledge, 2013. P. 343–391.
2. *Поддьяков А.Н.* Принцип нетранзитивности превосходства в разных парадигмах // *Вопросы психологии*. 2019. № 2. С. 3–16.
3. *Vandermeer J., Perfecto I.* Intransitivity as a dynamic assembly engine of competitive communities // *PNAS*. 2023. V. 120. No. 15. Art. e2217372120.
4. *Verdu M., Alcantara J.M., Navarro-Cano J.A., et al.* Transitivity and intransitivity in soil bacterial networks // *ISME J*. 2023. V. 17. P. 2135–2139.
5. *Arcones M.A., Kvam P.H., Samaniego F.J.* Nonparametric estimation of a distribution subject to a stochastic precedence constraint // *J. Amer. Stat. Assoc.* 2002. V. 97. No. 457. P. 170–182.
6. *Boland P.J., Singh H., Cukic B.* The stochastic precedence ordering with applications in sampling and testing // *J. Appl. Probab.* 2004. V. 41. No. 1. P. 73–82.
7. *Лепский А.Е.* Стохастическое и нечеткое упорядочивание методом минимальных преобразований // *АиТ*. 2017. № 1. С. 59–79.  
*Lepskiy A.E.* Stochastic and fuzzy ordering with the method of minimal transformations // *Autom. Remote Control*. 2017. V. 78. No. 1. P. 50–66.
8. *Steinhaus H., Trybula S.* On a paradox in applied probabilities // *Bull. de l'Acad. Polon. des Sci.* 1959. V. 7. P. 67–69.
9. *Trybula S.* On the paradox of three random variables // *Zastos. Matem.* 1961. V. 5. No. 4. P. 321–332.
10. *Demler O.V., Demler I.A.* Non-transitivity of the Win Ratio and the Area Under the Receiver Operating Characteristics Curve (AUC): a case for evaluating the strength of stochastic comparisons. Available at: <https://arxiv.org/abs/2309.01791> (accessed November 10, 2023).

11. *Gardner M.* The paradox of the nontransitive dice and the elusive principle of indifference // *Sci. Amer.* 1970. V. 223. No. 6. P. 110–114.
12. *Gardner M.* On the paradoxical situations that arise from nontransitive relations // *Sci. Amer.* 1974. V. 231. No. 6. P. 120–125.
13. *Лебедев А.В.* Проблема нетранзитивности для трех непрерывных случайных величин // *АиТ.* 2019. № 6. С. 91–103.  
*Lebedev A.V.* The nontransitivity problem for three continuous random variables // *Autom. Remote Control.* 2019. V. 80. No. 6. P. 1058–1068.
14. *Лебедев А.В.* Нетранзитивные триплеты непрерывных случайных величин и их приложения // *Информатика и ее применения.* 2019. Т. 13. № 3. С. 20–26.
15. *Горбунова А.В., Лебедев А.В.* Эффекты стохастической нетранзитивности в системах массового обслуживания // *Управление большими системами.* 2020. Т. 85. С. 23–50.
16. *Gorbunova A.V., Lebedev A.V.* Nontransitivity of tuples of random variables with polynomial density and its effects in Bayesian models // *Math. Comput. Simulat.* 2022. V. 202. P. 181–192.
17. *Poddiakov A., Lebedev A.V.* Intransitivity and meta-intransitivity: meta-dice, levers and other opportunities. // *Eur. J. Math.* 2023. V. 9. Art. N 27. P. 1–17.
18. *Хамханова Д.Н., Шарипова С.М.* Исключение нетранзитивных подмножеств из результатов определения весовых коэффициентов показателей качества хлебо-булочных изделий методом ранжирования // *Фундаментальные исследования.* 2013. № 11-2. С. 184–187.
19. *Хамханова Д.Н., Шарипова С.М.* Понятие «нетранзитивные подмножества» в перерабатывающих отраслях агропромышленного комплекса / *Образование и наука. Материалы национальной конференции.* 2019. С. 301–305.
20. *Шарипова С.М.* Исследование нетранзитивных подмножеств в результатах экспертных измерений // *Дисс. ... канд. техн. наук.* ЮЗГУ. Улан-Удэ. 2014.
21. *Зеляк А.А., Шокин Я.В.* Исследование нарушения принципа транзитивности предпочтений в рамках развития поведенческой экономической теории // *РИСК: Ресурсы. Информация. Снабжение. Конкуренция.* 2012. № 4. С. 212–216.
22. *Волкова А.А., Панов С.А., Шокин Я.В.* Анализ феномена нетранзитивности предпочтений потребителей в свете современных нейроэкономических исследований // *Вест. Челяб. ГУ.* 2019. № 9 (431). С. 131–142.
23. *Шокин Я.В., Волкова А.А.* Обзор приложений нетранзитивности потребительских предпочтений в экономике // *Тр. ИСА РАН.* 2022. Т. 72. № 4. С. 62–67.
24. *Ларичев О.И.* Свойства методов принятия решений в многокритериальных задачах индивидуального выбора // *АиТ.* 2002. № 2. С. 146–158.  
*Larichev O.I.* Properties of the Decision Methods in the Multicriteria Problems of Individual Choice // *Autom. Remote Control.* 2002. V. 63. No. 2. P. 304–315.
25. *Зутлер И.А.* Выбор последовательными сравнениями как непрерывное марковское блуждание // *АиТ.* 2011. № 12. С. 60–74.  
*Zutler I.A.* The sequential alternative search as a continuous Markov random walk // *Autom. Remote Control.* 2011. V. 72. No. 12. P. 2477–2491.
26. *Trybula S.* On the paradox of  $n$  random variables // *Zastos. Matem. (Appl. Math.)* 1965. V. 8. No. 2. P. 143–156.
27. *Usiskin Z.* Max–min probabilities in the voting paradox // *Ann. Math. Stat.* 1964. V. 35. No. 2. P. 857–862.

28. *Богданов И.И.* Нетранзитивные рулетки // Матем. просвещение. 2010. Сер. 3. Вып. 14. С. 240–255.
29. *Komisarski A.* Nontransitive random variables and nontransitive dice // Amer. Math. Monthly. 2021. V. 128. No. 5. P. 423–434.
30. *Nelsen R.* An introduction to copulas. NY, USA. Springer, 2006.
31. *Thangalevu K., Brunner E.* Wilcoxon–Mann–Whitney test for stratified samples and Efron’s paradox dice // J. Statist. Plann. Inference. 2007. V. 137. No. 3. P. 720–737.

*Статья представлена к публикации членом редколлегии И.В. Родионовым.*

Поступила в редакцию 13.11.2023

После доработки 21.06.2024

Принята к публикации 09.01.2025

# Интеллектуальные системы управления, анализ данных

© 2025 г. Н.А. СКВОРЦОВ (nskv@mail.ru)  
(Федеральный исследовательский центр  
«Информатика и управление» РАН, Москва)

## УПРАВЛЕНИЕ КАЧЕСТВОМ ДАННЫХ ПРИ РЕШЕНИИ ЗАДАЧ В ИССЛЕДОВАТЕЛЬСКИХ ИНФРАСТРУКТУРАХ НАД НЕОДНОРОДНЫМИ ИСТОЧНИКАМИ ДАННЫХ

Решение задач на основе доступных научных данных, особенно в контексте открытой науки и исследовательских инфраструктур, должно обеспечивать возможность их многократного повторного использования. Показатели качества данных являются важными характеристиками, влияющими не только на точность методов при решении исследовательских задач, но и на оценку пригодности данных, возможность решения конкретных научных задач, выбор методов работы с данными, их совместимость, возможность отождествления объектов и другие аспекты повторного использования. При этом требуется оценка различных показателей качества данных на разных уровнях агрегации – от целых наборов данных до отдельных значений. В данном исследовании представлен подход к комплексному управлению качеством данных на основе их спецификаций, а также требований к качеству данных и метаданных. Обсуждаются различные показатели оценки качества данных, включая точность, полноту и происхождение. Разработанный подход применен на примере решения задач с использованием множественных источников данных в области звездной астрономии.

*Ключевые слова:* качество данных, повторное использование данных, формальные спецификации, нефункциональные требования.

**DOI:** 10.31857/S0005231025040057, **EDN:** CARLWN

### 1. Введение

Научные исследования неизбежно сталкиваются с необходимостью оценки качества доступных данных, которые используются в процессе решения задач. На разных этапах научного исследования возникает потребность в оценке качества данных. При поиске и выборе подходящих наборов данных важно оценить их применимость, основываясь на информации о качестве данных в этих наборах. При создании выборок, включающих только подходящие данные, а также при очистке данных и их улучшении с целью подготовки к исследованию необходимо оценивать качество данных, связанных с конкретными объектами или характеристиками. При использовании научных методов важно учитывать качество исходных данных как для оценки качества

получаемых результатов, так и для оценки качества самих применяемых методов. Таким образом, показатели качества данных являются важным и даже необходимым элементом при решении различных исследовательских задач. Отсутствие информации о качестве данных или ее игнорирование может существенно снизить качество результатов, вплоть до их ошибочности, или сделать проведение исследования невозможным.

Во многих дисциплинах объем научных данных неуклонно растет. В условиях неоднородности источников научных данных и их разнообразия возникают проблемы совместимости данных, полученных разными методами, созданных с различными целями и имеющих разные требования к качеству.

В одних источниках информация о качестве представленных данных может содержаться в описании наборов данных, сопроводительной документации или быть известной из внешних исследований. В других источниках данные сопровождаются оценками качества, включенными в структуру самих наборов. В некоторых случаях информация о качестве данных отсутствует, но его можно оценить с помощью статистических методов, анализируя сами данные или их выборки. Методы оценки качества данных также могут существенно различаться. В зависимости от решаемых задач важными могут быть разные виды показателей качества и различные критерии их оценки. В некоторых случаях необходимо по-разному оценивать качество решений, исходя из информации о качестве исходных данных. Таким образом, неоднородность наблюдается и в самих метаданных, касающихся качества данных, и в их применении.

Для управления научными данными и поддержки исследований создаются исследовательские инфраструктуры, которые аккумулируют данные, предоставляют сервисы и метаданные, обеспечивая их повторное использование. Ввиду своей важности подходы к управлению качеством данных в исследовательских инфраструктурах требуют тщательных исследований. При работе с большими объемами данных и множеством неоднородных источников необходимо стремиться к автоматизированному управлению их качеством. Поэтому метаданные, касающиеся качества данных, должны быть четко определены, доступны и понятны как для человека, так и для машины.

В данной статье ставится задача разработки представления и применения различных показателей качества данных в исследовательских инфраструктурах. В следующем разделе представлен анализ состояния исследований в данной области. В разделе 3 предложена классификация подходов к управлению качеством данных. Описаны спецификации метаданных и принципы их использования в исследовательских инфраструктурах. Раздел 4 содержит пример управления качеством данных при решении задач в звездной астрономии с использованием данных больших многоцветных фотометрических обзоров неба. Сделаны общие выводы о развитии исследовательских инфраструктур для повышения эффективности научных исследований.



## 2. Состояние исследований в области управления качеством данных

Проблемы качества столь чувствительны в национальной и глобальной экономике, что для их решения давно прорабатывались подходы, методы и стандарты. В информатике вопросы качества данных, как часть общей проблемы качества объектов реального мира, играют столь же немаловажную роль. Базы данных и информационные системы рассматриваются как самостоятельные объекты [1], которые, как и другие объекты реального мира, могут иметь допустимое или недопустимое состояние, которое может оцениваться показателями качества данных. При этом данные в них отражают состояние объектов реального мира, и необходимо оценивать как соответствие состояния информационных объектов характеристикам объектов реального мира, так и возможность идентификации объектов реального мира по данным. Игнорирование проблем качества данных может привести к серьезным последствиям не только в информационной сфере, но и негативно сказаться на других сферах человеческой деятельности.

Оценка качества объектов реального мира основывается на данных об их качественных и количественных характеристиках. Таким образом, качество объектов (изделий, материалов и др.) всегда связано с качеством данных о них. Более того, в современной научной парадигме, основанной на извлечении новых знаний из научных данных, сбор большого объема данных разного качества об исследуемых объектах из различных источников требует оценки и учета их качества в исследованиях, что, в конечном счете, определяет развитие науки.

К данным, используемым в человеческой деятельности, могут предъявляться требования. Качество данных (такие характеристики, как точность и достоверность) об объектах реального мира может определять применимость и особенности использования как самих данных, так и объектов исследования. Низкое качество данных, даже если сами исследуемые объекты обладают высоким качеством, может затруднить решение задач и помешать правильному использованию как данных, так и объектов исследования.

С давних пор преобладающим подходом в оценке качества данных стали многокритериальные модели, которые включают наборы показателей качества. Уже в ранних исследованиях, таких как [2], обсуждается важность и востребованность различных показателей качества данных. Среди основных выделяются такие показатели, как точность, полнота, целостность и актуальность, а также достоверность, недвусмысленность, надежность и объем данных. При этом набор используемых показателей зависит от поставленных целей в конкретной предметной области исследования. В некоторых работах подчеркивается связь качества данных с качеством продукции и объектов реального мира, что подтверждает эффективность таких подходов к оценке качества данных. С этой точки зрения ценность данных определяется их потребителем, что обращает внимание на такие показатели качества, как значимость и корректность данных [3].

В процессе исследований устоялся подход, в котором модель качества определяется как пространство показателей качества, рассматриваемых в виде измерений. В [4] методология оценки качества определяется моделью, включающей определенный набор показателей, методы их оценки и, возможно, методы их обобщения. В каждой конкретной модели в зависимости от решаемых задач могут использоваться различные методы оценки показателей качества, такие как алгоритмы, правила, эвристики или модели машинного обучения, обеспечивающие решение задач, относящихся к определенным аспектам качества данных.

Так, например, полнота данных может быть оценена как полнота кортежей (заполнение всех атрибутов), атрибутов (оценка количества отсутствующих значений среди значений атрибута в кортежах) или как покрытие всех существующих объектов данного типа. Измерение точности может определяться как погрешность значений, синтаксическая точность (совпадение используемых имен для одного объекта) или семантическая точность (непротиворечивость фактов). Объем данных может определяться занимаемым объемом памяти или ожидаемым количеством возвращаемых кортежей и т.д. Другими словами, методы и реализации оценки качества могут зависеть как от характеристик наборов данных или выборок, так и от определенных ограничений предметной области, к которой относятся данные. Семантика того или иного показателя, несмотря на одинаковые названия, может существенно различаться в разных задачах. Соответственно требования к качеству данных могут формулироваться в исследовательских задачах как в терминах наборов данных или хранилищ данных, так и в терминах предметной области прикладной задачи.

Похожие подходы определены в действующих стандартах качества данных. С управлением качеством данных связаны международные и соответствующие им российские национальные стандарты, в частности ГОСТ Р 56214-2014 [5] (ISO/TS 8000-1: 2011) и комплекс сопутствующих стандартов. Эти стандарты соответствуют устоявшимся принципам моделирования качества данных и вводят такие показатели качества, как точность, полнота (или завершенность) и происхождение (или источник) данных, при этом не ограничивая конкретные реализации методов оценки качества по этим критериям. Кроме того, стандарты определяют концептуальные схемы управления качеством и принципы разработки критериев для оценки этих показателей, включая принципы однозначного синтаксического и семантического кодирования. Качество данных в рамках данных стандартов указывает на степень, в которой данные удовлетворяют требованиям потребителей и соответствуют установленным критериям. Устанавливаются принципы предъявления требований и повышения качества данных, позволяющие потребителям запрашивать данные надлежащего качества и точно определять, соответствуют ли полученные данные установленным стандартам.

Интересен также стандарт ГОСТ Р 57773-2017 [6] (ISO 19157:2013), который закрепляет практику управления качеством при работе с источниками пространственных данных, имеющими разное разрешение и детализацию. Этот стандарт позволяет выбирать наиболее подходящие по качеству данные для решения задач в требуемых масштабах.

Проблемы доступности научных данных, подходящих для исследований, возникают во многих областях с интенсивным использованием данных. В связи с этим активно развивается направление исследовательских инфраструктур, которые объединяют исследовательские данные, сервисы и инструменты, позволяя многократно их использовать.

Одной из идеологических основ их развития стали руководящие принципы FAIR-данных [7], которые предлагают направление для обеспечения обнаруживаемости (Findability, F), доступности (Accessibility, A), интероперабельности (Interoperability, I) и, как следствие, возможности повторного использования (Reusability, R) исследовательских данных. При разработке исследовательских инфраструктур и принципов их работы вопросы обеспечения качества данных часто сводятся к подходам оценки хранилищ данных с точки зрения соответствия принципам FAIR.

Крупные междисциплинарные инициативы научного сообщества, направленные на обеспечение и управление данными, а также на их анализ, задаются вопросами обеспечения качества данных, в основном, с позиции принципов FAIR, смещая акцент с управления качеством данных на оценку качества инфраструктуры управления данными для возможности улучшения качества данных.

В инициативе RDA разработана модель FAIR Data Maturity Model [8], которая определяет набор показателей, их приоритеты и методы для оценки следования принципам FAIR. Она используется в качестве общего подхода для оценки различных методологий. В частности, проект FAIRsFAIR [9] был создан для популяризации принципов FAIR для данных, создаваемых исследователями. Были разработаны популярные разъяснения, примеры и варианты решений, которые способствуют интероперабельности и повторному использованию данных, а также инструменты, стандарты и практики для обеспечения управления данными в различных научных дисциплинах. На базе модели FAIR Data Maturity Model и разработок FAIRsFAIR реализован инструмент F-UJI [10] для оценки степени соответствия принципам FAIR научных наборов данных и предоставления рекомендаций. Таким образом, подобные разработки создают альтернативный подход к оценке качества данных. Оценка следования принципам FAIR дает осязаемую оценку качества управления данными, а не самих данных, в основном. С другой стороны, это может обесценить сами принципы FAIR как некоторое стратегическое направление в развитии управления данными. Ведь эти принципы остаются перспективными, пока в достаточной мере невыполнима декларируемая ими возможность автономного управления дан-

ными машиной (machine-actionability), т.е. переход от обработки машиной predetermined наборов метаданных и инструкций к обеспечению корректной интерпретации данных и метаданных, с которыми машина ранее не работала.

Инициатива ELIXIR специализируется на биомедицинских и биологических информационных ресурсах. Здесь вопросы качества данных также чаще обсуждаются именно в контексте соответствия принципам FAIR, в частности в рекомендациях FAIR CookBook [11]. Однако специальных рекомендаций об управлении качеством данных не выработано.

Исследовательская инфраструктура ELIXIR [12, 13] (в частности, платформы Data и Interoperability в ее составе) поддерживает хранение больших объемов данных, развитое управление метаданными, контроль целостности данных, механизмы обнаружения дубликатов и данных и другие средства, используемые для управления качеством данных. При этом специализированных средств для спецификации качества данных, помимо оценки выполнения принципов FAIR, не предусмотрено.

В рамках инициативы ESIP, посвященной управлению данными в области наук о Земле, проводились исследования, связанные с управлением качеством данных. Рабочие группы Data Quality Working Group (DQWG) и Information Quality Cluster (IQC) разработали рекомендации применения лучших практик и стандартов в этой области, принципов управления жизненным циклом данных для обеспечения их целостности, качества и повторного использования, а также способствовали внедрению этих рекомендаций среди провайдеров данных, разработчиков программных систем и научных групп [14–16]. Эти решения во многом связаны со сложившейся практикой работы с пространственными данными, такими как стандартизация атрибутов и метаданных, использование флагов качества в данных, следование специализированным стандартам.

Также в рамках ESIP разработана матрица оценки зрелости научных данных, объединяющая в себе основные принципы курирования данных при долгосрочном хранении и эксплуатации. Частью этих принципов является управление качеством данных [17].

Международные стандарты управления качеством данных также продолжают развиваться. Комплекс стандартов управления качеством пространственных данных находится в процессе обновления [18].

Большинство распространенных продуктов, связанных с управлением качеством данных, является инструментами очистки данных в составе систем интеграции. Задачами, решаемыми в таких продуктах, являются выявление ошибок и дубликатов данных, приведение типов и стандартизация представления значений, решение проблемы отсутствующих данных и др. Однако такие продукты также могут предоставлять средства оценки качества данных и варианты алгоритмов вычисления показателей качества. К таким продуктам относится, в частности, Talend Data Quality [19], являющийся частью систе-

мы интеграции реляционных баз данных. В этом инструменте используются бизнес-правила на основе SQL-запросов для представления требований к данным и мониторинга качества данных по таким показателям, как полнота, точность и согласованность данных с возможностью выявления конкретных проблемных областей. IBM InfoSphere Quality Stage [20], входящий в линейку продуктов IBM для поддержки процессов управления данными, обеспечивает непрерывный мониторинг событий качества в потоках данных, исправление ошибок, очистку данных, обогащение метаданных.

В основе Semantic Data Quality Management (SDQM) [21] лежит анализ стандартов ISO 9000 и исследования моделей качества данных с наборами показателей, а также технологии семантического Веба [22], включая RDF-технологии [23]. Семантические подходы на основе RDF, языка с расширяемой семантикой для описания ресурсов, позволяют связывать с ресурсами, идентифицируемыми в глобальном информационном пространстве, метаданные, семантика которых определяется словарями или онтологиями в различных пространствах имен, задавать требования к ресурсам в терминах этих словарей и находить релевантные ресурсы. Более того, модель RDF стала наиболее часто используемой для определения схем метаданных. Для определения модели качества вводится словарь управления качеством данных [24], повторно используются модель происхождения данных для веб-ресурсов [25] и другие словари.

Big Data Quality Management Framework (BDQMF) [26] предназначен для управления качеством больших данных на всех этапах их жизненного цикла. BDQMF охватывает весь процесс управления качеством данных: от определения требований и их преобразования до анализа данных и мониторинга качества. Основные подходы включают создание профиля качества данных с целевыми показателями качества, оценку и улучшение качества на каждом этапе, валидацию и оптимизацию правил качества для повышения точности данных, мониторинг и визуализацию данных после обработки.

С точки зрения принципов FAIR проблемы качества данных должны рассматриваться в контексте обеспечения возможности повторного использования данных (R) [27]. Основной из принципов (R1) в этом направлении указывает на необходимость обильно описывать данные точными и актуальными сведениями (атрибутами). Под такими сведениями подразумеваются ресурсы и метаданные, необходимые для оценки возможности повторного использования данных, а не просто описание их семантики. Эти сведения могут включать информацию о происхождении данных, лицензионные условия их использования и другие нефункциональные свойства, которые тем не менее играют важную роль при выборе данных исследователями для решения своих научных задач [28]. В эту категорию также следует включать метаданные, основанные на показателях качества данных.

К стандартам качества данных, разработанным консорциумом W3C, можно отнести стандарт PROV [29], который устанавливает модель описания про-

исхождения данных. Этот стандарт определяет множество нефункциональных метаданных, которые фиксируются при любой манипуляции с данными. Они включают сведения об авторстве, способе получения, преобразованиях данных, информацию о том, как данные использовались в качестве исходных при создании рассматриваемых данных, и многое другое. Однако на практике доступные данные чаще всего сопровождаются минимальной информацией, ограниченной сведениями об авторстве, принадлежности и времени создания наборов данных, и редко используют все богатые возможности, предоставляемые этим стандартом.

Помимо этого, в рамках консорциума W3C был разработан подход к спецификации качества данных DQV (Data Quality Vocabulary) [30] на основе RDF-технологий. Определения в этой модели позволяют описывать словарь показателей качества наборов данных, включая наборы показателей (dimension) и методы оценки показателей качества (metric), такие как оценки объема, полноты, точности и других характеристик содержимого каталогов данных. Для связи с описываемыми каталогами данных используются спецификации в формате DCAT (Data Catalog Vocabulary) [31].

Спецификация DCAT позволяет описывать каталоги, наборы данных, более общие концепции ресурсов, их представления в различных форматах и источниках, отдельные записи, а также сервисы данных, которые предоставляют доступ к данным через программные интерфейсы для запросов и получения нужных частей данных. Таким образом, спецификации качества данных могут быть связаны как с целыми каталогами и наборами данных, так и с их отдельными фрагментами. Рекомендации DCAT развиваются, в версии 2 добавлены некоторые возможности идентификации данных и обратная связь с моделью DQV, а в недавней версии 3 добавлена поддержка версий и серий наборов данных. Обратная связь с моделью DQV обеспечивает расширение возможностей спецификации качества данных при развитии DCAT.

В целом подходы к построению моделей качества данных можно считать устоявшимися и достаточно гибкими для дальнейшего развития в рамках существующих концепций. Однако, несмотря на это, такие подходы редко применяются для описания данных в виде четко определенных моделей. Метаданные, связанные с качеством данных, чаще всего присутствуют в документации в произвольной форме или включены непосредственно в наборы данных, но не выделены как отдельные метаданные качества и плохо структурированы. Это особенно заметно, например, в случае с погрешностями измерения характеристик объектов. Часто значения характеристик и их погрешности представляются в каталогах как равноценные атрибуты объекта.

### **3. Метаданные качества и подход к управлению качеством данных**

Спецификации качества данных в формате DQV [30] предлагают RDF-схему, которая позволяет задавать структуру для описания качества данных.

Эта схема, среди прочего, включает:

- наборы показателей качества (Dimension): характеристики, которые используются для оценки различных аспектов качества данных;
- категории показателей (Category): классификация показателей по определенным признакам, что облегчает их организацию и использование;
- значения показателей качества (QualityMeasurement): конкретные значения, которые отражают уровень качества по заданным показателям;
- методы оценки показателей качества (Metric): методы вычисления или формирования показателей качества. Они описывают, каким именно способом было оценено качество данных, а также типы данных, которые используются для представления значений показателей качества;
- наборы данных (Dataset) и их представления (Distribution): связывание метаданных качества с конкретными наборами данных или их копиями в определенных форматах;
- сервисы данных (DataService): связывание метаданных качества с сервисами, предоставляющими доступ к данным.

Модель метаданных качества для спецификации качества может быть построена на основе спецификаций, подобных тем, что предлагаются в DQV. Однако такая базовая модель требует значительного расширения, чтобы учитывать спецификацию метаданных качества на различных уровнях агрегации данных, определение действий с данными в зависимости от их качества, установление требований к качеству искомых или создаваемых данных и другие важные особенности.

Для формирования необходимых видов спецификаций, включаемых в модель качества, была проведена классификация метаданных качества по различным критериям. Связывание элементов этой классификации позволяет определить спецификации показателей качества, установить их связь с описываемыми данными, определить действия с данными на основе их качества, а также сформулировать требования к данным, которые необходимо учитывать.

1. Классификация по измерениям показателей качества:

- объем;
- полнота;
- точность;
- происхождение;
- актуальность;
- надежность;
- и др.

2. Классификация по типу значений показателей качества:

- булевый флаг или битовый вектор;
- вещественное значение;

- набор категориальных значений градаций качества;
- набор категориальных значений с различием разновидностей фактов нарушения или выполнения требований качества;

– те же типы с пустыми значениями (неизвестное качество).

### 3. Классификация по способу спецификации значений:

- константный;
- формулы или правила;
- табличный.

### 4. Классификация по источнику метаданных:

- метаданные набора данных;
- метаданные в составе схемы данных;
- внешние репозитории метаданных;
- метаданные в публикациях;
- оценка на основании эксперимента;
- статистическая оценка данных;
- оценка на представительной выборке данных;
- изменчивые метаданные (требующие переоценки).

### 5. Классификация по способу агрегации данных:

- набор данных;
- отношение;
- выборка данных (срез, условия запроса);
- связанная семантика сущностей или процессов;
- кортеж;
- конкретная сущность или процесс;
- атрибут;
- связанная семантика атрибутов;
- значение атрибута.

### 6. Классификация действий по улучшению качества данных:

– удаление данных в соответствии с уровнем агрегации (набор данных, выборка, кортеж, сущность, атрибут, значение);

- назначение метаданных качества (показатель, вес);
- выбор качественных данных;
- обогащение данных из внешних источников;
- методы усреднения с учетом качества данных;
- методы усреднения без учета качества данных;
- игнорирование качества данных (включение всех данных без учета качества).

### 7. Классификация требований к качеству данных:

- ограничения на состав и качество исходных данных;



- требования к улучшению качества исходных данных;
- декларируемые необходимые оценки качества результатов;
- характеристика зависимости качества результатов от качества и характеристик исходных данных;
- необходимость оценки качества результатов или невозможности их оценить.

В случаях, когда необходимые метаданные на уровне наборов данных отсутствуют или требуют корректировки на подмножествах данных (например, срезах по параметрам), их можно оценить статистически на всех данных или на их представительных выборках, а затем дополнить существующие метаданные в том же формате.

Метаданные качества на уровне кортежей или значений атрибутов обычно включаются непосредственно в состав данных, представленных в каталогах или наборах данных, в виде дополнительных атрибутов отношений.

Метаданные качества, относящиеся к определенному уровню агрегации данных, распространяются на вложенные уровни агрегации. Таким образом, любой используемый кортеж или значение данных можно связать с метаданными качества, собранными с разных уровней представления данных, включая метаданные набора данных в целом, метаданные атрибутов, значений и др. При формулировании требований к метаданным качества могут использоваться ограничения как на метаданные целых наборов данных, так и на метаданные конкретных значений данных или их происхождения.

Часто применяются подходы к оценке веса качества на основе показателей качества для принятия решений о дальнейших действиях. Вес может выражаться как числовыми значениями, так и специальными обозначениями, такими как отсутствие оценки (naп) или указание на необходимость удаления данных (del). Если значения атрибутов качества отсутствуют, весу качества присваивается значение naп, что не влияет на общий вес качества кортежа. Если хотя бы одно значение веса для кортежа обозначено как del, это означает, что кортеж должен быть исключен из анализа, и другие значения весов не учитываются. Веса качества кортежей, определенные на уровне выборки или набора данных, влияют на общий вес качества кортежа. Наличие нескольких весов для кортежа обобщается путем вычисления их среднего значения. Эти же принципы применяются к весам качества отдельных значений атрибутов. Таким образом, веса качества кортежей и атрибутов также распространяются на веса каждого из значений атрибутов.

#### **4. Пример решения задач с применением спецификаций качества**

В качестве примера применения модели управления качеством данных была разработана модель качества, и решалась проблема улучшения качества данных в области звездной астрономии на основе множества больших

многоцветных фотометрических обзоров неба. Для исследования разных типов звезд, их спектров, восстанавливаемых по фотометрическим данным в разных диапазонах излучения, а также их наблюдаемых параметров и оценочных значений астрофизических параметров была поставлена задача перекрестного отождествления и слияния данных об одних и тех же объектах из различных фотометрических каталогов. Пример анализа представлен на данных каталогов SDSS [32], UKIDSS [33], GALEX [34, 35].

Исследование каталогов показывает множество факторов, влияющих на качество данных и отражающихся на качестве отождествления объектов.

- Каталоги имеют разное покрытие неба в связи с расположением обсерваторий, проводивших наблюдения.

- Каждый каталог представлен данными, полученными посредством различных наборов фотометрических фильтров, имеющихся в телескопах и обеспечивающих наблюдения в разных диапазонах спектра излучения объектов.

- Телескопы обладают разным оптическим разрешением и соответственно имеют различную точность позиционирования (калибровки) и различения объектов. Ожидается, что точность различения объектов в наблюдениях в плоскости Галактики может оказаться ниже, поскольку количество объектов в поле наблюдения больше.

- Телескопы имеют разные пределы минимальной и максимальной звездных величин объектов: величины блеска ниже минимального регистрируются в рамках погрешности, а величины выше максимального оказываются засвеченными и не отражают реальные значения.

- Каталоги включают различные наборы флагов качества наблюдений, а также флагов для классификации объектов, обнаружения артефактов и других сведений о наблюдениях.

- Наблюдения, проводимые в разные эпохи, отражаются на позиции объектов вследствие их перемещения и на интенсивности блеска в результате его переменности.

- Искажения данных могут также быть связаны с различными условиями наблюдения: временем года и суток, положением объекта относительно зенита и погодными условиями.

- Расположение объекта на краях области наблюдения влияет на уровень зашумления по сравнению с наблюдениями в центре области, а также может приводить к появлению бликов, связанных с оборудованием.

Задача перекрестного отождествления каталогов и обзоров неба остается актуальной в астрономии и обычно решается для пары каталогов с разными требованиями к качеству отождествления. В соответствии с требованиями конкретных задач проводится предобработка данных или удаление наблюдений с низким качеством. При этом процедура может варьироваться в зависимости от специфики каждого случая.

Исходя из анализа проблемы, были поставлены следующие задачи.

1. Оценка покрытия каталогами: необходимо определить, в достаточной ли степени каталоги охватывают площадки в определенных направлениях обзора неба. Важно, чтобы в одной площадке были данные из нескольких обзоров с разными наборами фильтров, что позволит охватить большинство диапазонов излучения в спектре объектов.

2. Метод перекрестного отождествления: необходимо разработать метод решения задачи перекрестного отождествления обзоров разного качества (разрешения), включающий отождествление множественных наблюдений объектов как внутри одного каталога, так и между различными каталогами.

3. Оценка радиусов отождествления: необходимо решить задачу оценки радиусов отождествления объектов между каталогами в зависимости от направления наблюдения относительно плоскости Галактики.

4. Оценка качества данных: требуется оценить качество данных об объектах наблюдений на основании значений метаданных каталогов и флагов, присутствующих в структурах каталогов.

5. Формирование списков отождествления: необходимо разработать подход к формированию списков отождествления объектов и слиянию данных для получения кортежей, описывающих блеск объектов в разных диапазонах излучения.

6. Представление результатов слияния: требуется представить результаты слияния в виде каталога и связанных с ним метаданных.

#### *4.1. Проблемы полноты данных*

Задача 1, связанная с проблемой полноты данных в обзорах, предстает в двух аспектах.

Во-первых, речь идет о покрытии данными используемых каталогов большинства направлений на небе. Это покрытие существенно влияет на решение таких задач, как оценка позиционной точности каталогов и возможность оценки поглощения в межзвездной среде в различных направлениях на небе, что также зависит от угла наблюдения относительно плоскости Галактики. Показатели полноты данных в заданных направлениях оцениваются статистически на основе данных каталогов. В этом контексте важно не процентное покрытие неба, а покрытие крупных площадок в разных направлениях. Для конкретной площадки можно подбирать состав каталогов, который обеспечивает хорошее покрытие.

Во-вторых, полнота данных, представленных фотометрическими каталогами, оценивается с точки зрения покрытия большинства диапазонов наблюдаемого спектра излучения звезд. Эта характеристика имеет решающее значение для оценки спектра на основе данных фотометрии и зависит от набора фильтров, используемых в телескопах. Она определяется набором полей в структуре каталогов.

Очевидно, что данные ни одного из больших фотометрических обзоров в отдельности не соответствуют этим двум требованиям. Именно поэтому

возникает необходимость интеграции и отождествления ряда каталогов, чтобы данные, полученные в результате слияния каталогов, были пригодны для оценки спектра звезд, их дальнейшей классификации и параметризации.

#### *4.2. Проблемы точности данных*

В задаче 2 проблема ошибок отождествления множественных наблюдений объектов в каталоге с низким разрешением (GALEX) и затем усреднения ошибочных результатов отождествления до одного кортежа решается с помощью выбора последовательности отождествления каталогов. Сначала все кортежи этого каталога отождествляются с каталогами более высокого качества. Затем, если необходимо получить усредненный кортеж для множественных наблюдений каталога, усредняются результаты перекрестного отождествления с более качественными каталогами. При перекрестном координатном отождествлении объектов из разных каталогов оценивается совместная точность, рассчитываемая как среднеквадратическая ошибка, учитывающая обе неопределенности. Таким образом, достигается меньшая оценка совместной точности (среднеквадратической ошибки) и, соответственно, снижается вероятность неверного отождествления.

При решении задачи 3 радиус отождествления объектов обычно принимается как константа: для каталогов SDSS и UKIDSS – 1 угловая секунда, для GALEX – 3 угловых секунды. Эти значения связаны с позиционной точностью и оптическим разрешением каталогов (точностью различения объектов). Разные значения точности координат можно найти в статьях о каталогах, однако описание качества координат остается неоднозначным. Более того, желательно оценивать радиус отождествления объектов в зависимости от направления на небе. Поэтому обычный подход оказывается недостаточным, и возникает необходимость в определении оптимального радиуса отождествления.

Задача выбора радиуса отождествления решается на основе статистической оценки, которая максимизирует количество уникальных вариантов координатного отождествления для радиуса в конкретном направлении наблюдения относительно плоскости Галактики [36]. Таким образом, по данным или по представительным выборкам определяется точность позиционирования в каталогах, и формируются метаданные спецификаций точности для них. Эти метаданные могут быть представлены таблично в зависимости от площадки в определенном направлении на небе и впоследствии используются для каждого значения координат объектов площадки в качестве его точности.

#### *4.3. Проблемы надежности данных*

Задача 4 связана с оценкой точности данных в разных каталогах, которая зависит от целого ряда факторов:

— позиционная точность объектов: зависит от разрешающей способности оборудования, с помощью которого велись наблюдения неба;

— расположение объектов в области наблюдения: объекты, находящиеся в центре области наблюдения, имеют более высокую точность, чем те, что расположены на краю;

— засвеченность фона объектов: яркие звезды, расположенные близко к наблюдаемым объектам, могут создавать помехи;

— артефакты: такие как планеты или спутники, пролетающие в кадре объектов, а также блики от линз;

— засвеченность ярких объектов: яркие звезды на верхней границе звездных величин, определяемых оборудованием, могут искажать данные об их блеске;

— обнаружение объектов: объекты на нижней границе звездных величин могут оставаться незамеченными;

— точность значений блеска: для каждого значения блеска указывается его точность в определенном диапазоне излучения.

Некоторые атрибуты каталогов представляют собой флаги, которые указывают на тип объекта, дополнительные характеристики его наблюдений и их качество. При подготовке данных учитываются ограничения, накладываемые на атрибуты каталогов, не относящиеся к качеству данных, в соответствии с семантикой задачи.

— Атрибуты  $F_r$  и  $N_r$  в GALEX указывают угловой размер объекта на полувысоте свечения. Ограничением отсекаются кортежи со значениями этих атрибутов, превышающими 0,003 (для звезд).

— Атрибут  $class$  в SDSS предоставляет информацию о типе объекта: звезда, галактика, артефакт и др. Ограничением отсекаются кортежи со значениями, отличными от 6 (звезда).

Далее необходимо оценить качество кортежей в целом как наблюдений, так и значений характеристик объектов в них. В каталоге GALEX используются следующие флаги дополнительной информации о наблюдениях и качестве данных.

— Атрибуты  $F_{exf}$  и  $N_{exf}$  – векторы битовых значений, указывающие на различные виды некачественных наблюдений: объекты, имеющие соседей, смешанные с другими объектами, усеченные границей наблюдения и др. При значениях этих флагов, не равных 0, значению блеска присваивается вес качества  $del$ .

— Атрибуты  $F_{afl}$  и  $N_{afl}$  – векторы битовых значений, указывающие на различные виды артефактов: блики, пятна и др. При значениях этих флагов, не равных 0, значению блеска также присваивается вес качества  $del$ .

— Атрибуты  $nS/G$  и  $fS/G$  – вещественные значения, оценивающие, является ли объект звездой или галактикой. Можно сказать, что это классификация объекта с указанием степени ее надежности. Назначается вес качества кортежа, равный 1 для значений параметров, превышающих 0,5; в противном случае вес равен 0.

В каталоге SDSS используется один существенный флаг качества кортежа:

– Атрибут Q вводит категориальные градации и оценивает качество наблюдения: 1 – низкое, 2 – приемлемое, 3 – высокое. При значении атрибута, равном 1, назначается вес качества кортежа, равный 0; при значении 2 – вес, равный 0,5; при значении 3 – вес, равный 1.

В каталоге UKIDSS используются следующие флаги качества:

– Атрибут cl – смешанный категориальный и градационный показатель, оценивающий, является ли объект звездой, галактикой или шумом. Значения атрибута могут указывать на отнесение к тому или иному классу с указанием степени надежности. Присваивается значение веса качества кортежа del при всех значениях, кроме  $-2$  и  $-1$  (звезды).

– Атрибут  $r^*$  – вещественное значение, отчасти дублирующее и уточняющее значение атрибута cl, оценивающее вероятность того, что объект является звездой. При значении атрибута, превышающем 0,7, присваивается значение веса качества кортежа 1; при значении ниже 0,3 – значение 0; иначе значение 0,5.

– Атрибуты pG и pN оценивают подобную вероятность для галактик и шума и не рассматриваются в данной задаче. Кортежи, классифицированные как галактика или шум с помощью атрибута cl, удаляются.

То, каким именно образом формируются веса качества, кортежей или значений на основании значений флагов, было задано постановкой задачи во взаимодействии со специалистами в области звездной астрономии. При отсутствии значений атрибутов качества весу качества присваивается значение nap. Таким образом, на основании флагов были определены веса качества кортежей и некоторых значений атрибутов в кортежах. Обобщение весов качества для получения общих оценок качества кортежей или значений производится в соответствии с принципами, приведенными в предыдущем разделе. Обобщенные веса качества данных учитываются в дальнейшем при слиянии данных.

#### *4.4. Слияние данных*

При решении задачи 5, связанной с формированием списков отождествления объектов, применяются принципы работы с отождествляемыми объектами из нескольких каталогов, предложенные в [37]. Используются связи отождествлений, которые исключают транзитивные связи отождествления объектов между несколькими каталогами. Слияние данных производится на основании кортежей, входящих в связи отождествлений, и их обобщенных весов качества.

В зависимости от решаемых задач над отождествленными объектами слияние данных может выполняться по разным принципам. Например, для решения задачи параметризации объектов важны только качественные данные, и любые некачественные наблюдения удаляются из рассмотрения. Однако, если необходимо сохранить некачественные объекты для обеспечения

полноты данных и повышения их качества, такие данные могут быть включены в анализ.

Для идентификации объектов после их слияния и их позиционирования используются координаты из самого точного каталога (в примере это SDSS). Для отождествленных множественных наблюдений одного каталога, собираемых в один кортеж, координаты могут усредняться без учета весов.

Флаги в основном касаются качества наблюдений блесков объектов на небе. Данные блесков при слиянии могут быть усреднены по данным наблюдений в одном диапазоне излучения.

В большинстве каталогов присутствуют поля, указывающие на погрешность ( $\sigma$ ) измерения блесков ( $m$ ). Для оценки погрешности результата слияния при выборе только надежных кортежей (с высоким весом качества:  $r = 1$ ) может использоваться взвешенное среднее значение. При этом вес обратно пропорционален квадрату погрешности, что позволяет минимизировать суммарную дисперсию. Если при решении задачи используются не только надежные кортежи, необходимо учитывать оценку веса качества ( $r$ ) кортежа или значения совместно с погрешностью значения блеска. В этом случае при вычислении среднего взвешенного (2) вес, обратный квадрату погрешности, умножается на вес качества значения (1).

$$(1) \quad w = \frac{r}{\sigma^2},$$

$$(2) \quad m = \frac{\sum_i m_i w_i}{\sum_i w_i}.$$

Погрешность результата в этом случае можно оценить следующим образом (3):

$$(3) \quad \sigma = \sqrt{\frac{1}{\sum_i w_i}}.$$

#### 4.5. Формирование результата

Решение задачи 6 заключается в формировании результирующего отношения. Оно создается путем соединения отождествленных кортежей из разных каталогов с обобщенной оценкой атрибутов, для которых были найдены множественные значения. Результат должен объединять в общем кортеже следующие элементы:

- координаты из самого точного каталога, которые также используются для идентификации объекта;
- усредненные взвешенные оценки блесков для каждого диапазона излучения, соответствующего диапазонам каталогов;
- оценки погрешностей для каждого из значений блеска.

Данное отношение может быть представлено в виде самостоятельного каталога, который должен быть обильно снабжен метаданными. Такие метадан-

ные должны включать описание происхождения данных и их качество, что позволит оценивать применимость данных при решении задач и обеспечит возможность их повторного использования.

Описание происхождения данных нового каталога может включать следующие элементы:

- на уровне каталога: ссылки на каталоги, данные которых использовались при формировании результата;
- на уровне атрибутов: ссылки на атрибуты (координаты, блески в соответствующих диапазонах), из данных которых формировались значения атрибутов;
- ссылку на описание метода отождествления кортежей;
- ссылку на описание метода выбора (для координат) и усреднения значений атрибутов (для блесков).

Описание качества данных каталога может включать:

- на уровне отношения информацию о полноте покрытия неба и полноте заполнения атрибутов (охваченных диапазонов наблюдения) в зависимости от направления на небе;
- информацию об оценке позиционной точности результирующих данных;
- информацию о качестве соединенных кортежей;
- информацию о присутствии в атрибутах каталога взвешенных оценок погрешностей блесков и связи их с атрибутами самих значений блеска.

Благодарим О.Ю. Малкова (ИНАСАН) за постановку задачи, используемой для демонстрации представленного подхода.

## 5. Заключение

Исследованы подходы к спецификации качества данных в исследовательских инфраструктурах с множественными неоднородными источниками данных разного качества. Стандарты описания качества данных опираются на многокритериальные модели, обычно не ограничивающие состав и методы оценки показателей качества. Предложено развитие принципов спецификации качества данных для определения способов хранения и доступа к метаданным качества, уровня агрегации оцениваемых данных и учета нефункциональных требований к данным при решении задач и действий над данными. Эти принципы проиллюстрированы при решении задачи в области звездной астрономии. На разных этапах ее решения оцениваются различные показатели качества, включая полноту, точность и надежность данных, используются разные источники метаданных качества, такие как сведения о точности значений физических величин в каталогах, флаги качества кортежей, гарантированное качество из документации к каталогам, статистические оценки на основе выборки данных каталогов. В зависимости от полученных оценок показателей качества используются разные источники данных, отфильтро-



вываются данные, не подходящие по качеству, изменяется последовательность действий в алгоритмах решения задач, формируются и сохраняются метаданные как источников данных, так и промежуточных и окончательных результатов решения задачи.

## СПИСОК ЛИТЕРАТУРЫ

1. *Wand Y., Wang R.* Anchoring data quality dimensions in ontological foundations // Communications of the ACM. New York: ACM, 1996. V. 39. No. 11. P. 86–95.
2. *Ballou D., Pazer H.* Modeling data and process quality in multi-input, multi-output information systems // Management Sci. 1985. V. 31. No. 2. P. 150–162. <https://doi.org/10.1287/mnsc.31.2.150>
3. *Wang R., Strong D.* Beyond accuracy: What data quality means to data consumers // J. Management Inform. Syst. 1996. V. 12. No. 4. P. 5–33. URL: <http://www.jstor.org/stable/40398176>
4. *Batini C., Scannapieco M.* Data quality: concepts, methodologies and techniques. Heidelberg: Springer, 2006. 262 p. <https://doi.org/10.1007/3-540-33173-5>
5. ГОСТ Р 56214-2014. Качество данных. Часть 1. Обзор. М.: Стандартинформ, 2015.
6. ГОСТ Р 57773-2017. Пространственные данные. Качество данных. М.: Стандартинформ, 2017.
7. *Wilkinson M., Dumontier M., Aalbersberg I., et al.* The FAIR Guiding principles for scientific data management and stewardship // Sci. Data 2016. V. 3. Article 160018. <https://doi.org/10.1038/sdata.2016.18>
8. FAIR data maturity model. Specification and guidelines. Version 1.0. RDA FAIR Data Maturity Model Working Group. Geneva: Zenodo, 2020. <https://doi.org/10.15497/rda00050>
9. FAIRsFAIR. Fostering FAIR Data Practices in Europe. URL: <https://www.fairsfair.eu/>
10. *Devaraju A., Mokrane M., Cepinskas L., et al.* From conceptualization to implementation: FAIR Assessment of Research Data Objects // Data Sci. J. 2021. V. 20. No. 1. Article 4. <https://doi.org/10.5334/dsj-2021-004>
11. The FAIR cookbook for FAIR doers. URL: <https://faircookbook.elixir-europe.org/>
12. *Harrow J., Drysdale R., Smith A., et al.* ELIXIR: providing a sustainable infrastructure for life science data at European scale // Bioinformatics. Oxford: Oxford University, 2021. V. 37. No. 16. P. 2506–2511. <https://doi.org/10.1093/bioinformatics/btab481>
13. ELIXIR Platforms. URL: <https://elixir-europe.org/platforms>
14. Recommendations from the Data Quality Working Group. NASA ES DSWG, 2019. URL: <https://www.earthdata.nasa.gov/esdis/esco/standards-and-practices/recommendations-from-the-data-quality-working-group>
15. Data Quality Working Group’s comprehensive recommendations for data producers and distributors. NASA ES DSWG, 2019. URL: <https://www.earthdata.nasa.gov/s3fs-public/imported/ESDS-RFC-033.pdf>
16. ESIP Information Quality Cluster. Earth Science Information Partners (ESIP). URL: [http://wiki.esipfed.org/index.php/Information\\_Quality](http://wiki.esipfed.org/index.php/Information_Quality)

17. *Peng G., Privette J., Kearns E., et al.* A unified framework for measuring stewardship practices applied to digital environmental datasets // *Data Sci. J.* 2015. V. 13. No. 2. P. 231–253. <https://doi.org/10.2481/dsj.14-049>
18. ISO 19157-1:2023 Geographic information - Data quality. Part 1. General requirements. Geneva: ISO, 2023. URL: <https://www.iso.org/standard/78900.html>
19. *Sirotnak C., Cook J.* The total economic impact of Talend. Cost savings and business benefits enabled by Talend Solutions. Cambridge: Forrester, 2023. URL: <https://www.talend.com/lp/the-total-economic-impact-of-talend/>
20. *Chien M., Medd J.* Magic Quadrant for Augmented Data Quality Solutions. Stamford: Gartner, 2024. URL: <https://www.gartner.com/en/documents/5257863>
21. *Fürber C.* Data quality management with semantic technologies. Thesis. Wiesbaden: Springer Gabler, 2016. <https://doi.org/10.1007/978-3-658-12225-6>
22. *Berners-Lee T., Hendler J., Lassila O.* The Semantic Web // *Scientific American* 2001. V. 284. No. 5. P. 34–43. URL: <https://www.jstor.org/stable/26059207>
23. *Cyganiak R., Wood D., Lanthaler M. (eds.)*. RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation. Wakefield: W3C, 2014. URL: <http://www.w3.org/TR/rdf11-primer/>
24. *Fürber C., Hepp M.* Towards a vocabulary for data quality management in Semantic Web architectures // *Proceedings of the 1st International Workshop on Linked Web Data Management (LWDM2011)*. New York: ACM, 2011. P. 1–8. <https://doi.org/10.1145/1966901.1966903>
25. *Hartig O., Zhao J.* Provenance Vocabulary Core Ontology Specification. San Diego: SourceForge, 2012. URL: <https://trdf.sourceforge.net/provenance/ns.html>
26. *Taleb I., Taleb, Serhani M., Bouhaddioui C., et al.* Big data quality framework: a holistic approach to continuous quality management // *J. of Big Data* 2021. V. 8. Article 76. <https://doi.org/10.1186/s40537-021-00468-0>
27. *Gallo R.* Data quality with FAIR principles, an introduction. The Hyve, 2024. URL: <https://www.thehyve.nl/articles/data-quality-with-fair-principles>
28. *Skvortsov N.* The principles of data reuse in research infrastructures // *Proceedings of the International Conference Common Digital Space of Scientific Knowledge: Problems and Solutions (CDSSK 2020)*. Aachen: CEUR WS, 2021. V. 2990. P. 62–74. URL: <https://ceur-ws.org/Vol-2990/rpaper6.pdf>
29. PROV-Overview: An overview of the PROV family of documents. W3C Working Group Note. Wakefield: W3C, 2013. URL: <http://www.w3.org/TR/prov-overview/>
30. Data on the Web Best Practices: Data quality vocabulary. W3C Working Group Note. Wakefield: W3C, 2016. URL: <https://www.w3.org/TR/vocab-dqv/>
31. *Albertoni R., Isaac A. (eds.)*. Data catalog vocabulary (DCAT), Version 3. W3C Recommendation. Wakefield: W3C, 2024. URL: <https://www.w3.org/TR/vocab-dcat/>
32. *Alam S., Albareti F., Prieto C., et al.* The eleventh and twelfth data releases of the Sloan Digital Sky Survey: Final data from SDSS-III // *Astrophys. J. Suppl. Ser.* 2015. V. 219. No. 1. P. 12. <https://doi.org/10.1088/0067-0049/219/1/12>
33. *Lawrence A., Warren S., Almaini O., et al.* The UKIRT Infrared Deep Sky Survey (UKIDSS) // *Mon. Not. R. Astron. Soc.* 2007. V. 379. No. 4. P. 1599–1617. <https://doi.org/10.1111/j.1365-2966.2007.12040.x>

34. *Bianchi L., Herald J., Efremova B., et al.* GALEX catalogs of UV sources: statistical properties and sample science applications: hot white dwarfs in the Milky Way // *Astrophys. Space Sci.* 2011. V. 335. No. 1. P. 161–169. <https://doi.org/10.1007/s10509-010-0581-x>
35. *Bianchi L., Shiao B., Thilker D.* Revised catalog of GALEX ultraviolet sources. I. The All-Sky Survey: GUVcat\_AIS // *Astrophys. J. Suppl. Ser.* 2017. V. 230. No. 2. P. 24. <https://doi.org/10.3847/1538-4365/aa7053>
36. *Malkov O., Dluzhnevskaya O., Karpov S., et al.* Cross catalogue matching with Virtual Observatory and parameterization of stars // *Open Astronomy* 2012. V. 21. No. 3. P. 319–330. <https://doi.org/10.1515/astro-2017-0390>
37. *Gray J., Szalay A., Budavari T., et al.* Cross-Matching Multiple Spatial Observations and Dealing with Missing Data. Microsoft Technical Report, MSR-TR-2006-175. Redmond: Microsoft Research, 2006. <https://doi.org/10.48550/arXiv.cs/0701172>

*Статья представлена к публикации членом редколлегии А.А. Галяевым.*

Поступила в редакцию 29.11.2024

После доработки 10.01.2025

Принята к публикации 14.01.2025

# Оптимизация, системный анализ и исследование операций

© 2025 г. О.В. СЕНЬКО, д-р физ.-мат. наук (senkoov@mail.ru),  
А.А. ДОКУКИН, канд. физ.-мат. наук (dalex@ccas.ru)  
(Федеральный исследовательский центр  
«Информатика и управление» РАН, Москва),  
Ф.А. МЕЛЬНИК (melnik.tedor@gmail.com)  
(Московский государственный университет им. М.В. Ломоносова)

## ИСПОЛЬЗОВАНИЕ АНСАМБЛЕЙ С УВЕЛИЧЕННОЙ ДИВЕРГЕНЦИЕЙ В ПРОСТРАНСТВЕ ПРОГНОЗОВ В РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМАХ

Рассматривается метод дивергентного решающего леса, основанный на достижении более высокой дивергенции в пространстве прогнозов по сравнению со стандартным случайным решающим лесом за счет включения на каждом шаге в ансамбль нового дерева  $T_x$ , которое строится исходя из условий минимизации специального функционала, являющегося разностью квадратичной ошибки  $T_x$  и квадрата расхождения прогнозов  $T_x$  и текущего ансамбля. Метод является развитием аналогичных ранее разработанных методов, которые предназначены для прогнозирования числовых переменных. В работе представлены результаты применения метода дивергентного решающего леса для решения задач классификации, возникающих при создании рекомендательных систем. Исследована зависимость эффективности прогноза от глубины деревьев и одного из ключевых параметров алгоритма, регулирующего вклад двух составляющих в минимизируемый функционал. Исследования показали, что точность предлагаемой технологии заметно превышает точность случайного решающего леса и близка к точности метода CatBoost.

*Ключевые слова:* ансамблевый метод, машинное обучение, рекомендательные системы.

**DOI:** 10.31857/S0005231025040061, **EDN:** CAPHAM

### 1. Введение

Методы, основанные на использовании ансамблей регрессионных или решающих деревьев, являются одной из основных ветвей современного машинного обучения и активно используются при решении разнообразных прикладных задач [1]. Можно выделить два популярных направления ансамблевых алгоритмов: случайные леса [2] и градиентный бустинг [3]. Для обеих технологий разработаны методы, предназначенные для решения как задач

автоматической классификации, так и задач прогнозирования числовых переменных. При прогнозировании целевой числовой переменной  $Y$  по признакам  $X_1, \dots, X_n$  обучение производится по выборке  $S = \{s_1 = (y_1, x_1), \dots, s_m = (y_m, x_m)\}$ , где  $y_j$  и  $x_j$  – значения целевой переменной  $Y$  и вектора значений признаков  $X_1, \dots, X_n$  объекта  $s_j$  соответственно. Задачи классификации с двумя непересекающимися классами также могут рассматриваться как задачи прогнозирования бинарной числовой переменной с обучением по выборке аналогичного вида. Сформулируем основные различия между методами, основанными на градиентном бустинге, и случайными лесами. В методе случайных лесов отдельные деревья строятся независимо. При этом дерево на шаге  $k$  нацелено на прогнозирование целевой переменной  $Y$  и строится по выборке  $S_k$ , получаемой с использованием процедур бэггинга [4] и метода случайных подпространств [5]. Процедура бэггинга заключается в построении новой выборки как случайной выборки с возвращениями из  $S$ . Метод случайных подпространств предполагает использование случайного подмножества исходного множества признаков  $X_1, \dots, X_n$  фиксированного размера при обучении каждого нового дерева случайного леса. Выходной прогноз ансамбля  $\hat{A}_k$ , состоящего из  $k$  деревьев, вычисляется как средний прогноз по всем деревьям, вошедшим в него, т.е.  $\hat{A}_k = \frac{1}{k} \sum_{j=1}^k A_j$ . В методе градиентного бустинга оптимальный алгоритм ищется как линейная комбинация регрессионных деревьев  $\hat{A}_N = T_0 + \alpha_1 \times T_1 + \dots + \alpha_N \times T_N$ , где исходный алгоритм  $T_0$  обычно вычисляет прогноз, тождественно равный среднему значению  $Y$ . На каждом шаге в текущую линейную комбинацию  $\hat{A}_{k-1} = A_0 + \alpha_1 \times T_1 + \dots + \alpha_{k-1} \times T_{k-1}$  добавляется новое слагаемое  $\alpha_k \times T_k$ . При этом выбор последнего производится исходя из цели минимизации функции потерь для линейной комбинации  $\hat{A}_k = T_0 + \alpha_1 \times T_1 + \dots + \alpha_k \times T_k$ . Последнее достигается за счет использования процедуры градиентного спуска.

Пусть  $L(\hat{A}_{k-1}, S)$  – некоторая функция потерь, зависящая от прогнозов, вычисляемых  $\hat{A}_{k-1}$  и истинных значениях  $Y$  на объектах  $S$ . Сопоставим прогнозам, вычисляемым для объектов  $S$ , переменные  $z_1, \dots, z_m$ . Согласно методу градиентного спуска предполагаемый минимум  $L$  будет достигаться для прогноза на  $k$ -м шаге для объекте  $s_j$  при условии

$$\hat{A}_k(s_j) = \hat{A}_{k-1}(s_j) - \eta \frac{\partial L}{\partial z_j} \Big|_{z_j = \hat{A}_{k-1}(s_j)}.$$

Однако прогноз в указанном виде невозможно вычислить для объектов, для которых истинное значение  $y_j$  неизвестно, поскольку функция потерь и ее частные производные определены только при известных значениях  $Y$ . Поэтому авторами метода было предложено использовать вместо величин  $g_j$  их прогнозы, вычисленные с помощью регрессионного дерева, обученного по выборке  $\{(g_1, x_1), \dots, (g_m, x_j)\}$ . Именно это дерево и является деревом  $T_k$ , добавляемым в линейную комбинацию. При обучении дерева  $T_k$  использовались бэггинг и метод случайных подпространств.

В настоящее время существует несколько модификаций градиентного бустинга, включающих XGBoost [6], LightGBM [7], CatBoost [8]. При этом метод случайного леса остается практически неизменным с момента своего создания. Вместе с тем отсутствуют убедительные доказательства того, что используемая в случайных лесах схема генерации ансамбля действительно является оптимальной при использовании простого усреднения в качестве агрегирующей процедуры.

## 2. Построение лесов с увеличенной дивергенцией

В [9–11] был предложен новый подход, нацеленный на построение ансамбля исходя из условия минимизации потерь прогнозов, вычисляемых как средние прогнозы по ансамблю. Иными словами, была поставлена задача подбора деревьев в ансамбль таким образом, чтобы потери были по возможности минимальными при использовании в качестве прогноза  $\hat{A}_k$ . При этом в качестве функционала потерь использовалась обычная среднеквадратичная ошибка:

$$L(S, A) = \frac{1}{m} \sum_{j=1}^m \left( y_j - \hat{A}_k(x_j) \right)^2.$$

Было показано, что средний квадрат ошибки для алгоритма  $A_k$  вычисляется по формуле

$$(1) \quad L(S, \hat{A}_k) = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^k (y_j - T_i(x_j))^2 - \frac{1}{km} \sum_{j=1}^m \sum_{i=1}^k \left( \hat{A}_k(x_j) - T_i(x_j) \right)^2.$$

Можно поставить задачу поиска такого набора деревьев  $T_1, \dots, T_k$ , при котором достигается минимум потерь (1). Данная задача является чрезвычайно трудоемкой. Однако может быть использован простой эвристический подход, когда на каждом шаге в ансамбль добавляется дерево, которое одновременно хорошо аппроксимирует зависимость и максимально удалено по вычисляемым прогнозам от текущего ансамбля. Для реализации такого подхода в [9–11] было предложено на шаге  $k$  добавлять дерево  $T_k$ , при котором по возможности минимален был бы функционал

$$(2) \quad Q(T_k, \hat{A}_k, \mu) = \frac{1}{m} \sum_{j=1}^m (y_j - T_k(x_j))^2 - \mu \frac{1}{m} \sum_{j=1}^m \left( \hat{A}_k(x_j) - T_k(x_j) \right)^2,$$

где  $\mu$  – параметр из интервала  $(0, 1)$ . Функционал  $Q(T_k, \hat{A}_k, \mu)$  может быть переписан в виде

$$(3) \quad Q(T_k, \hat{A}_k, \mu) = \frac{1}{m} \sum_{j=1}^m (y_j - T_k(x_j))^2 - \theta \frac{1}{m} \sum_{j=1}^m \left( \hat{A}_k(x_j) - T_k(x_j) \right)^2, \quad \theta = \frac{k^2 \mu}{(k+1)^2}.$$

Отметим, что минимум функционала (2) достигается, если каждой точке  $x_j$  сопоставить прогноз  $t_j^o$ , равный значению вспомогательной переменной  $t_j$ , при котором достигается минимум функции

$$(y_j - t_j)^2 - \theta \left( \hat{A}_{k-1}(x_j) - t_j \right)^2.$$

Нетрудно показать, что необходимые условия экстремума выполняются в случае, когда

$$(4) \quad t_j^o = \frac{y_j - \theta \hat{A}_{k-1}(x_j)}{1 - \theta}.$$

Очевидно, что  $\theta \rightarrow 1$  при  $k \rightarrow \infty$  и  $\mu = 1$ , что приводит к неустойчивости оценок по формуле (4). С этим и связано использование дополнительного множителя  $\mu \in (0, 1)$ . Для поиска оптимального ансамбля в [10, 11] был предложен подход, в котором на шаге  $k$  в ансамбль добавляется новое дерево  $T_k$ , прогнозирующее величины  $t_j^o$ , рассчитываемые по (4) в зависимости от векторных описаний. Дерево  $T_k$  обучается по выборке  $\{(t_1^o, x_1), \dots, (t_m^o, x_m)\}$ . Использование гиперпараметра  $\mu$ , оказывающего влияние на эффективность построенных ансамблей, является эвристическим приемом. Теоретически обоснованный способ расчета оптимального значения  $\mu$  отсутствует. Предполагается, что оптимальное значение  $\mu$  может быть подобрано в результате экспериментов, что и является одной из целей настоящего исследования.

### 3. Применение для автоматической классификации

Изложенный выше подход основан на использовании квадратичных функций потерь и предназначен для решения задач прогнозирования числовых переменных. Однако метод может быть использован также для решения задач автоматической классификации. При решении задач автоматической классификации для оценки точности аппроксимации зависимости принято использовать кросс-энтропию, т.е. величину, пропорциональную взятому со знаком минус логарифму функции правдоподобия, рассчитанной исходя из распределения Бернулли. В случае задачи бинарной классификации целевая величина  $Y$  принимает значения из множества  $\{0, 1\}$ , где равенство  $Y = 1$  указывает на принадлежность целевому классу  $K$ . Предположим, что некоторое дерево  $T$  вычисляет вероятности принадлежности классу  $K$  объектов из  $S = \{s_1 = (y_1, x_1), \dots, s_m = (y_m, x_m)\}$ . Пусть  $p_j$  – вероятность принадлежности объекта  $s_j$  целевому классу, вычисляемая деревом  $T$  от его описания  $x_j$ . Потери в смысле кросс-энтропии для дерева  $T$  на выборке  $S$  оцениваются по формуле

$$(5) \quad L(T, S) = -\ln \prod_{j=1}^m p_j^{y_j} (1 - p_j)^{1-y_j} = -\sum_{j=1}^m [y_j \ln p_j + (1 - y_j) \ln(1 - p_j)].$$

Потери по формуле (5) являются аналогом квадратичных потерь для дерева  $T$  на выборке  $S$  и соответствуют левому слагаемому в (2). Для оценки с использованием отклонения добавляемого в ансамбль дерева  $T_k$  от ансамбля  $A_k$  может быть предложен следующий подход. Пусть  $p_j^k$  – вероятность принадлежности объекта  $s_j$  целевому классу, вычисляемая деревом  $T_k$  от его описания  $x_j$ ,  $\hat{p}_j^k$  – вероятность принадлежности объекта  $s_j$  целевому классу, вычисляемая ансамблем  $A_k$ . Отклонение добавляемого дерева от ансамбля оценим по формуле

$$D(T_k, A_k, S) = -\ln \prod_j^m (p_j^k)^{\hat{p}_j^k} (1 - p_j^k)^{1 - \hat{p}_j^k} = \sum_{j=1}^m \left[ -\hat{p}_j^k \ln p_j^k - (1 - \hat{p}_j^k) \ln(1 - p_j^k) \right].$$

Для поиска оптимальных значений вероятностей  $p_j^k$  может быть использован тот же самый подход, что использовался в случае квадратичных потерь. На первом этапе ищутся величины  $p_1^o, \dots, p_m^o$ , для которых достигается минимум функционала

$$Q(T_k, A_k, S) = L(T, S) - \mu D(T_k, A_k, S).$$

Далее новое, добавляемое в ансамбль дерево  $T_k$  обучается по выборке

$$\{(p_1^o, x_1), \dots, (p_m^o, x_m)\}.$$

Ансамбль  $A_k$  на шаге  $k$  является неизвестным. Поэтому вместо  $Q(T_k, A_k, S)$  может быть использован близкий ему функционал

$$(6) \quad Q(T_k, A_{k-1}, S) = L(T, S) - \mu D(T_k, A_{k-1}, S).$$

Нетрудно показать, что минимум функционала (6) достигается при

$$(7) \quad p_j^o = \frac{y_j - \mu \hat{A}_{k-1}(x_j)}{1 - \mu}.$$

К сожалению, использование (7) приводит к нарушению условия  $p_j^o \in [0, 1]$ . Данное условие может быть сохранено при переходе к задаче условной оптимизации. Однако последнее приводит к существенному усложнению алгоритма. Более простым является решение, основанное на использовании (7) с последующим обучением нового дерева по выборке

$$(8) \quad \{(p_1^o, x_1), \dots, (p_m^o, x_j)\}.$$

Для разделения классов оценка величины  $p_j$  сравнивается с порогом, для поиска которого используются средства ROC (Receiver Operating Characteristic) анализа. Указанный метод далее будем называть дивергентным решающим лесом (Divergent Decision Forest, DDF).



#### 4. Исследование эффективности ансамблей с увеличенной дивергенцией при построении рекомендательных систем

Предложенный в предыдущих разделах метод зависит от набора гиперпараметров, в которые входят как множитель  $\mu$ , так и гиперпараметры, используемые при построении отдельных деревьев. Успешное применение большинства методов машинного обучения зависит от выбранных значений гиперпараметров. Вместе с тем точные теоретические оценки для выбора гиперпараметров обычно отсутствуют. Поэтому их оптимальные значения приходится искать через эксперименты с данными. Целью настоящей работы являлся экспериментальный поиск оптимальных значений гиперпараметров для дивергентного решающего леса. Исследования проводились в рамках решения задач прогнозирования пользовательских предпочтений, возникающих при создании рекомендательных систем. Выбор данного класса задач связан как с широким распространением рекомендательных систем в различных отраслях экономики, так и с интенсивным использованием в этой области методов машинного обучения [12]. Рассматривались три задачи оценивания предпочтений интернет-пользователей при выборе компьютерной игры и две задачи оценивания предпочтений интернет-пользователей при выборе стикера в социальных сетях. Оценивание проводилось на основе информации о взаимодействии пользователей с соответствующими каталогами. Для оценки использовался показатель HR5 (hitrate at 5) – доля пользователей, для которых среди первых пяти рекомендаций оказался хотя бы один релевантный объект. Используются следующие обозначения:  $k$  – число пользователей,  $p$  – число объектов,  $n$  – число признаков,  $M$  – число строк в обучающей выборке. Характеристики рассмотренных задач представлены в табл. 1.

**Таблица 1.** Характеристики рассмотренных задач

| Задача   | $k$  | $p$ | $n$ | $M$    |
|----------|------|-----|-----|--------|
| game0    | 414  | 187 | 41  | 4132   |
| game1    | 1706 | 190 | 41  | 16 327 |
| game2    | 3888 | 398 | 55  | 29 065 |
| sticker1 | 2685 | 118 | 40  | 27 032 |
| sticker2 | 5942 | 197 | 41  | 44 613 |

Целью работы было изучение зависимости точности алгоритма в смысле показателя HR5 от глубины деревьев и параметра  $\mu$ . Результаты для задач game0 и sticker1 представлены в табл. 2 и 3 соответственно. Значение  $\mu = 0$  соответствует обычной модели случайного решающего леса (RF).

Из таблиц видно, что при глубине деревьев, не превышающей 7, HR(5) возрастает с ростом  $\mu$ . Вместе с тем при больших значениях  $\mu$  при глубине деревьев более 7 показатель HR(5) снижается.

В табл. 4 представлено сравнение предлагаемой технологии со стандартными случайными решающими лесами, а также с методом CatBoost. Стандарт-

**Таблица 2.** Связь показателя  $HR(5)$  с глубиной деревьев и параметром  $\mu$  для задачи `game0`

|                   | $\mu = 0,0$  | $\mu = 0,25$ | $\mu = 0,5$  | $\mu = 0,75$ | $\mu = 0,9$  |
|-------------------|--------------|--------------|--------------|--------------|--------------|
| <i>depth</i> = 3  | 0,584        | 0,594        | 0,599        | 0,613        | 0,640        |
| <i>depth</i> = 5  | 0,596        | 0,604        | <b>0,623</b> | <b>0,630</b> | <b>0,657</b> |
| <i>depth</i> = 7  | 0,591        | 0,606        | 0,621        | 0,621        | 0,623        |
| <i>depth</i> = 9  | 0,592        | <b>0,618</b> | 0,606        | 0,606        | 0,570        |
| <i>depth</i> = 11 | <b>0,606</b> | 0,601        | 0,606        | 0,611        | 0,493        |

**Таблица 3.** Связь показателя  $HR(5)$  с глубиной деревьев и параметром  $\mu$  для задачи `sticker1`

|                   | $\mu = 0,0$  | $\mu = 0,25$ | $\mu = 0,5$  | $\mu = 0,75$ | $\mu = 0,9$  |
|-------------------|--------------|--------------|--------------|--------------|--------------|
| <i>depth</i> = 3  | 0,447        | 0,449        | 0,448        | 0,458        | 0,514        |
| <i>depth</i> = 5  | 0,482        | 0,488        | 0,498        | <b>0,516</b> | <b>0,527</b> |
| <i>depth</i> = 7  | 0,475        | 0,491        | <b>0,501</b> | 0,508        | 0,513        |
| <i>depth</i> = 9  | 0,479        | 0,492        | 0,500        | 0,507        | 0,477        |
| <i>depth</i> = 11 | <b>0,482</b> | <b>0,494</b> | 0,499        | 0,494        | 0,467        |

**Таблица 4.** Лучшее значение  $HR(5)$  для каждой из моделей

|                        | RF    | DDF ( $\mu \neq 0$ ) | catboost     |
|------------------------|-------|----------------------|--------------|
| <code>games0</code>    | 0,606 | 0,657                | <b>0,664</b> |
| <code>games1</code>    | 0,642 | <b>0,654</b>         | 0,637        |
| <code>games2</code>    | 0,513 | 0,550                | <b>0,562</b> |
| <code>stickers1</code> | 0,482 | <b>0,527</b>         | 0,513        |
| <code>stickers2</code> | 0,337 | 0,372                | <b>0,385</b> |
| среднее                | 0,516 | 0,552                | 0,5522       |

ным решающим лесам соответствует  $\mu = 0$ . Из таблицы видно, что точность в смысле метрики  $HR5$  для предлагаемой технологии заметно превышает точность случайного решающего леса и близка к точности метода CatBoost.

## 5. Заключение

Разработан метод дивергентного решающего леса, предназначенный для решения задач бинарной классификации. Метод основан на развитии подходов, ранее разработанных для решения задач регрессии [9–11]. На примере задач, связанных с созданием рекомендательных систем, проведено исследование метода дивергентного решающего леса (DDF), основанного на достижении более высокой дивергенции в пространстве прогнозов по сравнению со стандартным случайным решающим лесом. Показана существенная зависимость эффективности от глубины деревьев и от коэффициента  $\mu$ , регулирующего относительные вклады составляющих, отвечающих за аппроксимацию целевой переменной и дивергенции ансамблей. При этом при небольшой

глубине деревьев точность алгоритма возрастает с ростом  $\mu$ . Такого эффекта не наблюдается при большой глубине деревьев. Однако в целом DDF выигрывает у стандартного случайного решающего леса, соответствующего  $\mu = 0$ . Эксперименты показали, что эффективность DDF приближается к широко известной бустинговой модели CatBoost. Метод DDF основан на последовательной генерации наборов деревьев. Поэтому его быстродействие близко к быстродействию обычных случайных лесов и вариантов градиентного бустинга. Одной из проблем метода является выбор оптимального значения гиперпараметра  $\mu$ . Данная проблема может быть решена с помощью известных средств выбора гиперпараметров. Однако проведенные эксперименты указывают, что более высокая эффективность достигается при высоких значениях гиперпараметра, т.е. при  $\mu$ , равном 0,8 или 0,9. В отличие от методов решающих и регрессионных деревьев, обладающих высокой прозрачностью и интерпретируемостью, методы, использующие большие ансамбли деревьев, эти свойства, к сожалению, утрачивают. Последнее касается и метода DDF. Данный недостаток может быть компенсирован с помощью различных известных методов достижения интерпретируемости. Для повышения прозрачности могут быть использованы методы интеллектуального анализа данных, а также статистический анализ.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Hastie T., Tibshirani R., Friedman J.* The Elements of Statistical Learning Data Mining, Inference, and Prediction. Springer Series in Statistics. New York: Springer, 2009.
2. *Breiman L.* Random Forests // Machine Learning. 2001. V. 45. No. 1. P. 5–32.
3. *Friedman J.* Stochastic gradient boosting // Comput. Statist. Data Anal. 2002. V. 38. No. 4. P. 367–378.
4. *Breiman L.* Bagging predictors // Machine Learning. 1996. No. 24. P. 123–140.
5. *Tin Kam Ho.* The random subspace method for constructing decision forests // IEEE Transact. Patt. Machine Intelligen. 1998. V. 20. No. 8. P. 832–844.
6. *Chen T., Guestrin C.* XGBoost: A Scalable Tree Boosting System // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13–17 / eds. Krishnapuram B., Shah M., Smola A.J., Aggarwal C.C., Shen D., Rastogi R. 2016. P. 785–794.
7. *Ke G., Meng Q., Finley T. et al.* LightGBM: A Highly Efficient Gradient Boosting Decision Tree // NIPS'17: Proc. 31st Int. Conf. Neural Inform. Proc. Syst. 2017. P. 3149–3157.
8. *Hancock J.T., Khoshgoftaar T.M.* CatBoost for big data: an interdisciplinary review // J. Big Data. 2020. V. 7. No. 94.
9. *Zhuravlev Yu.I., Senko O.V., Dokukin A.A., Kiselyova N.N., Saenko I.A.* Two-Level Regression Method Using Ensembles of Trees with Optimal Divergence // Dokl. Math. 2021. V. 103. P. 1–4.
10. *Докучкин А.А., Сенько О.В.* Новый двухуровневый метод машинного обучения для оценивания вещественных характеристик объектов // Изв. РАН. Теория и системы управления. 2023. No. 4. P. 17–24.

11. *Senko O.V., Dokukin A.A., Kiselyova N.N., et al.* New Two-Level Ensemble Method and Its Application to Chemical Compounds Properties Prediction // Lobachev. J. Math. 2023. V. 44. No. 1. P. 188–197.
12. *Roy D., Dutta M.* A systematic review and research perspective on recommender systems // J. Big Data. 2022. V. 9. No. 59.

*Статья представлена к публикации членом редколлегии А.А. Галляевым.*

Поступила в редакцию 29.11.2024

После доработки 09.01.2025

Принята к публикации 14.01.2025

# Оптимизация, системный анализ и исследование операций

© 2025 г. Д.Е. ШАПОШНИКОВ, канд. физ.-мат. наук (shaposhnikov@unn.ru)  
(Нижегородский государственный университет им. Н.И. Лобачевского)

## МИНИМАКСНЫЙ ПОДХОД ДЛЯ УЧЕТА КАЧЕСТВЕННЫХ ПРЕДПОЧТЕНИЙ ПРИ МНОГОКРИТЕРИАЛЬНОЙ ОЦЕНКЕ

Поддержка принятия решений в аналитических системах на основе использования больших данных предполагает формирование интегральных оценок объектов по всему множеству параметров или некоторому их подмножеству. В статье рассматривается проблема получения многокритериальной (многопараметрической) оценки объектов и подход, который предполагает использование весовых коэффициентов важности при наличии качественной и, возможно, неполной информации об относительной важности тех или иных частных критериев. Рассматривается фундаментальный принцип возможности различных количественных оценок относительной предпочтительности частных критериев для различных объектов популяции при сохранении системы предпочтений всего множества объектов. Используемый подход предполагает, что лицо, принимающее решение, формулирует качественную информацию об относительной предпочтительности частных критериев в виде необязательно полного графа предпочтений. Для каждого объекта весовые коэффициенты рассчитываются автоматически по принципу гарантированного результата путем решения оптимизационной задачи с использованием обобщенных логических критериев максимального риска и максимальной осторожности. Для частных случаев систем предпочтений приведены аналитические соотношения и алгоритмы расчета весовых коэффициентов. Данная методика обеспечивает использование дополнительной качественной информации о предпочтениях тех или иных критериев, получение численных значений весовых коэффициентов значимости и решение задачи многокритериальной оценки на основе принципа гарантированного результата.

*Ключевые слова:* многокритериальная оптимизация, многокритериальная оценка, интерактивная процедура, весовые коэффициенты важности, качественная информация и предпочтительности.

**DOI:** 10.31857/S0005231025040075, **EDN:** SAMJHV

### 1. Введение

Задачи многокритериальной оценки широко распространены в различных сферах деятельности и в различных предметных областях. Под проблемой оценки в данном случае понимается необходимость получения интегральной

оценки объектов совокупности для последующего окончательного выбора лицом, принимающим решение. Объективная сложность таких задач заключается в невозможности достижения наилучших значений по всем критериям одновременно и в выборе компромиссного решения, т.е. решения, которое невозможно улучшить ни по одному из критериев, не ухудшив значения остальных. Одним из наиболее распространенных подходов к решению таких задач является формирование численной оценки значения для каждого допустимого решения и, следовательно, решение однокритериальной задачи оптимизации или, в случае дискретного множества допустимых решений, задачи выбора [1].

Известны подходы к задаче многокритериального выбора, учитывающие качественную информацию о предпочтениях лиц, принимающих решения (ЛПР) [2]. Основной особенностью данных методов является требование к ЛПР вводить достаточно большое количество необходимых параметров (точных числовых значений), отсутствие (хотя бы частичное) которых делает применимость этих методов невозможной. Описанная методика предполагает использование ровно того объема информации, который лицо, принимающее решение, готово предоставить (в конкретном случае – отсутствие таковой вообще) [3].

Одним из наиболее распространенных методов решения многокритериальных задач является использование обобщенного критерия оптимальности, включающего веса важности, отражающие представление лица, принимающего решение, об относительной важности тех или иных критериев оптимальности [2]. Важность критериев здесь понимается в смысле аксиоматической теории важности, которая позволяет предположить, что если имеется дополнительная информация вида “ $i$ -й критерий предпочтительнее  $j$ -го критерия ( $Q_i \succ Q_j$ )”, то для весовых коэффициентов должно выполняться соотношение  $w_i > w_j$ . При этом, конечно, лицу, принимающему решение, требуются точные численные значения весовых коэффициентов.

Весовые коэффициенты относительной важности частных критериев могут быть назначены лицом, принимающим решение, различными способами [4]. Все известные методы индивидуальной экспертизы требуют от лица, принимающего решение (или эксперта), полных и точных ответов (чаще всего с численными оценками) на вопросы, которые задает выбранный метод. Описанный подход позволяет использовать неполную и/или качественную информацию о предпочтениях тех или иных критериев.

## 2. Постановка задачи многокритериальной оценки

В общем виде задачу многокритериального выбора можно сформулировать следующим образом.

ЛПР сформировал набор некоторого числа допустимых альтернатив, из которых ему (ЛПР) необходимо сделать единственный и окончательный выбор. Выбор осуществляется как отбор некоторого подмножества альтернатив

(в частном случае одной альтернативы) таких, что ЛППР считает их наилучшими и равноценными друг другу по наилучшему качеству.

Также на практике рассматривается и решается задача многокритериальной оценки вариантов, в результате которой ЛППР стремится упорядочить варианты по предпочтению для дальнейшего анализа. Очевидно, что данная задача (формирование итоговой оценки) является обобщением предыдущей задачи (задачи выбора) и не противоречит практике.

Традиционно рассматриваются следующие компоненты модели принятия решений:

- исходный набор вариантов (решений, альтернатив), из которых осуществляется выбор;
- принцип оптимальности, на основе которого выбирается наилучший вариант (или наилучшие варианты);
- лицо, принимающее решение (ЛППР), определяющее процесс поиска решения, а также эксперты и консультанты, которые ему в этом помогают.

Не ограничивая общности математической формулировки, будем считать, что множество допустимых решений представляет собой дискретный конечный набор альтернатив, каждая из которых обозначается числом

$$(1) \quad D = \{1, \dots, m\}.$$

Задача выбора заключается в выборе варианта  $x^* \in D = \{1, \dots, m\}$ , который является наилучшим (оптимальным) с точки зрения лица, принимающего решение.

Поскольку в большинстве практических случаев в задачах многокритериального выбора оптимального решения не существует, применим термин «рациональное решение», который означает наличие понятных другим людям рациональных причин, приведших к выбору данного решения из множества допустимых.

Многие подходы к решению задач многокритериального выбора предполагают использование задачи оценки [3]. Задача оценки заключается в определении для каждой альтернативы числового значения, характеризующего качество (полезность, эффективность и т.д.) этого варианта с точки зрения предметной области и с точки зрения относительной важности частных критериев в соответствии с индивидуальным мнением ЛППР. При этом для сопоставимости альтернатив эта обобщенная характеристика должна иметь направление и масштаб.

В дальнейшем будем считать, что шкалы измерения частных критериев оптимальности определены и численное значение оценки по варианту может быть получено для каждого варианта и каждого критерия.

Также будем считать, что все частные критерии приведены к безразмерному виду и к единой шкале измерений  $[\alpha, \beta]$ , при этом  $0 \leq \alpha < \beta$ .

Общая идея оценки заключается в построении скалярной функции  $F(Q(x))$ , обладающей свойством упорядочивать исходные варианты в

соответствии с их предпочтением. Процедура построения такой функции называется сверткой или объединением векторного критерия оптимальности (как вариант сверткой или объединением частных критериев):

$$(2) \quad \min_{x \in D} Q_1(x), \min_{x \in D} Q_2(x), \dots, \min_{x \in D} Q_n(x),$$

где

$$(3) \quad D = \{1, \dots, m\},$$

преобразовывается в однокритериальную (скалярную) задачу

$$(4) \quad \min_{x \in D} \{F(Q_1(x), \dots, Q_n(x))\}.$$

Функция  $F(\cdot)$  называется обобщенным критерием оптимальности задачи (2). Для учета относительной важности частных критериев используются веса относительной важности  $w = (w_1, \dots, w_n)$ . В качестве обобщенного критерия могут использоваться различные функции, в частности обобщенные логические критерии оптимальности:

критерий «максимальной осторожности»

$$(5) \quad F_{L_{\min}}(w, Q_1(x), \dots, Q_n(x)) = \min_{1 \leq i \leq n} (w_i Q_i(x)),$$

критерий «максимального риска»

$$(6) \quad F_{L_{\max}}(w, Q_1(x), \dots, Q_n(x)) = \max_{1 \leq i \leq n} (w_i Q_i(x)).$$

Весовые коэффициенты  $w_j, j = 1, \dots, n$ , отражают точку зрения ЛПР на относительную важность частных критериев оптимальности. Важность критериев здесь понимается в смысле аксиоматической теории важности, которая позволяет предположить, что если дополнительная информация вида « $i$ -й критерий не менее важен, чем  $j$ -й критерий ( $Q_i \succeq Q_j$ )», то для весовых коэффициентов и справедливо следующее соотношение:

$$(7) \quad Q_i \succeq Q_j \Leftrightarrow w_i \geq w_j.$$

Опять же для решения исходной задачи выбора требуется точное задание весовых коэффициентов важности в числовой форме. При этом обычно предполагается, что весовые коэффициенты важности принадлежат области допустимых значений следующего вида (далее будем называть ее областью допустимых значений весовых коэффициентов):

$$(8) \quad D_w^0 = \left\{ w \in R^n \mid w_i \geq 0, i = 1, \dots, n; \sum_{i=1}^n w_i = 1 \right\}.$$



Известно, что выбор весовых коэффициентов важности частных критериев в строгом соответствии с (8) может привести к определенной проблеме: если для некоторого частного критерия  $Q_i$  значение соответствующего весового коэффициента важности равно нулю (что не противоречит построению области  $D_w$ ), то этот критерий, по сути, будет вычеркнут из рассмотрения и не будет иметь значения его численная величина. Математически это приведет к слабой эффективности получаемых решений (доказательство в [5]).

Кроме того, известно, что сумма весов важности может быть любым положительным числом (например, равной 100, если удобно работать в процентной терминологии).

Таким образом, в итоге приходим к следующей формулировке области допустимых значений весовых коэффициентов важности  $w$ :

$$(9) \quad D_w^1 = \left\{ w \in R^n \mid w_i \geq w_0 > 0, \quad i = 1, 2, \dots, n; \quad \sum_{i=1}^n w_i = R \right\}.$$

Для решения исходной задачи выбора лицу, принимающему решение, необходимо назначить точные значения весов важности в числовой форме из области допустимых значений.

Подходы и методы назначения весовых коэффициентов широко описаны в литературе. Среди них следующие [6]:

- упорядочение критериев по важности;
- определение соотношения весовых коэффициентов, при этом лицо, принимающее решение, указывает соотношение  $w_i/w_j$  в числовой форме;
- построение таблиц на основе попарного сравнения критериев по важности;
- метод определения весов с помощью набора последовательных сравнений (метод Черчмена–Акоффа);
- методы, использующие информацию о качестве оптимальных значений частных критериев;
- теоретико-игровые методы назначения весовых коэффициентов и др.

Эти методы имеют ряд существенных недостатков в практическом применении.

1. Утверждение (7) подразумевает «переход» от качественной информации ( $Q_i \succeq Q_j$ ) к количественной информации ( $w_i \geq w_j$ ). Очевидно, что это большая неопределенность, которая в конечном итоге приведет к разным решениям.

2. Все перечисленные методы назначения численных значений весовых коэффициентов важности являются последовательными алгоритмами, требующими от лица, принимающего решение, точной количественной информации на каждом шаге этого алгоритма. Как только лицо, принимающее решение, отказывается отвечать на очередной вопрос, весь алгоритм завершается и численные значения  $w$  не вычисляются.

### 3. Анализ и использование качественной информации об относительной важности конкретных критериев оптимальности – минимаксный подход

В ряде работ [6, 7] сформулированы основная идея и концепция данного подхода, которые заключаются в следующем.

1. Предполагается, что при использовании обобщенных критериев оптимальности с весовыми коэффициентами важности их значения могут быть разными для каждого допустимого решения области  $D$ .

2. Лицо, принимающее решение, может формулировать предпочтения на множестве частных критериев в виде набора отношений, не обязательно для всех возможных пар частных критериев.

3. Значения весов важности вычисляются автоматически на основе принципа гарантированного результата, т.е. для каждого допустимого решения определяется такой набор  $w \in D_w$ , который обеспечивает «наихудшее» значение обобщенного критерия для всех возможных значений весовых коэффициентов.

Рассмотрим более подробно частный, но часто встречающийся случай, когда лицо, принимающее решение, устанавливает для некоторых  $L$  пар частных критериев (не обязательно для всех  $C_n^2$  возможных пар) взаимное предпочтение  $i$ -го частного критерия над  $j$ -м на всем множестве  $D$  допустимых решений:

$$(10) \quad e_l = \{Q_i \succeq Q_j\}, \quad l = 1, \dots, L \leq \frac{n(n-1)}{2}.$$

Информация (10) является качественной, так как из нее следует, что критерий  $i$  важнее критерия  $j$ , но невозможно сказать, насколько или во сколько раз. Тогда, принимая во внимание соотношение (10), диапазон допустимых значений весовых коэффициентов важности  $D_w$ , определяемый как (8), сужается согласно соотношениям, введенным лицом, принимающим решения:

$$(11) \quad D_w^2 = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 > 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq w_j \rangle_{e_l}, \quad l = 1, \dots, L \end{array} \right. \right\}.$$

В некоторых случаях лицо, принимающее решение, имеет возможность уточнить информацию о взаимосвязи весовых коэффициентов  $w_i$  и  $w_j$ , связанных с бинарным отношением  $\{Q_i \succeq Q_j\}$ , используя параметр  $g_l \geq 1$ :

$$(12) \quad D_w^3 = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 > 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq g_l w_j \rangle_{e_l}, \quad l = 1, \dots, L \end{array} \right. \right\}.$$

В ряде работ [7–10] был предложен и развит подход, в котором весовые коэффициенты важности частных критериев являются неконтролируемыми факторами и их значения могут быть разными для разных альтернатив.

Иногда в процессе принятия решений целесообразно учитывать зависимость весовых коэффициентов от значения частного критерия в каждой точке области допустимых решений. Предполагая, что ЛПР не может точно определить численные значения весовых коэффициентов  $w$ , можно рассматривать их как неконтролируемые факторы и, применяя принцип гарантированного результата, перейти к следующей модели принятия решений:

$$(13) \quad x^* = \arg \min_{x \in D} \max_{w \in D_w^3} F(q(x), w),$$

где

$$(14) \quad D = \{1, \dots, N_D\},$$

$$(15) \quad D_w^3 = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 \geq 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq g_l w_j \rangle_{e_l}, \quad l = 1, \dots, L \end{array} \right. \right\}.$$

#### 4. Представление качественной информации в виде графа

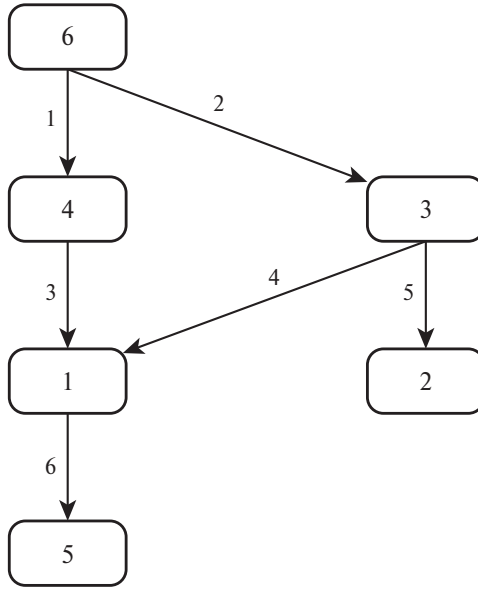
Качественную информацию о предпочтениях эксперта можно представить в виде ориентированного графа  $G(V, E)$ , где  $V$  – множество вершин, соответствующих частным критериям,  $E$  – множество ребер, соединяющих  $i$ -ю вершину с  $j$ -ой вершиной тогда и только тогда, когда выполняется соотношение  $(Q_i \succeq Q_j)$ .

В дальнейшем будем предполагать, что информационные сообщения (7) удовлетворяют условию транзитивности, т.е. в графе  $G(V, E)$  нет замкнутых циклов. Разобьем множество  $I$  вершин графа на слои (ярусы) следующим образом:

- к первому слою ( $s = 1$ ) относятся все вершины, не включающие ни одной дуги;
- ко второму слою ( $s = 2$ ) – те и только те вершины, включающие дуги из вершин первого яруса;
- к третьему слою ( $s = 3$ ) – те и только те вершины, включающие дуги из вершин первого и второго ярусов и т.д.

В этом случае вершины последнего слоя ( $s = S \leq N$ ) не будут иметь ни одной исходящей дуги.

В частном случае отсутствия качественной информации (7) граф  $G(V, E)$  будет представлять собой совокупность  $N$  изолированных вершин. А в частном случае линейного упорядочения частных критериев граф  $G(V, E)$  будет иметь число слоев, равное числу вершин ( $S = n$ ).



Пример графа предпочтений.

В качестве примера рассмотрим граф предпочтений (рисунок), содержащий шесть вершин (критериев) и дополнительную качественную информацию, введенную лицом, принимающим решения:  $e_1 = \{Q_6 \succeq Q_4\}$ ,  $e_2 = \{Q_6 \succeq Q_3\}$ ,  $e_3 = \{Q_4 \succeq Q_1\}$ ,  $e_4 = \{Q_3 \succeq Q_1\}$ ,  $e_5 = \{Q_3 \succeq Q_2\}$ ,  $e_6 = \{Q_1 \succeq Q_5\}$ .

Для каждой вершины графа  $v_i \in \{1, \dots, n\}$ , соответствующей частному критерию  $Q_i$ , введем следующие обозначения:  $V_i$  – множество вершин графа  $G(V, E)$ , из которых есть путь в вершину  $i$ , включая ее саму;  $n_i$  – мощность множества  $V_i$ .

Пусть  $p$  – произвольный путь в графе  $G(V, E)$ :  $p = \{l_1, \dots, l_n(p)\}$ , где  $l_i \in E$  – дуга, входящая в путь  $p$ ;  $n(p)$  – число дуг в пути  $p$ . Обозначим через  $P_i^k$  множество всех путей из вершины  $i$  в вершину  $k$ . Затем введем величины  $\bar{g}_k^i$  и  $\bar{\bar{g}}_i$  следующим образом:

$$(16) \quad \bar{g}_k^i = \begin{cases} \max_{p \in P_i^k} \prod_{l \in p} g_l, & P_i^k \neq \emptyset, \\ 1, & P_i^k = \emptyset \end{cases}$$

для всех  $i, k \in E$ ;

$$(17) \quad \bar{\bar{g}}_i = \max_{k \in E, k \neq i} \bar{g}_k^i.$$

Очевидно, что  $\bar{\bar{g}}_i$  является обобщенной оценкой относительной важности частного критерия  $i$  в условиях качественной информации (10). Ниже приведен пример расчета этих величин для графа, представленного

на рисунке. Для всех дуг графа назначим коэффициенты уточнения  $g_l$ ,  $l = 1, \dots, 6$  (табл. 1).

**Таблица 1.** Пример уточняющих коэффициентов

| Номер дуги $l$ | Коэффициент $g_l$ |
|----------------|-------------------|
| 1              | 1,2               |
| 2              | 1,3               |
| 3              | 1,0               |
| 4              | 1,5               |
| 5              | 1,3               |
| 6              | 2,0               |

Расчет значений (16) и (17) для всех вершин графа приведен в табл. 2.

**Таблица 2.** Пример вычисления численных значений характеристик вершин

| Вершина $i$ | Значение $\bar{g}_i^k (P_i^k \neq \emptyset)$  | $\bar{g}_i$ |
|-------------|--|-------------|
| 1           | $\bar{g}_1^5 = 2,0$  | 2,0         |
| 2           | –  | 1,0         |
| 3           | $\bar{g}_3^5 = 2,0 \times 1,5 = 3,0$ ; $\bar{g}_3^2 = 1,3$   | 3,0         |
| 4           | $\bar{g}_4^5 = 1,5 \times 1,0 = 1,5$   | 1,5         |
| 5           | –  | 1,0         |
| 6           | $\bar{g}_6^5 = \max\{1,2 \times 1,0 \times 2,0; 1,3 \times 1,5 \times 2,0\} = 3,9$ ;<br>$\bar{g}_6^2 = 2,0 \times 1,3 = 2,6$ | 3,9         |

*Алгоритм 1.*

1. Полагаем  $s = S$ ;  $\bar{g}_i^k = 0$ ,  $i, k \in \{1, \dots, n\}$ .
2. Для всех вершин  $i$  слоя  $s$  полагаем  $\bar{g}_i = 1$ .
3. Полагаем  $s = s - 1$ . Если  $s = 0$  (т.е. все слои пройдены), переход на шаг 5, иначе переход на шаг 4.
4. Для каждой вершины  $k$ , принадлежащей слою  $s$ , производятся следующие вычисления. Пусть  $J_k$  является множеством вершин слоя  $(s + 1)$ , в которые есть дуга из вершины  $k$ . Тогда для каждой вершины  $j \in J_k$  полагаем:

$$\bar{g}_i^k = g_l, \quad \text{где } l \text{ – номер дуги } (k, j) \text{ в графе } G(V, E);$$

$$\bar{g}_k^i = \max_{j \in J_k} \left\{ \bar{g}_k^j; \bar{g}_j^i \right\}, \quad i \neq j;$$

$$\bar{g}_k = \max_{j \in J_k} \left\{ \bar{g}_k^j; \bar{g}_k \right\}$$

и переход к шагу 3.

5. Для всех  $\bar{g}_i^k$ , для которых  $\bar{g}_k = 0$ ,  $k, i \in \{1, \dots, n\}$ , полагаем  $\bar{g}_i^k = 1$ .

Конец алгоритма.

## 5. Вычисление весовых коэффициентов важности для обобщенного критерия максимальной осторожности

### 5.1. Общий случай

Рассмотрим решение задачи расчета весовых коэффициентов важности  $w \in D_w^4$  с использованием обобщенного логического критерия «максимальной осторожности»:

$$(18) \quad x^* = \arg \min_{x \in D} \max_{w \in D_w} \max_{1 \leq i \leq n} F(q(x), w).$$

$$(19) \quad D_w^1 = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 \geq 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq g_l w_j \rangle_{e_l}, \quad l = 1, \dots, L \end{array} \right. \right\}.$$

При фиксированных значениях переменных параметров  $x \in D$  значения векторного критерия  $q(x)$  являются определенными.

Применяя линейную модель для определения смешанных стратегий в матричной игре с нулевой суммой [12], в общем случае решение задачи (18)–(19) можно получить, решив следующую задачу линейной оптимизации:

$$(20) \quad \max v,$$

где

$$(21) \quad D_w = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 \geq 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq g_l w_j \rangle_{e_l}, \quad l = 1, \dots, L; \\ w_i q_i(x) \geq v, \quad i = 1, \dots, n \end{array} \right. \right\}.$$

Рассмотрим пример расчета весовых коэффициентов. Предположим, что в табл. 3 приведены значения частных критериев оптимальности для одного допустимого решения, нормированные на интервал  $[0, 1]$ . В этой же таблице приведены и рассчитанные весовые коэффициенты.

**Таблица 3.** Пример рассчитанных значений весовых коэффициентов важности для допустимого решения

| Частные критерии                 | $Q_1$  | $Q_2$  | $Q_3$  | $Q_4$  | $Q_5$  | $Q_6$  |
|----------------------------------|--------|--------|--------|--------|--------|--------|
| Значения частных критериев $q_i$ | 0,1    | 1,0    | 0,5    | 0,3    | 0,9    | 0,3    |
| Весовые коэффициенты $w_i$       | 0,1766 | 0,1767 | 0,2650 | 0,1767 | 0,0196 | 0,3445 |

Приведенное выше решение задачи расчета весовых коэффициентов важности частных критериев получено с использованием программного пакета LPSolv [13].

Приведенное выше решение получено путем решения линейной оптимизационной задачи симплекс-методом. Для частных случаев качественной информации о предпочтениях и, следовательно, области допустимых значений весовых коэффициентов важности расчет значений коэффициентов может быть осуществлен более быстрыми алгоритмами.

### 5.2. Вычисление весовых коэффициентов в частных случаях

Рассмотрим решение задачи расчета весовых коэффициентов важности  $w \in D_w^2$  (11) с использованием обобщенного логического критерия «максимальная осторожность»:

$$(22) \quad x^* = \arg \min_{x \in D} \max_{w \in D_w} F(q(x), w),$$

$$(23) \quad D_w^1 = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 \geq 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq w_j \rangle_{e_l}, \quad l = 1, \dots, L \end{array} \right. \right\}.$$

Для расчета значений весовых коэффициентов важности  $w$  в случае условия их неотрицательности, т.е. для случая  $w_0 = 0$ , можно применить алгоритм 2.

#### Алгоритм 2.

1. Граф  $G(V, E)$  разбивается на слои.
2. Каждой вершине  $j$  графа  $G(V, E)$ ,  $j = 1, \dots, n$ , присваиваются начальные значения  $w'_j = 0$ .
3. В качестве текущего слоя рассматриваем самый нижний слой:  $s = S$ .
4. Полагаем  $w'_j = \max \{w'_j, R/q_j(x)\}$  для всех вершин  $j$  слоя  $s$ .
5. Корректируются значения  $w'_j$  вершин вышележащих слоев:

$$w'_j = \max \left\{ w'_j, \max_{k \in V_j} w'_k \right\},$$

$V_j$  – множество вершин, из которых есть путь в вершину  $j$ .

6. Полагаем  $s = s - 1$ . Если  $s = 0$ , т.е., все вершины пройдены, переход к шагу 7. Иначе вычисления повторяются с шага 4.

7. Вычисляются значения весовых коэффициентов:

$$w_j^* = \frac{Rw'_j}{\sum_{i=1}^n w'_i}.$$

Конец алгоритма.

Рассмотрим обобщение экстремальной задачи (22)–(23) для случая  $w_0 \geq 0$ . Для решения этой задачи можно применить алгоритм 3.

*Алгоритм 3.*

1. Полагаем  $V' = V = \{1, \dots, n\}$ ;  $R' = R$ ;  $E' = E$ .

2. Применяем алгоритм 2 для графа  $G(V, E)$  с параметрами  $R = R'$  и получаем значения  $w'_j$ .

3. Если для всех  $j \in V'$  выполняется  $w'_j \geq w_0$ , тогда полагаем  $w_j^* = w'_j$ ,  $j \in V'$  и задача решена. Иначе переход на п. 4.

4. Для всех вершин  $m \in V'$ , для которых выполняется  $w'_m < w_0$ , производятся следующие действия:

– полагаем  $w'_m = w_0$ ;  $R' = R' - w_0$ ;

– исключаем вершину  $m$  из рассмотрения, для этого полагаем  $V' = V' \setminus m$  и исключаем из множества  $E'$  дуги, инцидентные вершине  $m$ , если таковые есть.

5. Вычисления повторяются с шага 1.

Доказательство корректности вышеприведенных алгоритмов содержится в [6].

## 6. Вычисление весовых коэффициентов важности для обобщенного критерия максимального риска

Рассмотрим решение задачи расчета весовых коэффициентов важности  $w \in D_w^4$  (6) с использованием обобщенного логического критерия «максимального риска»:

$$(24) \quad x^* = \arg \max_{w \in D_w} \max_{1 \leq i \leq n} F(q(x), w),$$

$$(25) \quad D_w = \left\{ w \in R^n \left| \begin{array}{l} w_i \geq w_0 \geq 0, \quad i = 1, \dots, n; \\ \sum_{i=1}^n w_i = R; \\ \langle w_i \geq g_l w_j \rangle_{e_l}, \quad l = 1, \dots, L \end{array} \right. \right\}.$$

При фиксированных значениях переменных параметров  $x \in D$  значения векторного критерия  $q(x)$  также определены.

Оптимальным решением задачи оптимизации (24)–(25) является вектор  $w^*$ , имеющий компоненты:

$$(26) \quad w_i^* = \begin{cases} \frac{(R' \bar{g}_i^r)}{\sum_{k \in V_r} \bar{g}_k^r} + \bar{g}_i w_0, & i \in V_r; \\ \bar{g}_i w_0, & i \notin V_r, \end{cases}$$

где

$$(27) \quad R' = R - w_0 \sum_{i=1}^n \bar{g}_i \geq 0,$$



значения  $\bar{g}_i^r$  и  $\bar{g}_i$  определяются из соотношений (16) и (17) соответственно;  $V_i$  – множество вершин графа  $G(V, E)$ , из которых есть путь в вершину  $i$ , включая саму вершину;  $n_i$  – мощность множества  $V_i$ , значение  $r$  получается из соотношений:

$$(28) \quad y_r = \max_{1 \leq k \leq n} y_k,$$

$$(29) \quad y_k = \max_{i \in V_k} \left( q_i(x) \left( \frac{(R' \bar{g}_i^r)}{\sum_{j \in V_r} \bar{g}_j^r} + \bar{g}_i w_0 \right) \right).$$

Доказательство формул (24)–(29) приведено в [11].

## 7. Заключение

Рассмотрена методика расчета весовых коэффициентов важности частных критериев при решении задач многокритериального выбора и оценки альтернатив. Фундаментальный принцип возможных различных значений весовых коэффициентов важности при постоянной системе качественных предпочтений позволяет лицу, принимающему решение, формулировать предпочтения в виде ориентированного графа, который в случае полноты представляет собой отношение строгого порядка. Данная методика может быть использована как при построении интерактивных систем принятия решений, так и для сведения многокритериальной задачи к однокритериальной для использования в различных методах оптимизации.

## СПИСОК ЛИТЕРАТУРЫ

1. Hillier F.S., Lieberman G.J. Introduction to operations research / 7th ed. McGraw-Hill Higher Education, 2001. ISBN 0-07-232169-5.
2. Steuer R.E. Multiple Criteria Optimization: Theory, Computations, and Application. New York: John Wiley Sons, 1986.
3. Батищев Д.И. Методы оптимального проектирования. М.: Радио и связь, 1984.
4. Matthias Ehrgott. Multicriteria Optimization. Springer, 2005. ISBN 3-540-21398-8.
5. Подиновский В.В., Ногин В.Д. Парето-оптимальные решения многокритериальных задач. М.: Наука, 1982.
6. Батищев Д.И., Шапошников Д.Е. Многокритериальный выбор с учетом индивидуальных предпочтений. Нижний Новгород: ИПФ РАН, 1994.
7. Batishchev D., Anuchin V., Shaposhnikov D. The Use of the Qualitative Information on the Importance of Particular Criteria for the Computation of Weighting Coefficients / Multiobjective Problems of Mathematical Programming. Lecture Notes in Economic and Mathematical Systems. V. 35. Springer-Verlag, 1991.

8. *Shaposhnikov D., Makarova J.* Multicriteria Problem of Wireless System Design Using Tricriteria Approach with Qualitative Information on the Decision Maker's Preference / Balandin D., Barkalov K., Gergel, V., Meyerov I. (eds). Mathematical Modeling and Supercomputer Technologies. MMST 2020. Communications in Computer and Information Science, V. 1413. Springer, Cham.
9. *Makarova J., Shaposhnikov, D.* Bicriterial problem of finding filter parameters using qualitative information on frequency interval preferences // Proceedings of the XXVI International Scientific and Technical Conference "Information Systems and Technology – 2020". NSTU, Nizhni Novgorod, 2020. P. 836–847.
10. *Shaposhnikov D., Makarova, J.* Multicriteria problem of calculation of wireless system parameters using qualitative information on the decision maker's preference // Proceedings of the XX International Conference "Mathematical modeling and supercomputer technologies", pp. 16–21. NNSU Publishing House, Nizhni Novgorod, 2020.
11. *Шапошников Д.Е., Костина И.В.* Применение обобщенного логического критерия для аппроксимации области эффективности в многокритериальных задачах оптимизации // Инженерный вестник Дона. 2014. № 4 / Электронный журнал: [ivdon.ru/ru/magazine/archive/n4y2014/2552](http://ivdon.ru/ru/magazine/archive/n4y2014/2552).
12. *Neumann J., Morgenstern O.* Theory of Games and Economic Behavior. Princeton, N.J. Princeton University Press, 1953.
13. <https://sourceforge.net/projects/lpsolve/>

*Статья представлена к публикации членом редколлегии А.А. Галляевым.*

Поступила в редакцию 29.11.2024

После доработки 11.01.2025

Принята к публикации 14.01.2025

---

---

## СОДЕРЖАНИЕ

### Линейные системы

- Никифоров В.О., Герасимов Д.Н., Дударенко Н.А.** Адаптивная компенсация внешних возмущений в многоканальных системах при управлении по выходу ..... 3

### Нелинейные системы

- Пашаев А.М., Искендеров А.Д., Мусаева М.А.** Задача оптимального управления коэффициентами нелинейного уравнения типа Шредингера ..... 22

### Стохастические системы

- Булычев Ю.Г.** Декомпозиционно-автокомпенсационное распознавание сигналов на базе принципов непрерывности, инвариантности, размножения и ранжирования с учетом регулярных и нерегулярных помех ..... 34
- Лебедев А.В.** Многомерные непрерывные распределения и копулы, порождающие нетранзитивные наборы зависимых случайных величин ..... 55

### Интеллектуальные системы управления, анализ данных

- Скворцов Н.А.** Управление качеством данных при решении задач в исследовательских инфраструктурах над неоднородными источниками данных ..... 71

### Оптимизация, системный анализ и исследование операций

- Сенько О.В., Докукин А.А., Мельник Ф.А.** Использование ансамблей с увеличенной дивергенцией в пространстве прогнозов в рекомендательных системах ..... 92
- Шапошников Д.Е.** Минимаксный подход для учета качественных предпочтений при многокритериальной оценке ..... 101

## C O N T E N T S

### Linear Systems

- Nikiforov V.O., Gerasimov D.N., Dudarenko N.A.** Adaptive Compensation of External Disturbances for MIMO Systems with Output Control . . . . . 3

### Nonlinear Systems

- Pashaev A.M., Iskenderov A.D., Musaeva M.A.** Problem of Optimal Control by Coefficients of the Nonlinear Schroedinger Type Equation . . . . . 22

### Stochastic Systems

- Bulychev Yu.G.** Decomposition-Autocompensation Signal Identification Based on the Continuity, Invariance, Multiplication, and Ranking Principles with Consideration of Regular and Irregular Noise . . . . . 34
- Lebedev A.V.** Multivariate Continuous Distributions and Copulas Generating Nontransitive Tuples of Dependent Random Variables . . . . . 55

### Intellectual Control Systems, Data Analysis

- Skvortsov N.A.** Data Quality Management in Problem Solving Using Research Infrastructures over Heterogeneous Data Sources . . . . . 71

### Optimization, System Analysis, and Operations Research

- Senko O.V., Dokukin A.A., Melnik F.A.** Using Ensembles with Enhanced Divergence in Forecast Space in Recommender Systems . . . . . 92
- Shaposhnikov D.E.** Minimax Approach for Using the Qualitative Preferences in the Multicriteria Evaluation . . . . . 101