

КИБЕРНЕТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

CYBERNETICS AND COMPUTER ENGINEERING

Выпуск 182

НАУЧНЫЙ ЖУРНАЛ

Основан в 1965 г.

Периодичность: 4 раза в год

КИЕВ 2015

СОДЕРЖАНИЕ

Информатика и информационные технологии

- Суровцев И.В.* Метод цифровой фильтрации электрохимических сигналов в хронопотенциометрии 4
- Баловсяк С.В., Фодчук И.М., Соловей Ю.Н., Луцик Я.В.* Многоуровневый метод повышения локального контраста и удаления неоднородного фона изображений 15
- Павлов А.В.* Подход к организации рекуррентно-параллельных вычислений в системе автоматизированной структурно-параметрической идентификации 27

Интеллектуальное управление и системы

- Копец М.М.* Оптимальное управление колебаниями балки переменного поперечного сечения 39
- Аралова Н.А.* Математическая модель краткосрочной и среднесрочной адаптации системы дыхания лиц, работающих в экстремальных условиях высокогорья 48

Медицинская и биологическая кибернетика

- Антомонов М.Ю.* Метод определения эффектов совместного действия факторов среды в эпидемиологических исследованиях для бинарных переменных 59
- Кривова О.А., Козак Л.М.* Комплексная оценка регионального демографического развития 70
- Настенко Е.А., Бойко А.Л., Павлов В.А., Тепляков К.И.* Синтез логистической регрессии на принципах самоорганизации моделей 85

Авторы номера 94

Системный указатель статей, опубликованных в 2015 году 95

КИБЕРНЕТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

CYBERNETICS AND COMPUTER ENGINEERING

Issue 182

SCIENTIFIC JOURNAL

Founded in 1965

Periodicity: 4 issues per year

KIEV 2015

TABLE OF CONTENTS

Informatics and Information Technologies

<i>Surovtsev I.V.</i> The Method of Digital Filtering of Electrochemical Signals in the Chronopotentiometry	4
<i>Balovsyak S.V., Fodchuk I.M., Solovay Yu.M., Lutsyk Ia.V.</i> Multilevel Method of Local Contrast Increase and Images Heterogeneous Background Removal	15
<i>Pavlov A.V.</i> Approach to Organization of Recurrent-And-Parallel Computations in Automized Structure-Parametric Identification System	27

Intellectual Control and Systems

<i>Kopets M.M.</i> Optimal Control by Vibrations of the Beam With Variable Cross-Section	39
<i>Aralova N.I.</i> Mathematical Model of the Short- And Medium-Term Adaptation of Respiratory System of the Persons Working in Extreme Conditions of High Mountains	48

Medical and Biological Cybernetics

<i>Antomonov M.Yu.</i> Methods of Determination of Environmental Factors Joint Impact in Epidemiological Studies for Binary Data	59
<i>Krivova O.A., Kozak L.M.</i> Complex Estimation of Regional Demographic Development	70
<i>Nastenko E.A., Boyko A.L., Nosovets O.K., Teplyakov K.I., Pavlov V.A.</i> Synthesis of Logistic Regression, Based on Self-Organisation Principles of Models	85

Our authors	94
--------------------------	----

System index of articles published in 2015	95
---	----

КІБЕРНЕТИКА ТА ОБЧИСЛЮВАЛЬНА ТЕХНІКА

CYBERNETICS AND COMPUTER ENGINEERING

Випуск 182

НАУКОВИЙ ЖУРНАЛ
Заснований у 1965 р.
Періодичність: 4 рази на рік

КИЇВ 2015

ЗМІСТ

Інформатика та інформаційні технології

<i>Суровцев І.В.</i> Метод цифрової фільтрації електрохімічних сигналів у хронопотенціометрії	4
<i>Баловсяк С.В., Фодчук І.М., Соловей Ю.М., Луцук Я.В.</i> Багаторівневий метод підвищення локального контрасту і видалення неоднорідного фону зображень	15
<i>Павлов А.В.</i> Підхід до організації рекурентно-паралельних обчислень в системі автоматизованої структурно-параметричної ідентифікації	27

Інтелектуальне керування та системи

<i>Копець М.М.</i> Оптиміальне керування коливаннями бруса змінного поперечного перерізу	39
<i>Арлова Н.І.</i> Математична модель короткострокової та середньострокової адаптації системи дихання осіб, які працюють в екстремальних умовах високогір'я	48

Медична та біологічна кібернетика

<i>Антомонов М.Ю.</i> Метод визначення ефектів сумісної дії факторів навколишнього середовища в епідеміологічних дослідженнях для бінарних змінних	59
<i>Кривова О.А., Козак Л.М.</i> Комплексна оцінка регіонального демографічного розвитку	70
<i>Настенко Е.А., Бойко А.Л., Павлов В.А., Тепляков К.І.</i> Синтез логістичної регресії на принципах самоорганізації моделей	85

Інформація про авторів	94
-------------------------------------	----

Системний показник статей, опублікованих у 2015 році	95
---	----

Информатика и информационные технологии

УДК 004.021:004.67

МЕТОД ЦИФРОВОЙ ФИЛЬТРАЦИИ ЭЛЕКТРОХИМИЧЕСКИХ СИГНАЛОВ В ХРОНОПОТЕНЦИОМЕТРИИ

И.В. Суровцев

Международный научно-учебный центр информационных технологий и систем НАН Украины и МОН Украины

Рассмотрен метод цифровой фильтрации, основанный на использовании способа определения спектра аналогового сигнала по точкам экстремумов. При его использовании в хронопотенциометрии показано, что он позволяет выполнить высокочастотную обработку разных видов электрохимических сигналов, не искажая форму полезного сигнала и не приводя к размыванию границ диапазона изменения компонентов. Метод прост в реализации, обладает большой скоростью фильтрации и обеспечивает высокую надежность в технических системах измерения сигналов различной природы.

Ключевые слова: цифровая фильтрация, точки экстремумов сигнала, хронопотенциометрия.

Розглянуто метод цифрової фільтрації, заснований на використанні способу визначення спектру аналогового сигналу по точках екстремумів. При його використанні в хронопотенціометрії показано, що він дозволяє виконувати високочастотне оброблення різних видів електрохімічних сигналів, не спотворюючи форми корисного сигналу та не призводячи до розмиву меж діапазонів зміни компонентів. Метод простий у застосуванні, має велику швидкість фільтрації та забезпечує високу надійність у технічних системах вимірювання сигналів різної природи.

Ключові слова: цифрова фільтрація, точки екстремумів сигналу, хронопотенціометрія.

ВВЕДЕНИЕ

В технических системах электрохимического определения концентрации токсичных элементов методами инверсионной хронопотенциометрии полезный сигнал, поступающий от измерительного и вспомогательного электрода, представляет собой интегральную сумму последовательного измерения во времени монотонно возрастающих потенциалов инверсии отдельных элементов после их предварительного накопления [1].

Цифровая обработка электрохимических сигналов по созданной информационной технологии определения концентрации предусматривает выполнение целенаправленного преобразования структуры данных измерения. Преобразование состоит из цифровой фильтрации исходного

сигнала, получения обратного сигнала и его дифференцирования. Полученный сигнал интенсивности рассматривают как линейную сумму компонентов измерения, при этом площадь компонента представляет собой значение времени инверсии, которое используется при определении концентрации элемента. Точность расчета концентрации напрямую связана с точностью определения значения времени инверсии и зависит от качества моделирования сигнала интенсивности, а также от правильности выбора диапазонов изменения каждого компонента [2].

Анализ исходных сигналов показывает, что создание метода цифровой фильтрации, который способен выполнить предварительную обработку всех видов электрохимических сигналов на фоне разнообразных высокочастотных помех при сохранении формы полезного и преобразованного сигнала, является довольно сложной и многоплановой задачей.

ПОСТАНОВКА ЗАДАЧИ

Прежде чем перейти к решению задачи, рассмотрим наиболее характерные виды сигналов при выполнении хронопотенциометрических методов анализа.

Представим на рис. 1 фрагменты сигналов инверсии при измерении концентрации четырех разных токсичных элементов с одинаковой длительностью в 0,15 с.

Необходимо отметить, что видов сигналов инверсии гораздо больше, и в продолжительных по времени сигналах один вид может плавно переходить в другой. Вид сигнала не связан с химическим элементом, а зависит от продолжительности (от 0,05 до 800 с.) и условий проведения измерения. Это, прежде всего, концентрация и диапазон инверсии анализируемого элемента, химический состав фонового раствора, тип и техническое состояние используемых электродов, а также другие факторы, трудно поддающиеся формализации.

В общем виде задача предварительной обработки связана с выделением полезного сигнала $f_0(t)$ по реализации $f(t) = f_0(t) + \zeta(t)$ на фоне аддитивной помехи $\zeta(t)$. Можно предположить, что помехи $\zeta(t)$ могут быть разных типов, одновременно присутствующие и независимо искажающие полезный сигнал [3–5]. Поэтому рассмотрим их как сумму аддитивных случайных помех

$$\zeta(t) = \zeta_1(t) + \zeta_2(t) + \dots + \zeta_k(t) \quad (1)$$

и построим цифровой фильтр в виде последовательности k фильтров.

К разным типам случайных помех можно отнести:

ζ_1 — ошибки квантования аналого-цифрового преобразования сигнала;

ζ_2 — ошибки округления вычислений;

ζ_3 — шум устройства измерения или внешние воздействия («наводки»);

ζ_4 — гармонические сетевые помехи на частоте 50 Гц;

ζ_5 — импульсные одиночные помехи малой длительности («выбросы»);

ζ_6 — значительные электромагнитные помехи (молния, работающий электродвигатель, мобильный телефон).

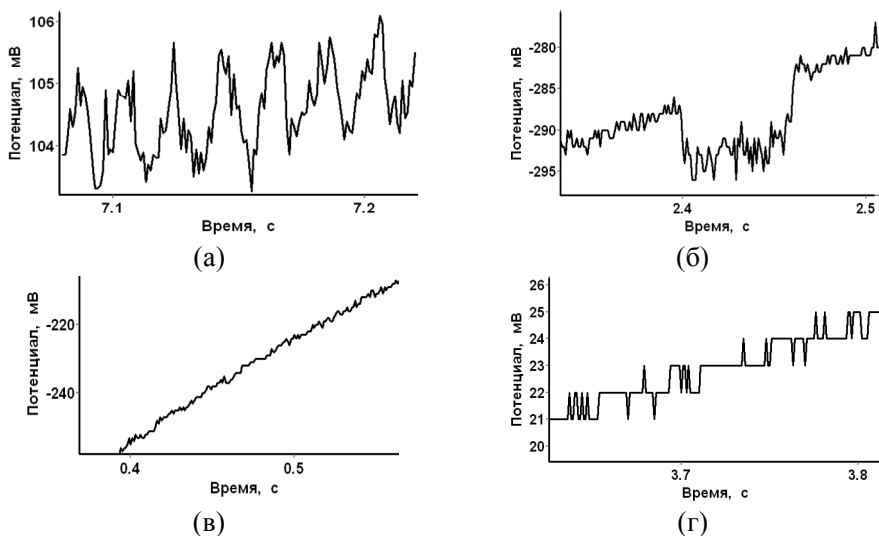


Рис. 1. Фрагменты измерения сигнала инверсии: а — ионов мышьяка, б — ионов свинца, в — ионов меди, г — ионов ртути

При создании реальных систем электрохимических измерений случайные помехи ($\zeta_1 + \zeta_2 + \zeta_3$) объединяют и рассматривают как «белый» шум [3, 5] с равномерным распределением, нулевым математическим ожиданием и ограниченной величиной дисперсии. «Белый» шум и гармонические сетевые помехи ζ_4 устраняют разнообразными методами цифровой фильтрации. Одиочные «выбросы» ζ_5 обрабатывают специальными программами, которые учитывают свойства и характер измеряемых сигналов. Значительные электромагнитные помехи ζ_6 являются неустраняемыми, их ограничивают на системном уровне при организации самого процесса анализа.

Рассмотрим возможность использования известных методов цифровой фильтрации для обработки указанных на рис. 1 видов сигналов инверсии.

На рис. 1а показан высокочастотный «белый» шум и гармонические сетевые помехи. Для цифровой фильтрации таких сигналов используют различные фильтры, например, широко известный фильтр нижних частот Баттерворта [3] с максимально плоской амплитудно-частотной характеристикой в полосе пропускания. Однако практическое его использование ограничено, поскольку он не обеспечивает полное подавление частот в полосе заграждения и искажает сигнал в полосе пропускания [4].

На рис. 1б сигнал содержит одиочный «выброс», который обрабатывается отдельно, после устранения «белого» шума и гармонических сетевых помех.

На рис. 1в сигнал может быть обработан одним из известных методов сглаживания. Например, с использованием метода экспоненциального сглаживания, предложенного Р. Брауном; модификацией метода скользящего среднего (алгоритм адаптивного сглаживания [4]); с помощью эрмитовых сплайнов и метода наименьших квадратов [6] и других. Однако использование всех этих методов приводит к искажению формы полезного

сигнала и к размыванию границ диапазона изменения компонентов на сигнале интенсивности, что существенно увеличивает ошибку определения концентрации элемента.

На рис. 1г для обработки сигнала можно использовать метод группировки данных, по которому исходный сигнал разбивается на равные участки во времени и в них определяются средние значения амплитуды [6]. Далее точки со средними значениями интерполируются по всему исходному сигналу. Однако такой подход применим только для простой формы сигнала и для него также присущи указанные выше ограничения. Кроме того, довольно сложно определить длину участка во времени для сигналов разной продолжительности.

Из сказанного следует, что трудно выбрать один какой-то существующий метод цифровой фильтрации для предварительной обработки одновременно всех видов электрохимических сигналов.

Как показали исследования, наиболее перспективным для выполнения высокочастотной предварительной обработки оказалось применение подхода, использованного при создании способа определения спектра аналогового сигнала [7]. Суть его заключается в использовании точек экстремумов для спектрального и корреляционного анализа, а также для цифровой фильтрации сигнала [8]. Основным элементом при построении спектра является значение временного интервала между соседними экстремумами и амплитудное приращение, соответствующее полуволне частоты. Главным преимуществом такого подхода является чрезвычайная простота программной реализации, ее высокая скорость и эффективность.

Использование точек экстремумов сигнала для выполнения цифровой фильтрации позволило преодолеть описанные трудности. Многократное и успешное использование такого подхода к задачам фильтрации и сглаживания не утратило актуальности, несмотря на появление новых методов предварительной обработки сигналов.

Цель работы — создать метод цифровой фильтрации, основанный на использовании способа определения спектра аналогового сигнала по точкам экстремумов, для технических систем электрохимического измерения, в которых значения полезного сигнала монотонно возрастают или убывают.

Математически эта проблема может быть сформулирована следующим образом. Пусть имеется некоторый полезный сигнал $s(t)$, монотонно возрастающий или убывающий в пределах временного интервала $[0, T]$ на фоне помехи $\xi(t)$. Иными словами, во времени t наблюдается случайный процесс $y(t) = s(t) + \xi(t)$. Требуется восстановить полезный сигнал $s(t)$ путем фильтрации исходного сигнала $y(t)$ с использованием точек экстремумов, когда величина помехи $\xi(t) \rightarrow 0$ на интервале $[0, T]$.

ПОСТРОЕНИЕ СПЕКТРА АНАЛОГОВОГО СИГНАЛА ПО ТОЧКАМ ЭКСТРЕМУМОВ

Прежде чем перейти к изложению основных принципов построения предложенного метода цифровой фильтрации, остановимся на вопросе построения спектра аналогового сигнала путем анализа его точек

экстремумов [8]. Для этого в исходном сигнале $y(t)$ с постоянным временем дискретизации $\Delta\tau$, выбранным согласно теореме Котельникова-Найквиста-Шеннона, и длиной выборки N определяют точки экстремумов с учетом их амплитудно-временных параметров, в которых $y'(t) = 0$.

За точку экстремума сигнала (t_j, y_j) принимается отсчет (t_i, y_i) , для которого выполняется условие:

$$(y_i - y_{i-1})(y_{i+1} - y_i) < 0, \quad (t_j, y_j) = (t_i, y_i), \quad i = 1 \dots N, \quad j = 1 \dots L, \quad (2)$$

где L — общее число точек экстремумов.

Измеряют временной интервал Δt_k между соседними точками экстремумов:

$$\Delta t_k = t_{j+1} - t_j. \quad (3)$$

Временной интервал Δt_k указывает на наличие в сигнале $y(t)$ полуволн k -й частоты f_k , поскольку

$$f_k = 1 / (2 \Delta t_k), \quad (4)$$

при этом текущую амплитуду Δy_k частоты f_k определяют по величине

$$\Delta y_k = |y_{j+1} - y_j| / 2. \quad (5)$$

Суммируют и усредняют по числу значений p_k приращения Δy_k , которые относятся к одинаковым временным интервалам Δt_k , и получают оценку амплитуды a_k спектральной составляющей частоты f_k :

$$a_k = (\sum \Delta y_{kj}) / p_k; \quad j = 1 \dots p_k. \quad (6)$$

В первой итерации обнаруживают высокочастотные компоненты сигнала по отсчетам (t_i, y_i) . В последующих итерациях анализируют в качестве исходных найденные в предыдущей итерации точки экстремумов (t_j, y_j) . Число итераций M зависит от длины массива L начальных точек экстремумов:

$$M \leq \log_2(L). \quad (7)$$

Как правило, в программах фильтрации вместо текущего времени отсчета t_i используют значение $i \Delta\tau$ или только порядковый номер i .

Для дальнейшего изложения, согласно выражению (6), назовем множество точек (f_k, a_k) оценок амплитуды a_k спектральной составляющей частоты f_k — *спектр амплитуд*, а множество точек (f_k, p_k) числа приращений p_k спектральной составляющей частоты f_k — *спектр числа приращений*.

АЛГОРИТМ ФИЛЬТРАЦИИ СИГНАЛА ПО ДВУМ ИЛИ ТРЕМ ЭКСТРЕМУМАМ

Рассмотрим метод цифровой фильтрации, основанный на использовании временных интервалов между точками экстремумов, для выполнения высокочастотной предварительной обработки электрохимических сигналов.

Существует два алгоритма этого метода — фильтрация сигнала по двум или по трем точкам экстремумов. Выбор необходимого алгоритма определяется для каждой технической системы отдельно путем опытного сравнения полученных результатов обработки измерения.

Для обоих алгоритмов метода цифровой фильтрации в исходном сигнале $y(t)$ определяют точки экстремумов (t_j, y_j) в первой итерации, рассчитывают спектр амплитуд (f_k, a_k) и спектр числа приращений (f_k, p_k) . Для выполнения фильтрации $y(t)$ и подавления высокочастотных аддитивных помех $\xi(t)$ задаются значением граничной частоты фильтра f_0 (частоты среза) или граничного временного интервала Δt_0 .

Фильтрацию проводят в два этапа. На первом этапе фильтруют высокочастотные составляющие, вызванные «белым» шумом, при этом считают, что граничный временной интервал фильтра равняется $\Delta t_0 = 3 \cdot \Delta \tau$. На втором этапе фильтрации используют исходное значение граничного интервала Δt_0 , которое соответствует частоте среза фильтра f_0 .

Алгоритм фильтрации сигнала по двум экстремумам включает в себя следующие шаги:

Шаг 1. По точкам экстремумов (t_j, y_j) находят временной интервал (3), который меньше временного интервала фильтра

$$\Delta t_k \leq \Delta t_0. \quad (8)$$

Шаг 2. Определяют среднюю точку временного интервала Δt_k :

$$y_m = (y_{j+1} + y_j) / 2, \quad t_m = (t_{j+1} + t_j) / 2. \quad (9)$$

Шаг 3. Производят децимацию (прореживание) сигнала, т.е. обнуляют все точки из временного интервала Δt_k и заменяют их одной точкой (t_m, y_m) , которую добавляют в сигнал фильтрации.

Шаг 4. Копируют точки (t_i, y_i) , не соответствующие условию (8), в сигнал фильтрации.

Шаг 5. По найденным координатам точек фильтрации выполняют восстановление и сглаживание сигнала для всех N исходных точек.

Разреженный сигнал фильтрации восстанавливают одним из известных способов, например, с помощью модифицированного алгоритма построения локальных кубических эрмитовых сплайнов по четырем точкам сигнала [6]. Модификация алгоритма заключается в уточнении интерполяции граничных диапазонов путем добавления двух точек в начало и конец сигнала. Их амплитуды определяются экстраполяцией кубическим полиномом по четырем крайним точкам. При этом выполнение интерполяции на граничных участках значительно улучшается.

Процесс фильтрации сигнала контролируют по спектру числа приращений (f_k, p_k) , при этом значения числа интервалов p_k для частот f_k , которые выше f_0 , должны стать нулевыми.

Фильтрация сигнала по трем экстремумам выполняется аналогично описанному алгоритму. Изменения в алгоритме касаются способа определения координат точки фильтрации (t_m, y_m) на шаге 2. В данном случае находят средневзвешенную точку временных интервалов Δt_{k-1} и Δt_k по трем экстремумам:

$$y_m = (y_{j+1} + 2y_j + y_{j+1}) / 4, \quad t_m = t_j \quad (10)$$

Как видно из сравнения выражений (9) и (10), в алгоритме фильтрации сигнала по трем экстремумам, временная средневзвешенная координата t_m

совпадает с координатой времени средней точки экстремума. При этом в алгоритмах фильтрации и интерполяции могут использоваться только порядковые номера точек сигнала, а не значения координат времени, что значительно упрощает их программную реализацию.

Рассмотрим пример выполнения предварительной обработки сигнала инверсии при измерении концентрации элемента методом хронопотенциометрии.

ПРИМЕНЕНИЕ АЛГОРИТМА ФИЛЬТРАЦИИ ПО ТРЕМ ЭКСТРЕМУМАМ ДЛЯ ОБРАБОТКИ СИГНАЛА ИНВЕРСИИ МЫШЬЯКА

Используем описанную техническую систему определения концентрации токсичных химических элементов в жидких растворах пробы методами инверсионной хронопотенциометрии.

Рассмотрим пример алгоритма фильтрации по трем экстремумам для высокочастотной обработки небольшого фрагмента сигнала инверсии мышьяка с полной длиной выборки $N = 98500$, частотой дискретизации $\Delta\tau = 0,001$ с и граничной частотой фильтра $f_0 = 35$ Гц в положительной области потенциалов.

На первом этапе используем временной интервал фильтрации равный $\Delta t_0 = 3 \cdot \Delta\tau$, что соответствует частоте фильтра $f_0 = 166,7$ Гц. Вид фрагмента исходного сигнала измерения и полученный сигнал фильтрации в этом случае представлен на рис. 2.

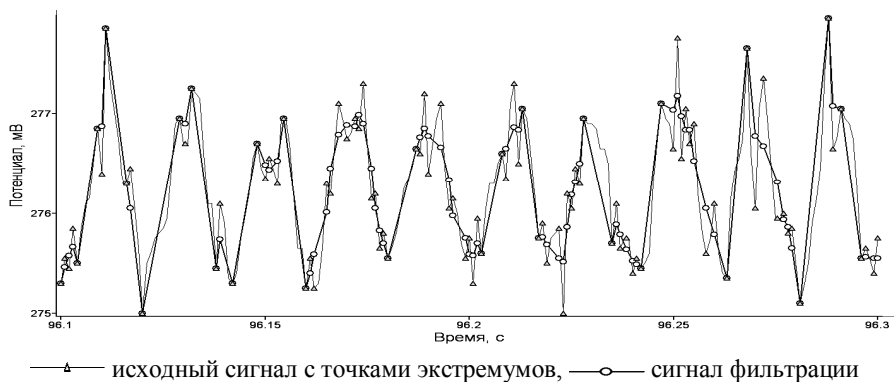


Рис. 2. Фильтрация сигнала по трем точкам экстремумов на первом этапе

Согласно спектру числа приращений, спектральные составляющие, которые больше частоты фильтра $f_0 = 166,7$ Гц, присутствуют в исходном сигнале. Поэтому выполнение первого этапа фильтрации необходимо и обычно его повторяют дважды, контролируя качество фильтрации по значениям спектра.

После интерполяции и сглаживания амплитуд сигнала для всех исходных точек отфильтрованный сигнал (толстая линия) после первого этапа представлен на рис. 3 на фоне исходного сигнала (тонкая линия).

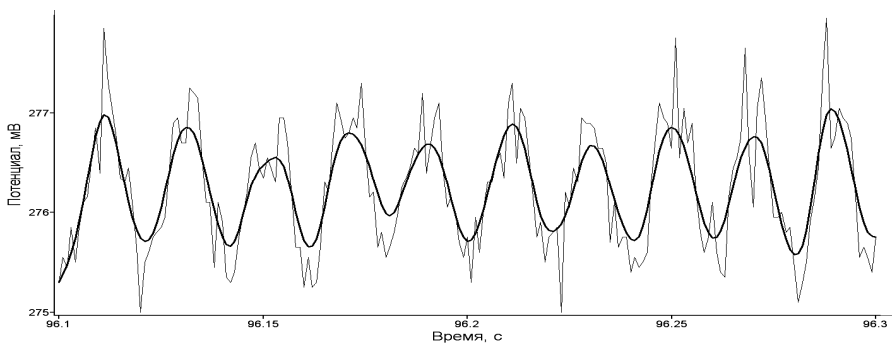


Рис. 3. Отфильтрованный сигнал после первого этапа фильтрации

Согласно спектру числа приращений, спектральные составляющие большие, чем заданная граничная частота фильтра $f_0 = 35$ Гц, наблюдаются в исходном сигнале. Поэтому выполнение второго этапа фильтрации необходимо.

После интерполяции и сглаживания амплитуд сигнала фильтрации второго этапа полностью отфильтрованный сигнал (толстая линия) представлен на рис. 4 на фоне исходного сигнала (тонкая линия).

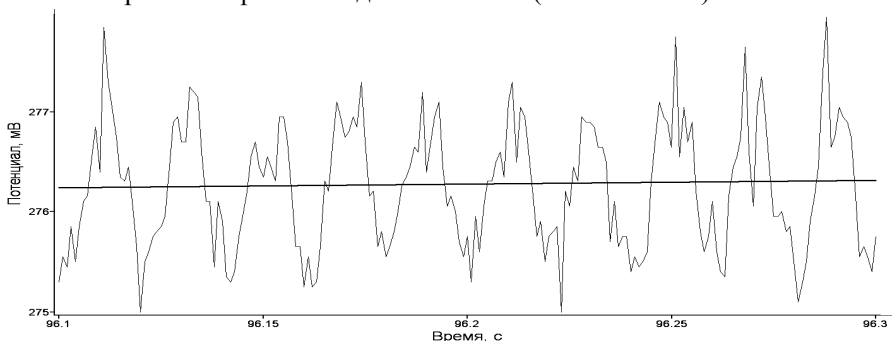


Рис. 4. Отфильтрованный и сглаженный сигнал по трем точкам экстремумов

Рассмотренный пример применения предложенного метода цифровой фильтрации продемонстрировал большую наглядность и простоту предварительной цифровой обработки реального электрохимического сигнала инверсии.

ИСПОЛЬЗОВАНИЕ МЕТОДА ЦИФРОВОЙ ФИЛЬТРАЦИИ В ТЕХНИЧЕСКИХ СИСТЕМАХ ИЗМЕРЕНИЙ

Реализация метода включена в специализированный базовый пакет программ «ТС-ПО», обеспечивающий решение ряда задач цифровой обработки сигналов при создании технических систем измерения, а именно: фильтрации, сглаживания, интерполяции, аппроксимации и моделирования.

Описанный метод цифровой фильтрации показал свою работоспособность, быстродействие, надежность и был многократно внедрен автором в программное обеспечение ряда систем измерений, действующих по настоящее время в разных областях техники.

Создан серийно выпускаемый анализатор солей тяжелых металлов «М-ХА1000-5», который выполняет измерение пяти тяжелых металлов (свинца, кадмия, цинка, меди и олова) с чувствительностью 0,001 мкг/мл. Более 300 приборов используются в аналитических лабораториях СЭС МЗ Украины.

В Международном научно-учебном центре информационных технологий и систем НАН Украины и МОН Украины разработан прибор «Анализатор ИХП», позволяющий определять содержание 14-ти токсичных элементов в растворе пробы (ртути, мышьяка, свинца, кадмия, цинка, меди, никеля, кобальта, олова, марганца, селена, йода, хрома и железа) с чувствительностью 0,0001 мкг/мл. В этом анализаторе удалось реализовать новые методы электрохимического анализа, защищенные патентами Украины: импульсной хронопотенциометрии [9], окислительной хронопотенциометрии [10], дифференциальной хронопотенциометрии [1]. Анализатор широко используется в научных лабораториях университетов и институтов для осуществления экологического мониторинга окружающей среды.

Ярким примером надежности и качества описанного метода цифровой фильтрации является его применение в действующих системах определения массы и нагрузок на оси при взвешивании автомобилей в движении на поосных весах в Международных таможенных пунктах пропуска Украины (МАПП): «Ягодин — Дорогуск РЛ» Волынской таможни; «Мамалыга — Крива МД», «Росошаны — Бричень МД» Черновицкой таможни. А также при создании системы динамического взвешивания автомобилей на платформенных весах, защищенной патентом Украины [11] и действующей на МАПП «Табаки–Мирное МД» Одесской таможни.

Метод также нашел применение в системах непрерывного дозирования руды, штыба и известняка на транспортерных весах для Побужского ферроникелевого комбината Николаевской области и в других технических системах.

Необходимо отметить, что описание приведенного метода цифровой фильтрации и конкретные результаты его применения в реальных системах измерений долгое время не освещались в научной литературе, так как способ построения спектра аналогового сигнала был закрыт для печати.

Выводы

Метод цифровой фильтрации, основанный на использовании точек экстремумов, дает возможность выполнить предварительную высокочастотную обработку разных видов электрохимических сигналов. Он не искажает форму полезного сигнала, не приводит к размыванию границ диапазона изменения компонентов на сигнале интенсивности и существенно минимизирует ошибку определения концентрации токсичных элементов методами хронопотенциометрии.

Круг задач, решаемых с помощью предложенного метода, не ограничивается приведенными техническими системами и может быть

значительно расширен. Объясняется это тем, что метод цифровой фильтрации по существу имеет универсальный характер и поэтому может быть использован для предварительной высокочастотной обработки сигналов физической или химической природы при создании систем измерения в различных предметных областях науки и техники.

1. Пат. 107412 Україна: МПК (2006) G01N 27/48. Пристрій для вимірювання концентрації токсичних елементів / І.В. Суровцев, С.К. Галімов, І.А. Мартинов, О.В. Бабак, В.М. Галімова; Заяв. 21.05.2013; Опубл. 25.12.2014 // Промислова власність. — 2014. — № 24.
2. Суровцев И. В., Татаринов А. Э., Галимов С. К. Моделирование дифференциальных хронопотенциограмм суммой нормальных распределений // Управляющие системы и машины. — 2009. — № 5. — С. 40–45.
3. Оппенгейм А.В., Шафер Р. В. Цифровая обработка сигналов / — М: Техносфера, 2006. — 856 с.
4. Файнзильберг Л.С. Информационные технологии обработки сигналов сложной формы. Теория и практика. — К.: Наук. думка, 2008. — 333 с.
5. Задирака В.К., Мельникова С. С. Цифровая обработка сигналов. — К.: Наук. думка, 1993. — 294 с.
6. Сплайни в цифровій обробці даних і сигналів / І.В. Шелевицький, М.О. Шутко, В.М. Шутко та ін. — Кривий Ріг: Видавничий дім. — 2008. — 232 с.
7. А.С. 845600 СССР. Способ определения спектра аналогового сигнала / В.И. Скурихин, И.Д. Пономарева, П.М. Сиверский, Г.В. Цепков. — 1981.
8. Цепков Г.В. Методы сжатия данных для быстрых корреляционно-спектральных преобразований // Вісн. Східноукр. нац. ун-ту ім. В. Даля. — 2013. — № 15 (204). — С. 222–229.
9. Пат. 96375 Україна: МПК (2006) G01N 27/48. Пристрій для вимірювання концентрації важких металів / І.В. Суровцев, І.А. Мартинов, В.М. Галімова, О.В. Бабак; Заяв. 02.06.2010; Опубл. 25.10.2011 // Промислова власність. — 2011. — № 20.
10. Пат. 104062 Україна: МПК (2006) G01N 27/48. Аналого-цифровий електрохімічний пристрій для вимірювання параметрів розчинів / І.В. Суровцев, В.А. Копілевич, В.М. Галімова, І.А. Мартинов, О.В. Бабак; Заяв. 28.05.2012; Опубл. 25.12.2013 // Промислова власність. — 2013. — № 24.
11. Пат. 106013 Україна: МПК (2006) G01G 19/02. Система поєсового зважування на платформних вагах / І.В. Суровцев, О.В. Бабак, О.Е. Татарінов, Ю.А. Крижановський; Заяв. 06.08.2013; Опубл. 10.07.2014 // Промислова власність. — 2014. — № 13.

UDC 004.021:004.67

THE METHOD OF DIGITAL FILTERING OF ELECTROCHEMICAL SIGNALS IN THE CHRONOPOTENTIOMETRY

I.V. Surovtsev

*International Research and Training Center for Information Technologies and
Systems of National Academy of Sciences of Ukraine and Ministry of Education and
Science of Ukraine (Kiev)*

Introduction. It is important to use methods of digital filtration of signals, that do not distort the form of signal and use its internal characteristics, such as points of extrema for systems of measuring the concentration of toxic elements in chronopotentiometry.

The purpose of research is to create a method digital filtering by using extrema points for performing high-frequency treatment of different types of electrochemical signals while maintaining the shape of the useful signal which increases monotonically.

Methods. The method of digital filtering is based on using of the method of determining the spectrum of the analog signal by points of extrema.

Results. Created method of high-frequency filtration of electrochemical signals has reduced errors in determining the concentration, since it does not distort the form of the useful signal and does not lead to a blurring of the boundaries of the components of measurement of elements. The method is actively used in existing devices measuring the concentration toxic elements in the systems of dynamic axle-by-axle weighting of automobiles and continuous dosing, as well as in many other technical systems of measurement.

Conclusion. The proposed method of digital filtering has substantially universal character and can be used for preliminary digital processing of very different physical or chemical signals.

Keywords: digital filtering, extrema points of signal, chronopotentiometry.

1. Surovtsev I.V., Galimov S.K., Martynov I.A., Babak O.V., Galimova V.M. *Device for measurement of concentration of toxic elements*. Patent 107412 Ukraine, Int.C1. (2006) G01N 27/48, 2014 (in Ukrainian).
2. Surovtsev I.V., Tatarinov A.E., Galimov S.K. The modeling of the Differential Chronopotentiograms by the Sum of Normal Distributions//*Control System and Computers* — 2009. — № 5. — pp.40–45 (in Russian).
3. Oppenheim A.V., Schaffer R.W. *Discrete-Time Signal Processing* — NJ: Prentice-Hall, 1999. — 860 p.
4. Fainzilberg L.S. *Information technologies of signal processing complex form. Theory and practice* — Kiev: Naukova dumka, 2008 — 333 p. (in Russian).
5. Zadiraka V.K., Melnikova S.S. *Digital signal processing* — Kiev: Naukova dumka, 1993. — 294 p. (in Russian).
6. Shelevitsky I.V., Shutko M.O., Shutko V.M., Kolganova O.O. *Splines in digital data processing and signals* — Kryvyi Rih: Vydavnychyy dim, 2008. — 232 p. (in Ukrainian).
7. Skurykhin V.I., Ponomareva I.D., Siverskij P.M., Tsepkov G.V. Method of determining the spectrum of the analogue signal. Patent 845600 SSSR, 1981 (in Russian).
8. Tsepkov G.V. Methods of data compression for quick spectrum and correlation transformations//*Visnyk Shidnoukrains'kogo nacional'nogo universytetu im. V.Dalja* — 2013. — № 15 (204). — pp. 222–229 (in Russian).
9. Surovtsev I.V., Martynov I.A., Galimova V.M., Babak O.V. *Device for measurement of concentration of heavy metals*. Patent 96375 Ukraine, Int.C1. (2006) G01N 27/48, 2011 (in Ukrainian).
10. Surovtsev I.V., Kopilevych V.A., Galimova V.M., Martynov I.A., Babak O.V. *Analog-digital electro-chemical device for measurement of parameters of solutions*. Patent 104062 Ukraine, Int.C1. (2006) G01N 27/48, 2013 (in Ukrainian).
11. Surovtsev I.V., Babak O.V., Tatarinov O.E., Kryzhanovskiy Y.A. *System for axle-by-axle weighing on platform scales*. Patent 106013 Ukraine, Int.C1. (2006) G01G 19/02, 2014 (in Ukrainian).

Получено 06.07.2015

МНОГОУРОВНЕВЫЙ МЕТОД ПОВЫШЕНИЯ ЛОКАЛЬНОГО КОНТРАСТА И УДАЛЕНИЯ НЕОДНОРОДНОГО ФОНА ИЗОБРАЖЕНИЙ

С.В. Баловсяк¹, И.М. Фодчук¹, Ю.Н. Соловей², Я.В. Луцик¹

¹*Черновицкий национальный университет имени Юрия Федьковича*

²*Буковинский государственный медицинский университет*

На основе многоуровневого подхода разработан метод повышения локального контраста и удаления неоднородного фона изображений с использованием нижней и верхней огибающих сигнала. Созданы алгоритмы для размещения локальных окон на изображении и коррекции огибающих сигнала. Тестирование разработанного метода на примере моделированных и медицинских рентгеновских изображений показало его высокую точность и быстродействие.

Ключевые слова: цифровая обработка изображений, повышение локального контраста, удаление неравномерного фона.

На основі багаторівневого підходу розроблено метод підвищення локального контрасту і видалення неоднорідного фону зображень з використанням нижньої та верхньої огинаючих сигналу. Створено алгоритми для розміщення локальних вікон на зображенні та корекції огинаючих сигналу. Тестування розробленого методу на прикладі модельованих і медичних рентгеновських зображень показало його високу точність і швидкодію.

Ключові слова: цифрове оброблення зображень, підвищення локального контрасту, видалення нерівномірного фону.

ВВЕДЕНИЕ

Повышение локального контраста и удаление неоднородного фона относятся к распространенным задачам цифровой обработки изображений [1–3]. Это объясняется тем, что в большинстве случаев неоднородный фон и контраст возникают на изображениях вследствие экспериментальных условий, которые сложно или невозможно изменить, например, из-за недостатков оптических систем и особенностей освещения. Поэтому глобальная обработка только в частных случаях может обеспечить высокое визуальное качество результата, а универсальным способом повышения контраста и удаления неоднородного фона изображений является их локальная обработка [4]. Такая обработка изображений значительно повышает их информативность и визуальное качество, а также увеличивает точность последующих этапов обработки изображений: распознавания, определения параметров объектов на изображениях и др.

К области использования рассмотренных локальных методов относится цифровая фотография, обработка изображений в технической и медицинской диагностике. Одним из наиболее распространенных методов повышения локального контраста и удаления неоднородного фона изображений является метод адаптивного повышения контрастности изображений [4]. В соответствии с данным методом для окрестности каждого пикселя в пределах заданного скользящего окна вычисляется значение локального контраста.

После этого выполняется преобразование яркости изображения в соответствии с рассчитанным значением локального контраста, учитывая статистические характеристики локальных окрестностей: энтропию, среднее значение, среднее квадратическое отклонение и др. Для расчета статистических характеристик локальных окрестностей также используются их гистограммы. К недостаткам рассмотренного локального метода, которые ограничивают его практическое использование, относятся: низкое быстродействие (поскольку оконная обработка выполняется для каждого пикселя изображения), сложность выбора параметров фильтрации (например, нужно определить, какие именно характеристики локальных окрестностей учитывать при расчете локального контраста) и погрешности в расчете интенсивности изображения-результата.

ПОСТАНОВКА ЗАДАЧИ

Учитывая недостатки существующих локальных методов, задача данной работы заключается в разработке метода повышения локального контраста и удаления неоднородного фона изображений, который обеспечивает высокое быстродействие, точность и простоту выбора параметров фильтрации. Для решения поставленной задачи в данной работе использован многоуровневый подход, при котором на основе начального изображения рассчитывается несколько дополнительных уровней изображения (например, изображения нижней и верхней огибающей сигнала), которые далее используются для повышения локального контраста и удаления неоднородного фона изображений. Максимальное и минимальное значение сигнала, которые используются для определения огибающих, рассчитываются локально в пределах выделенных окон. Быстродействие многоуровневого метода достигается за счет выбора расстояния между центрами окон больше одного пикселя, поэтому промежуточные точки огибающих рассчитываются путем аппроксимации.

Полученные огибающие сигнала требуют коррекции с целью их максимального приближения к реальным огибающим. Главным параметром, который влияет на качество локальной обработки изображений, является размер окна. Остальные параметры фильтрации изображения определяются на основе размера окна или устанавливаются по умолчанию.

Цель — разработка многоуровневого метода повышения локального контраста и удаления неоднородного фона изображений, который использует нижнюю и верхнюю огибающие сигнала, с максимальным быстродействием и точностью при минимальном количестве параметров фильтрации, а также программная реализация разработанного метода в системе MATLAB.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ЛОКАЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Математическая модель повышения локального контраста и удаления неоднородного фона предусматривает деление начального изображения f_n размером $M \times N$ пикселей на отдельные окна w , которые могут пересекаться

(рис. 1). При обработке изображений в оттенках серого в каждый элемент матрицы $f_n = (f_n(i, k))$, где $i = 1, \dots, M, k = 1, \dots, N$, записывается интенсивность соответствующего пикселя (нормированная к 1), а при обработке цветных изображений создаются три аналогичных матрицы для красного, зеленого и синего каналов [1]. Определение максимального и минимального значений сигнала происходит локально в пределах каждого окна w размером $M_w \times N_w$ пикселей, при этом размеры деталей на изображении не должны превышать размер окна. Центры окон размещены в узлах прямоугольной решетки с периодами S_w по ширине и S_H по высоте. Расстояния центров окон от края изображения должны превышать значение $N_w/2$ по ширине и $M_w/2$ по высоте, где

$$N_{w2} = \lceil N_w / 2 \rceil, M_{w2} = \lceil M_w / 2 \rceil. \quad (1)$$

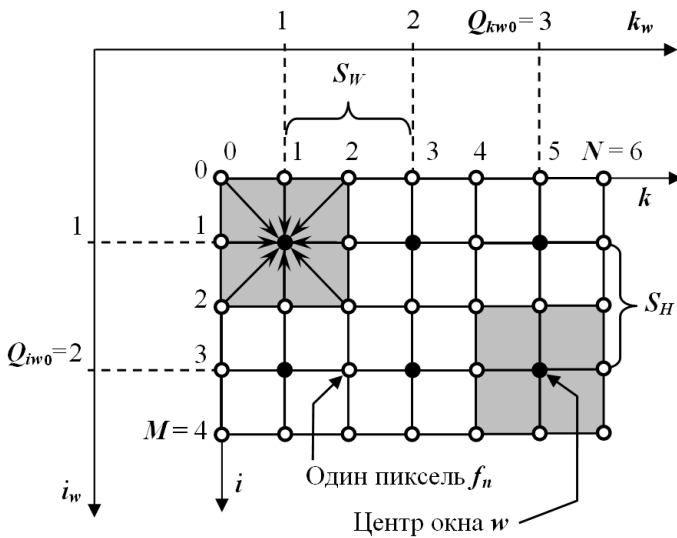


Рис. 1. Системы координат (СК), которые используются при разделении изображения на окна: ik — СК исходного изображения f_n , $i_w k_w$ — СК центров окон w ; $S_H = 2, S_w = 2, M_w = 3; N_w = 3$.

Принципы локальной обработки изображений продемонстрируем в тестовом режиме на примере моделированного изображения f_n (рис. 2б), полученного на основании эталонного изображения f (рис. 2а) путем добавления неоднородного фона и контраста.

Благодаря расчету положения окон w по формуле (1), удастся избежать краевых эффектов, поскольку размещение центра окна на краю изображения эквивалентно симметричному расширению изображения на величину $\lceil N_w / 2 \rceil$ и $\lceil M_w / 2 \rceil$ соответственно [1].

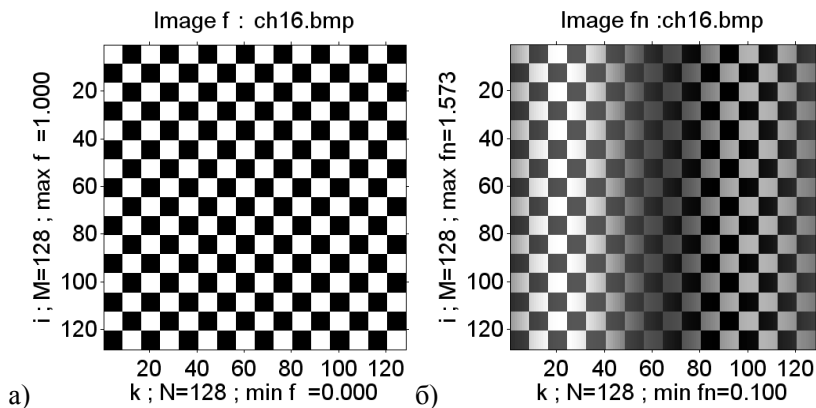


Рис. 2. Исходные изображения: а) эталонное изображение f (размер клетки 8×8 пикселей); б) модельное изображение f_n с неоднородным фоном и контрастом, полученное на основе f

Такое расширение изображения приводит к искусственному увеличению размеров объектов на границах и появлению артефактов при его фильтрации. Выполнение условия (1) приводит к тому, что расстояние от центров последнего столбца окон до предыдущего может отличаться от N_{w2} , а расстояние от центров последней строки окон до предыдущей — от M_{w2} (рис. 3а). При этом начальное количество строк Q_{iw0} и столбцов Q_{kw0} для окон w вычисляется по формуле

$$Q_{iw0} = \left[\frac{M - M_w - 1}{S_H} - \frac{1}{2M} \right] + 2, \quad Q_{kw0} = \left[\frac{N - N_w - 1}{S_W} - \frac{1}{2N} \right] + 2. \quad (2)$$

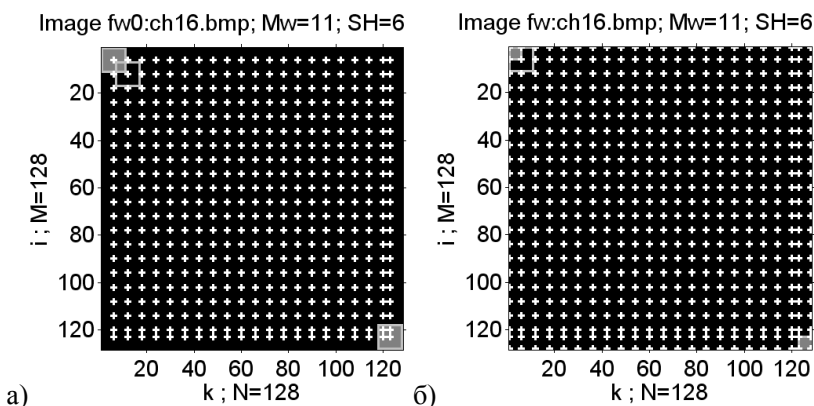


Рис. 3. Схема расположения центров окон w на изображении f_n (рис. 2б): а) $fw0$ — исходное расположение; б) fw — после расширения матриц w_{Dn0} и w_{Up0} ; центры окон показаны крестиками

Минимальные значения интенсивности изображения (в пределах окна w) сохраняются в прямоугольной матрице $w_{Dn0} = (w_{Dn0}(i_w, k_w))$, а максимальные значения интенсивности изображения — в матрице $w_{Up0} = (w_{Up0}(i_w, k_w))$, где

i_w — номер ряда окон, k_w — номер столбца окон; $i_w = 1, \dots, Q_{iw0}$, $k_w = 1, \dots, Q_{kw0}$ (рис. 1).

Далее на основе матриц w_{Dn0} и w_{Up0} осуществляется аппроксимация нижней и верхней огибающей сигнала. Для корректной аппроксимации выполняется расширение матриц w_{Dn0} и w_{Up0} путем добавления новых значений на границе изображения (рис. 3б), при этом матрицы w_{Dn0} и w_{Up0} симметрично расширяются до матриц w_{Dn} (рис. 4а) и w_{Up} (рис. 4б) соответственно.

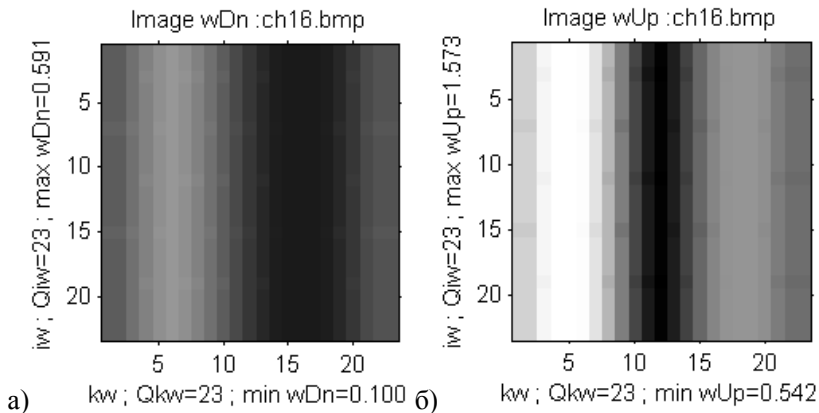


Рис. 4. Минимальные w_{Dn} (а) и максимальные w_{Up} (б) значения интенсивности изображения f_n (рис. 2б) в пределах окон w

РАСЧЕТ НИЖНЕЙ И ВЕРХНЕЙ ОГИБАЮЩИХ ИЗОБРАЖЕНИЯ

На основе минимальных w_{Dn} и максимальных w_{Up} значений интенсивности изображения выполняется расчет нижней g_{Dn} и верхней g_{Up} огибающих изображения путем аппроксимации (рис. 5). Размерность огибающих g_{Dn} и g_{Up} совпадает с размерностью входного изображения f_n . Аппроксимация выполняется с помощью кубических полиномиальных функций, а программно реализуется в системе MATLAB [5] функцией

$$\text{griddata}(pk, pi, pz, ka, ia, \text{method_A}), \quad (3)$$

где pk, pi — координаты центров окон относительно осей координат k и i , pz — минимальные w_{Dn} или максимальные w_{Up} значения интенсивности изображения в пределах окон, ka, ia — координаты пикселей огибающих изображения (после аппроксимации) относительно осей координат k и i , method_A — метод аппроксимации полиномиальными функциями: 'linear' — линейными; 'cubic' — кубическими и др.

В тестовом режиме вычисляются средние квадратические отклонения (СКО) R_{qDn} и R_{qUp} для аппроксимированных огибающих g_{Dn}, g_{Up} относительно эталонных огибающих f_{Dn}, f_{Up} .

Рассмотрим дальнейшую обработку на примере верхней огибающей изображения, обработка нижней огибающей выполняется аналогично. Для повышения точности расчета огибающих изображения выполняется аппроксимация двух верхних огибающих: g_{UpL} — линейным методом,

g_{Up} — кубическим. На основе значений g_{UpL} и g_{Up} вычисляется суммарная верхняя огибающая изображения g_{UpLC} по формуле

$$g_{UpLC}(i, k) = g_{UpL}(i, k) + (g_{Up}(i, k) - g_{UpL}(i, k)) \cdot k_{LC}, \quad (4)$$

где k_{CL} — коэффициент нелинейности; $i = 1, \dots, M, k = 1, \dots, N$.

При увеличении значения коэффициента k_{CL} нелинейность огибающей изображения g_{UpLC} также увеличивается.

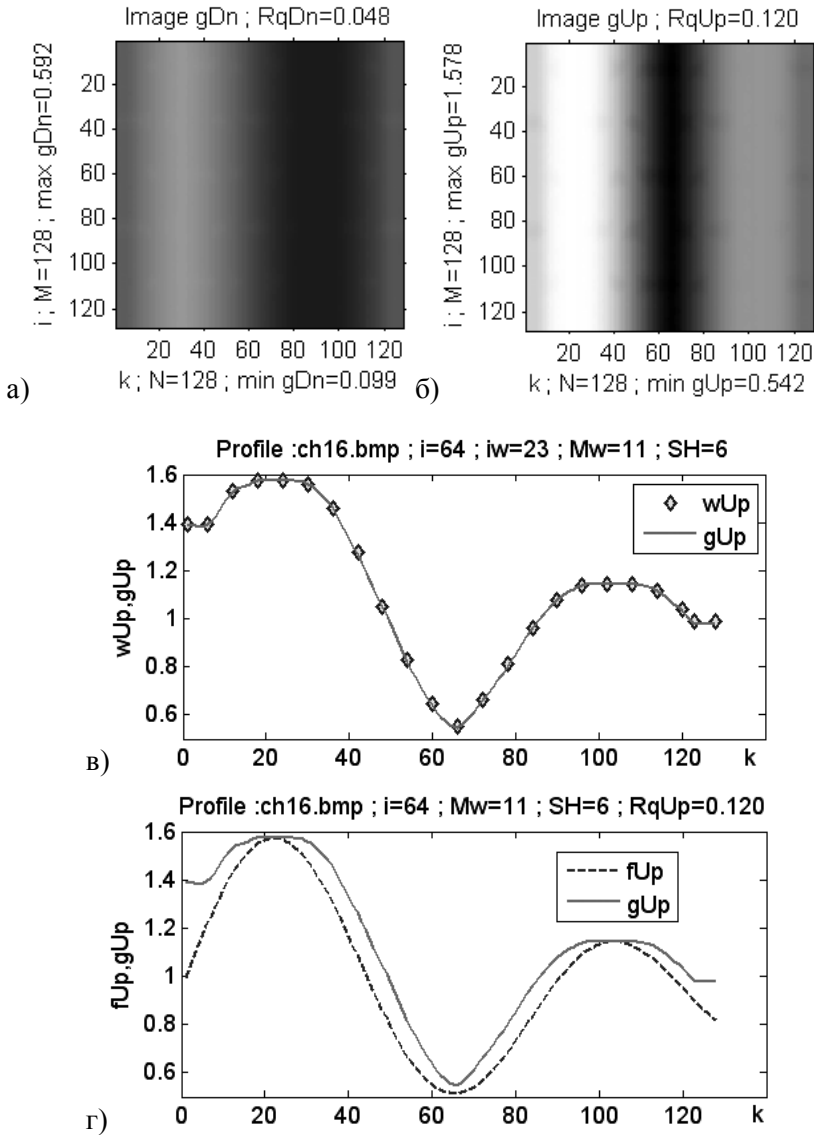


Рис. 5. Огибающие изображения f_n (рис. 2б): а) нижняя аппроксимированная огибающая gDn ; б) верхняя аппроксимированная огибающая gUp ; в) профили максимальных значений интенсивности изображения wUp и огибающей gUp ; г) профили аппроксимированной gUp и эталонной fUp огибающих

Для удаления высокочастотных всплесков на огибающей g_{UpLC} выполняется низкочастотная фильтрация с помощью свертки с ядром фильтра Гаусса G , в результате чего получается фильтрованная огибающая

$g_{Upf} = g_{UpLC} * G$. Наилучшие результаты фильтрации огибающей g_{Upf} (минимальное значение СКО R_{qUpf} огибающей g_{Upf} относительно эталонной f_{Up}) получены для коэффициента нелинейности $k_{CL}=3$ и среднего квадратического отклонения ядра фильтра Гаусса $\sigma_{wC} = M_w / 2$ (рис. 6а).

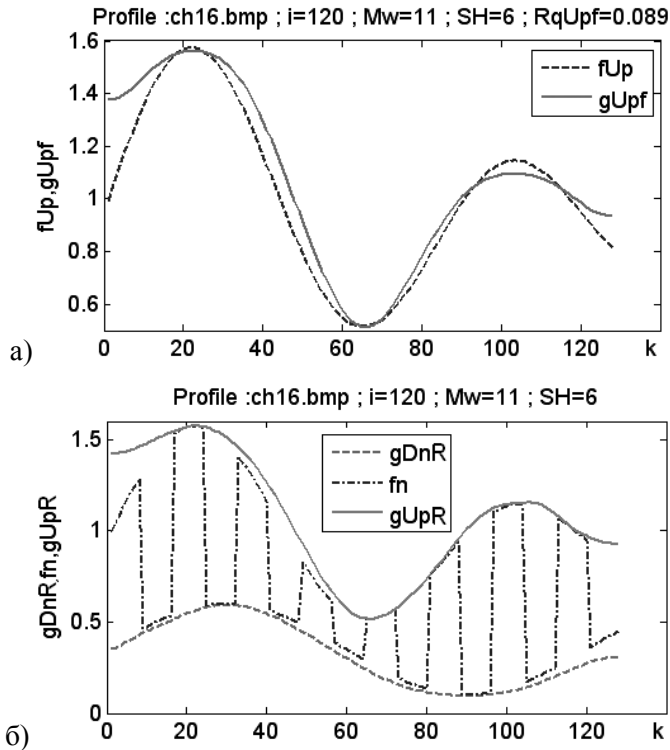


Рис. 6. Корректированные профили огибающих изображения f_n (рис. 2б): а) аппроксимированный профиль верхней огибающей после свертки g_{Upf} и профиль эталонной верхней огибающей f_{Up} ; б) профили результирующих огибающих (g_{UpR} , g_{DnR}) и исходного изображения f_n

Результирующая огибающая g_{UpR} (рис. 6б) рассчитывается на основе g_{Upf} по следующему алгоритму. Начальные значения g_{UpR} равны g_{Upf} , а если значение интенсивности пикселя изображения $f_n(i, k)$ превышает значение огибающей $g_{UpR}(i, k)$, то к g_{UpR} добавляется двумерная функция Гаусса со следующими параметрами: координаты центра (i, k) , СКО $\sigma_{wL} = \max(M_w / 3, N_w / 3)$, амплитуда $A_{wL} = (f_n(i, k) - g_{UpR}(i, k))$.

ПОВЫШЕНИЕ ЛОКАЛЬНОГО КОНТРАСТА И УДАЛЕНИЕ НЕОДНОРОДНОГО ФОНА

Изображение-результат g рассчитывается на основе результирующих огибающих (g_{UpR} , g_{DnR}) и начального изображения f_n . Используется три режима расчета g : 1) удаление неоднородного фона; 2) повышение локального контраста; 3) удаление неоднородного фона и повышение

локального контраста. В наиболее общем третьем режиме интенсивность g рассчитывается по формуле

$$g(i, k) = \frac{f_n(i, k) - q_{DnR}(i, k)}{g_{UpR}(i, k) - g_{DnR}(i, k)}. \quad (5)$$

В результате расчета интенсивности изображения по формуле (5) в значительной степени удаляется неоднородный фон и увеличивается локальный контраст (рис. 7). В тестовом режиме точность восстановления изображения g описывается СКО R_{qg} для интенсивности изображения g относительно эталонного изображения f .

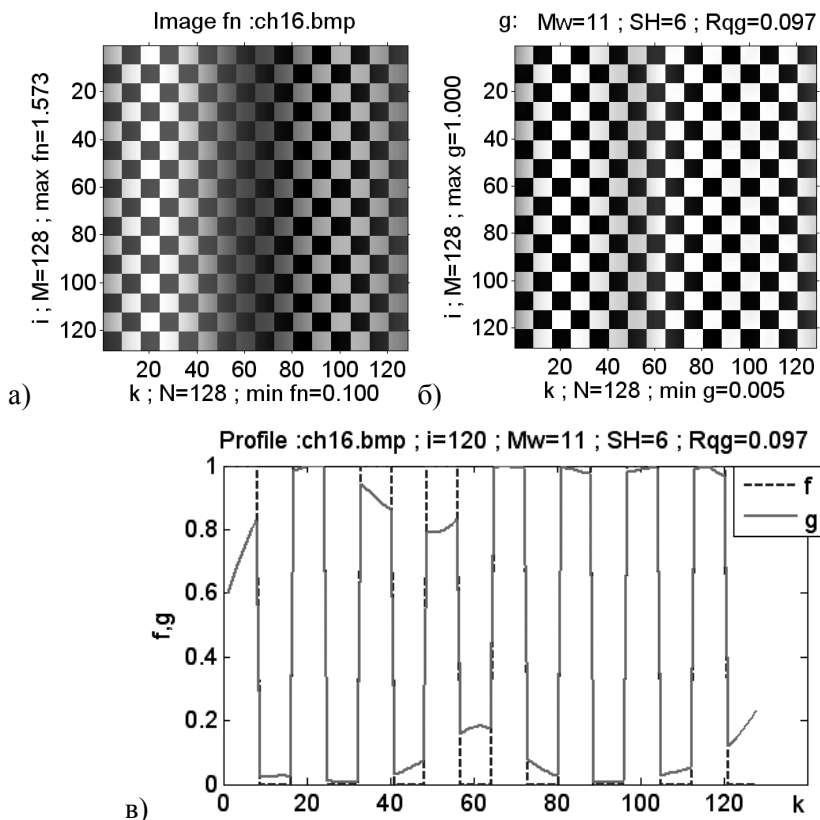


Рис. 7. Результаты обработки изображения: а) исходное изображение f_n (рис. 2б); б) изображение-результат g ; в) профили распределения интенсивности результата g и эталонного изображения f

Таким образом, многоуровневая локальная обработка начального изображения f_n заключается в расчете таких уровней:

- минимальных w_{Dn} и максимальных w_{Up} значений интенсивности изображения f_n в пределах окна w ;
- нижних g_{DnL} , g_{Dn} и верхних g_{UpL} , g_{Up} огибающих изображения, аппроксимированных линейными и кубическими полиномиальными функциями;
- нижней g_{DnLC} и верхней g_{UpLC} суммарных огибающих изображения;
- нижней g_{Dnf} и верхней g_{Upf} фильтрованных огибающих изображения;

- нижней g_{DnR} и верхней g_{UpR} результирующих огибающих изображения;
- изображения-результата g .

ВЫБОР ОПТИМАЛЬНЫХ РАССТОЯНИЙ МЕЖДУ ЦЕНТРАМИ ОКОН

Для выбора оптимального расстояния между центрами окон по ширине (S_W) и по высоте (S_H) выполнена обработка тестовых изображений при различных значениях S_H ($S_W = S_H$) без коррекции огибающих (рис. 8а). На основе минимума СКО R_{qg} для интенсивности изображения g относительно эталонного изображения f получены оптимальные значения:

$$S_{H0} = [M_w/2] + 1, S_{W0} = [N_w/2] + 1. \quad (6)$$

В случае $S_H < S_{H0}$ форма огибающей зависит от интенсивности деталей изображения, что приводит к увеличению СКО R_{qg} (рис. 8а). Осцилляции значений СКО R_{qg} при $S_H > S_{H0}$ объясняются периодичностью эталонного изображения f .

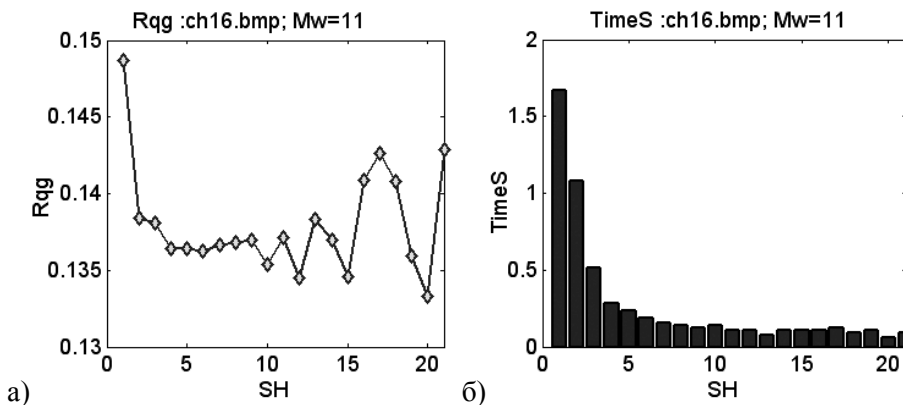


Рис.8. Результаты локальной обработки изображений при различных значениях шага S_H : а) зависимость СКО R_{qg} от S_H ; б) зависимость времени обработки $TimeS$ от S_H ; $S_{H0} = 6, S_{W0} = 6$.

Кроме того, поскольку S_{H0} , S_{W0} вычисляются по формуле (6) между окнами время $TimeS$ локальной обработки изображений (рис. 8б) сокращается по сравнению с распространенным методом адаптивного повышения контрастности изображений в $t_R \sim (M_w \times N_w)^2$ раз, например, при размере окна $M_w \times N_w = 11 \times 11$ пикселей, значение $t_R \approx 10$ раз.

ОБРАБОТКА РЕНТГЕНОВСКИХ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Для проверки возможностей разработанного метода выполнена обработка рентгеновских медицинских изображений органов брюшной полости (рис. 9), полученных с помощью комплекса рентгеновского диагностического (Toshiba, КРАС). При обработке таких изображений, используя рассчитанные огибающие сигнала, обеспечено одновременно удаление неоднородного фона и усиление локального контраста, что значительно повышает визуальное качество изображений и позволяет лучше выделить объекты на изображениях.

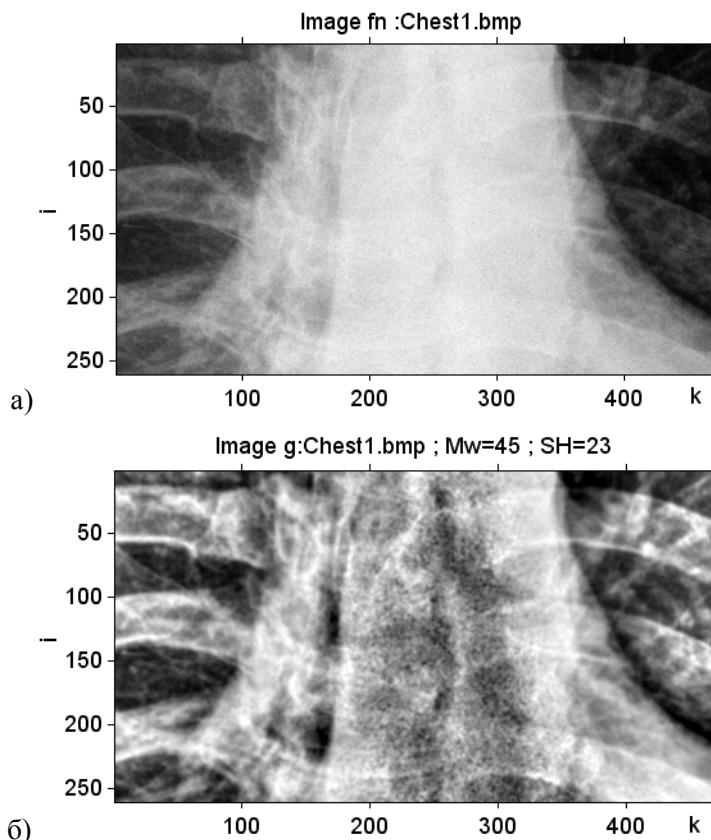


Рис. 9. Пример обработки фрагмента рентгеновского медицинского изображения позвоночника, полученного при рентгенографии органов брюшной полости (высота изображения $M = 260$, ширина $N = 470$ пикселей): а) исходное изображение f_n ; б) результат обработки g

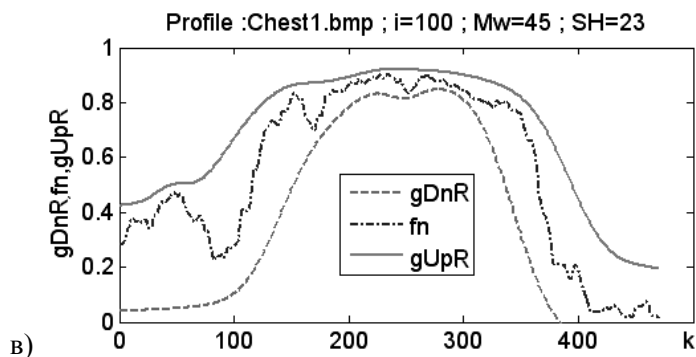


Рис. 9. в) профили распределения интенсивности изображения f_n и его огибающих $gDnR$, $gUpR$

Выводы

Разработан и программно реализован в системе MATLAB многоуровневый метод удаления неоднородного фона и повышения локального контраста изображений, в результате чего сделаны следующие выводы.

Краевые эффекты на изображении минимизируются благодаря расширению матриц минимальных и максимальных значений интенсивности начального изображения.

Обработка огибающих изображения, полученных методами линейной и кубической аппроксимации, а также их последующая свертка, увеличивают точность расчета фильтрованной огибающей.

Время обработки изображений сокращается по сравнению с попиксельной обработкой в $t_R \sim (M_w \times N_w)^2$ раз, например при размере окна $M_w \times N_w = 11 \times 11$ пикселей $t_R \approx 10$ раз.

Оптимальные расстояния между центрами окон по высоте и по ширине равны $S_{H0} = [M_w/2] + 1$ и $S_{W0} = [N_w/2] + 1$ соответственно, где M_w и N_w – размеры окна.

При многоуровневой обработке огибающих изображения увеличивается точность расчета результирующих огибающих, а соответственно, и изображения-результата.

Успешно удален неоднородный фон и усилен локальный контраст как для тестовых моделированных изображений, так и для медицинских рентгеновских изображений.

1. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений. — М.: Техносфера, 2005. — 1072 с.
2. Russ J.C. The image processing handbook. 6th ed. — CRC Press, 2011. — 817 p.
3. Особенности проектирования медицинской информационной системы поддержки принятия решений, основанной на интеллектуальном анализе данных / Г.В. Кнышов, А.В. Руденко, Е.А. Настенко и др. // Кибернетика и вычислительная техника. — 2014. — Вып. 177. — С. 79–87.
4. Бондина Н.Н., Муратов Р.Ю. Адаптивные алгоритмы фильтрации и изменения контраста изображения // Вестник НТУ «ХПИ», 2014. — № 35. — С.35–42.
5. Кетков Ю.Л., Кетков А.Ю., Шульц М.М. Matlab 7: программирование, численные методы. — СПб.: БХВ-Петербург, 2005. — 752 с.

UDC 004.9

MULTILEVEL METHOD OF LOCAL CONTRAST INCREASE AND IMAGES HETEROGENEOUS BACKGROUND REMOVAL

S.V. Balovsyak¹, Ia.V. Lutsyk¹, Yu.M. Solovay², I.M. Fodchuk¹

¹Yuriy Fedkovych Chernivtsi National University (Chernovtsy)

²Bukovinian State Medical University (Chernovtsy)

Introduction. The increase of local contrast and removal of heterogeneous background are the widespread problems of the digital image processing [1–4]. In existing local methods, such as the method of images adaptive contrast enhancement, a value of the local contrast is computed in vicinity of each pixel within a predetermined sliding window. The disadvantages of the existing local methods include poor performance, complicated selection of filter parameters and errors in the calculation of the intensity of the resulting image.

The purpose of the paper is to develop a multilevel method of local contrast increasing and removal of heterogeneous background of images with the high

performance and accuracy using the minimal number of filter parameters.

Methods. The signal envelopes are calculated by linear and cubic approximation.

Results. The multilevel method of removing heterogeneous background and local contrast enhancement of images within the window of the $M_w \times N_w$ pixels size has been elaborated and developed in the MATLAB system [5]. By means of the developed method the heterogeneous background has been successfully removed and local contrast has been increased for the test simulated and medical X-ray images.

Conclusions. Time of the image processing by the multilevel method is shortened comparing with per pixel processing in $t_R \sim (M_w \times N_w)^2$ time, for example at the window size of $M_w \times N_w = 11 \times 11$ pixels the value of $t_R \approx 10$ times. The optimal distance between the windows centres on height and width is equal to $S_{H0} = [M_w/2] + 1$ and $S_{W0} = [N_w/2] + 1$, respectively.

Keywords: digital image processing, local contrast increasing, heterogeneous background removal.

1. Gonzalez R., Woods R., Eddins S. Digital image processing. — M.: Technosphere, 2005. — 1072 p. (in Russian)
2. Russ J.C. The image processing handbook. 6th ed. — CRC Press, 2011. — 817 p.
3. Design features of medical information decision support system based on data mining / G.V. Knyshov, A.V. Rudenko, E.A. Nastenka & others // *Cybernetics and Computer Engineering*. — 2014. — Vol. 177. — P. 79–87. (in Russian)
4. Bondina N.N., Muratov R.Yu. Adaptive filtering and image contrast changing algorithms // *Vestnik NTU "KPI"*, 2014. — №35. — P.35–42. (in Russian)
5. Ketkov Y.L., Ketkov A.J., Schulz M. Matlab 7: programming, numerical methods. — SPb. : BHV-Petersburg, 2005. — 752 p. (in Russian)

Получено 20.10.2015

ПОДХОД К ОРГАНИЗАЦИИ РЕКУРРЕНТНО-ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ В СИСТЕМЕ АВТОМАТИЗИРОВАННОЙ СТРУКТУРНО-ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

А.В. Павлов

Международный научно-учебный центр информационных технологий и систем НАН Украины и МОН Украины

Работа нацелена на создание многопоточного программного приложения для решения прикладных задач прогнозирования, экстраполяции и аппроксимации на основе рекуррентных итерационных алгоритмов метода группового учёта аргументов. Предложен подход к организации рекуррентно-параллельных вычислений, позволяющий интегрировать различные методы и алгоритмы в современные программные приложения, что даёт возможность использовать первые широким кругом пользователей. Эффективность приложения продемонстрирована на примере решения задачи прогнозирования инвестиционной безопасности Украины. Средняя абсолютная ошибка в процентах полученных моделей на независимой выборке наблюдений укладывается в интервал $[-7; +7]$. По моделям получен прогноз на следующий год и дана его экономическая интерпретация.

Ключевые слова: многопоточное распараллеливание, паттерны проектирования, метод группового учёта аргументов, рекуррентные вычисления, библиотеки Qt и TBB.

Роботу націлено на створення багатопотокового програмного додатку для розв'язання прикладних задач прогнозування, екстраполяції та апроксимації на основі рекуррентних ітераційних алгоритмів методу групового урахування аргументів. Запропоновано підхід до організації рекуррентно-паралельних обчислень, що дозволяє інтегрувати різні методи та алгоритми в сучасні програмні додатки, даючи можливість використовувати перші широким колом користувачів. Ефективність додатку продемонстровано на прикладі розв'язання задачі прогнозування інвестиційної безпеки України. Середня абсолютна помилка у відсотках одержаних моделей на незалежній вибірці спостережень вкладається в інтервал $[-7; +7]$. За моделями одержано прогноз на наступний рік і надано його економічну інтерпретацію.

Ключові слова: багатопотокове розпаралелювання, патерни проектування, метод групового урахування аргументів, рекуррентні обчислення, бібліотеки Qt і TBB.

ВВЕДЕНИЕ

Разработка и оптимизация методов и алгоритмов решения задач моделирования по статистическим данным безусловно являются базовыми направлениями в науке. Однако даже самые эффективные алгоритмы и технологии теряют свою ценность, если остаются запрограммированными модулями, с которыми может работать только программист-разработчик. Для того, чтобы технология или метод применялись на практике, их следует интегрировать в некоторую программную систему, имеющую удобный, понятный и практичный интерфейс пользователя. Именно такие

программные системы позволяют раскрыть настоящую ценность алгоритмов и методов, заложенных в программное обеспечение.

ОБЗОР РАБОТ

Термин «рекуррентно-параллельные вычисления» появился сравнительно недавно. Впервые он упоминается в работе [1] и определяется как объединение рекуррентных и параллельных видов операций в единую парадигму параллельно-рекуррентных вычислений для перехода на более высокий уровень эффективности программных средств. Далее рекуррентно-параллельные вычисления применялись для оптимизации комбинаторного алгоритма с различными схемами генерации структур моделей [2], где показана высокая степень эффективности распараллеливания (94.2%) и загрузки вычислителей (88.4%). В работе [3] предложен один из самых быстродействующих итерационных алгоритмов метода группового учёта аргументов (МГУА), применяющий рекуррентно-параллельные вычисления. Эффективность его распараллеливания составила 97% на физических ядрах процессора Intel Core i3.

Недостатком работ [1–3] является то, что предложенные рекуррентно-параллельные алгоритмы и методы существуют сами по себе, они представляют незначительную ценность для пользователя; с ними может работать только программист-разработчик.

Цель — разработка системы моделирования на основе рекуррентно-параллельного алгоритма [3] для повышения эффективности применения разработанных методов и алгоритмов для решения задач прогнозирования, экстраполяции и аппроксимации.

ПОСТАНОВКА ЗАДАЧИ

Достижение поставленной цели включает большой спектр задач, начиная от организации процесса импортирования данных в систему до анализа, выдачи и сохранения результатов моделирования.

Данная работа является логическим продолжением работы [4], где был заложен основной каркас системы автоматизированной структурно-параметрической идентификации (САСПИ). Здесь мы сосредоточимся на решении двух задач: интеграция графического интерфейса пользователя (ГИП) с пакетом классов, предназначенных для построения моделей, использующих рекуррентно-параллельные вычисления; разработка удобного и понятного ГИП и его элементов для ввода данных, анализа, представления и выдачи результатов моделирования.

ПАТТЕРНЫ ПРОЕКТИРОВАНИЯ РЕКУРРЕНТНО-ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ В САСПИ

Интерфейс пользователя САСПИ построен на основе библиотеки Qt от Nokia [5]. Библиотека хорошо спроектирована, имеет множество

возможностей и предоставляет разработчику базовый набор элементов для построения систем любой сложности.

ГИП системы, реализованный без использования параллелизма, имеет недостаток, заключающийся в «замораживании» главного окна программы при выполнении каких-либо операций, вызываемых пользователем. Во избежание этого, все вычисления или любые другие задачи должны выполняться не в главном потоке, а в дополнительном.

Библиотека Qt имеет набор классов, служащих для создания и управления дополнительными потоками на высоком уровне. Главной задачей здесь является объединение двух механизмов параллелизма: механизм распараллеливания с помощью библиотеки ТВВ, выполняющий построение моделей [3] и механизм распараллеливания в библиотеке Qt, позволяющий приложению обрабатывать события пользователя во время выполнения вычислений в программе. Идея решения задачи следующая: каждая операция (включая вызов модуля построения моделей, создающий свои ТВВ-потоки), инициируемая пользователем в программе, будет выполняться в дополнительном Qt-потоке. Проиллюстрируем сказанное на рис. 1.

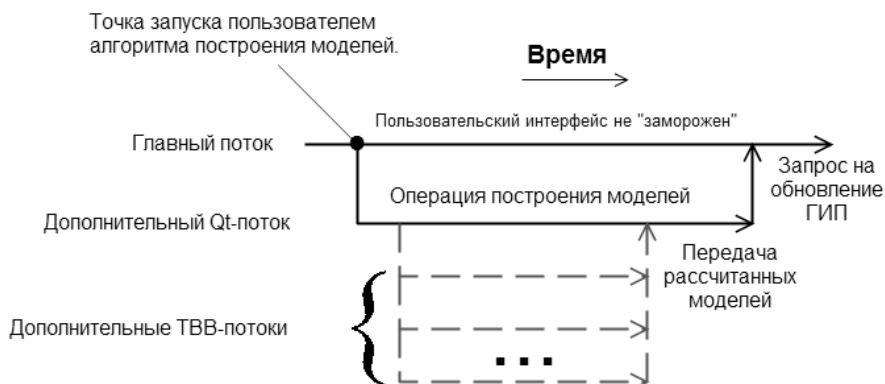


Рис. 1. Идея объединения двух механизмов параллелизма

Библиотека Qt спроектирована так, что классы, реализующие интерфейс пользователя, всегда находятся в главном потоке программы. Поэтому основным требованием, предъявляемым к паттерну проектирования, решающему данную задачу, является обновление ГИП в главном потоке, а не в дополнительном, где происходили вычисления. На рис. 2 представлен паттерн, удовлетворяющий предъявленному требованию.

Класс Thread предназначен для создания дополнительного Qt-потока и запуска задачи в этом потоке. Приведём реализацию метода run данного класса:

```
thread = new QThread;
task->moveToThread( thread );
QObject::connect( thread, SIGNAL( started() ), task,
SLOT( process() ) );
QObject::connect( thread, SIGNAL( finished() ), thread,
SLOT( deleteLater() ));
```

```

QObject::connect( task, SIGNAL( finished() ), thread,
SLOT( quit() ) );
QObject::connect( task, SIGNAL( finished() ), task,
SLOT( deleteLater() ) );
thread->start();

```

Для того чтобы какой-либо код мог быть вызван в параллельном Qt-потоке, он должен содержаться в слоте, поэтому любой класс-задача должен быть унаследован от QObject. Примерами таких классов-задач являются классы ParallelLoadData и ParallelRun. Согласно реализации метода Thread::run, эти классы должны иметь слот process(), который содержит код, запускаемый в параллельном потоке, а также сигнал finished(), говорящий об окончании выполнения параллельных операций.

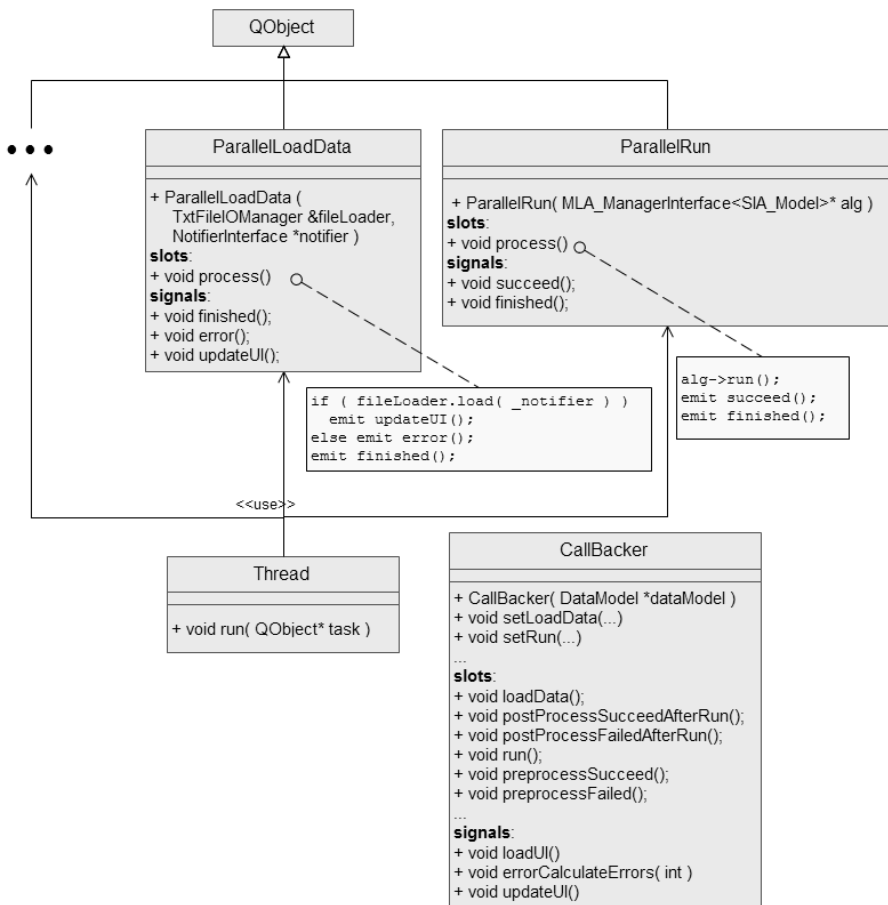


Рис. 2. Диаграмма классов параллельного вызова операций в САСПИ

Таким образом, классы **ParallelLoadData** и **ParallelRun** служат обёрткой для выполнения операций в параллельном потоке. Они делегируют операции загрузки данных и расчёта моделей классам `TxtFileManager` и `MLA_ManagerInterface` соответственно.

Дабы не усложнять класс `DataModel`, наследуя им класс `QObject`, а также логически отделить вызов `callback`-методов, был введён класс `CallBaker`. Его предназначение — обновлять данные модели `DataModel` с помощью «сигнал-слот» механизма после того, как были окончены вычисления какой-либо задачи в параллельном потоке. Следовательно, для каждой параллельной задачи `CallBaker` должен иметь: 1) метод установки параметров, которые будут использоваться при обновлении `DataModel`; 2) слот, который будет вызываться и обновлять `DataModel` после того, как параллельная задача была успешно/неуспешно выполнена.

Рассмотрим последовательность взаимодействия объектов системы при вызове операции построения моделей во времени. Процесс построения моделей состоит из трёх этапов: 1) предварительная обработка данных; 2) построение моделей; 3) расчёт ошибок моделей.

На рис. 3 приведена диаграмма последовательности операций для первого этапа. Метод `callBack_preprocSucceed()` конфигурирует алгоритм построения модели и запускает процесс её расчёта.

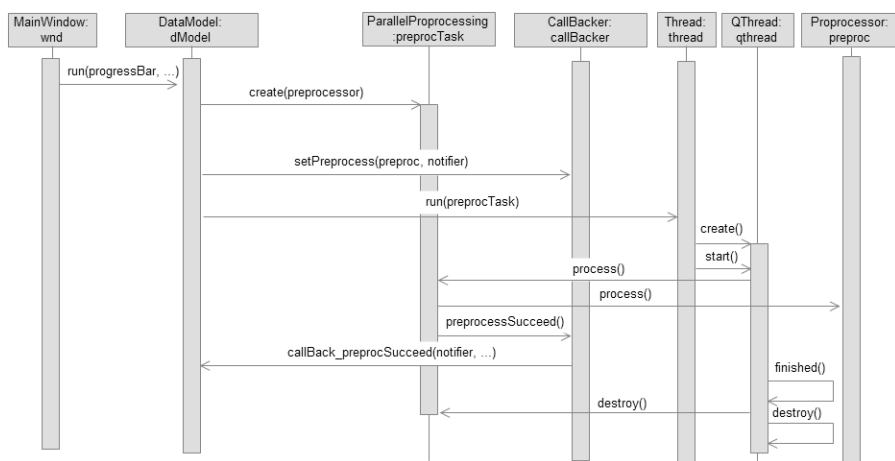


Рис. 3. Диаграмма последовательности операций для параллельной предварительной обработки данных

На рис. 4 представлена диаграмма последовательности операций при расчёте моделей в параллельных потоках. Метод `callBack_runSucceed()` удаляет алгоритм и обновляет объект модели данными построенными моделями. Метод `callBack_run()` удаляет фасад и препроцессор, и создаёт объекты для параллельного расчёта ошибок моделей (рис. 4).

На рис. 5 представлен завершающий этап процесса построения моделей — расчёт ошибок моделей. Метод `callback_postProcessSucceedAfterRun()` обновляет модель данных и вызывает метод `modelChanged()`, обновляющий ГИП.

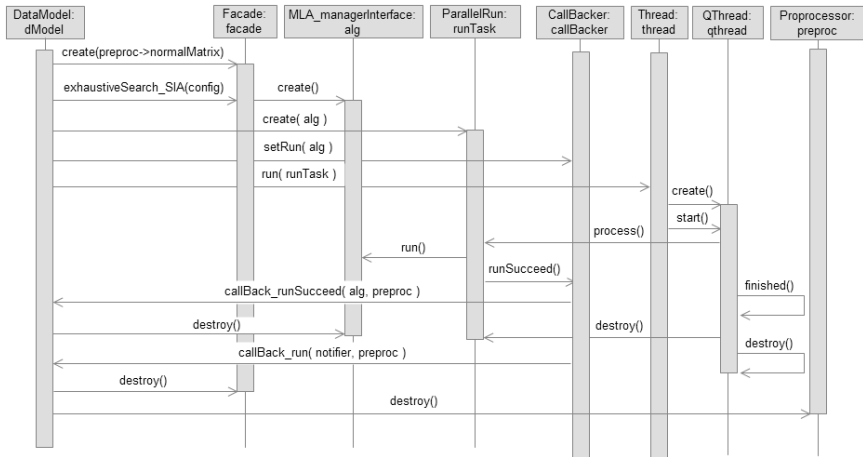


Рис. 4. Диаграмма последовательности операций для процесса расчёта моделей

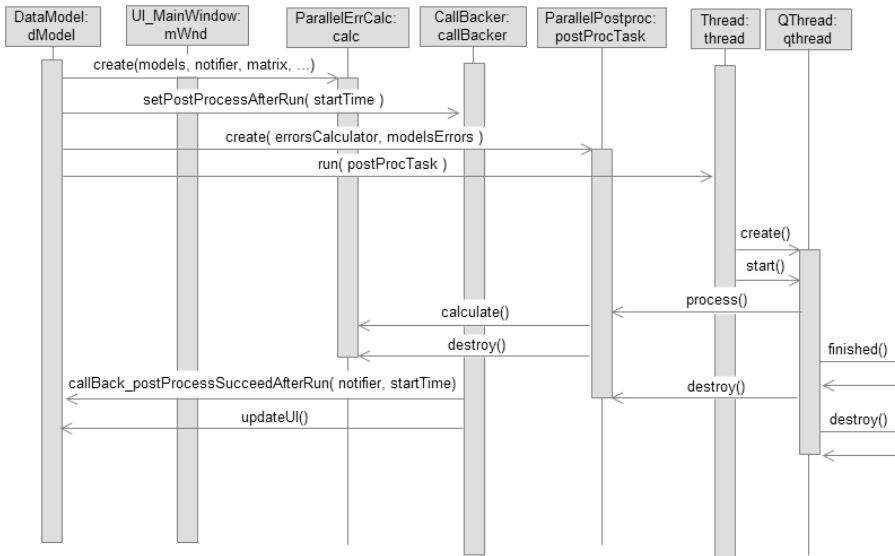


Рис. 5. Диаграмма последовательности операций для процесса расчёта ошибок моделей

Рассмотрим работу с ГИП САСПИ на примере решения задачи прогнозирования инвестиционной безопасности Украины.

ПРОГНОЗИРОВАНИЕ ИНВЕСТИЦИОННОЙ БЕЗОПАСНОСТИ УКРАИНЫ С ПОМОЩЬЮ САСПИ

Важным условием для устойчивого экономического развития страны служит активная и управляемая инвестиционная деятельность. Уровни инвестиционных возможностей, инвестиционной привлекательности и эффективности инвестиционных процессов — важные показатели развития национального хозяйства. Инвестиционная безопасность государства — это уровень соотношения между величинами инвестиций страны за границу и

полученными инвестициями, что удовлетворяет потребности внутренней экономики и поддерживает положительный платежный баланс государства.

В системе показателей инвестиционной безопасности Украины необходимо выделить следующие: x_1 — валовой внутренний продукт, млрд. грн.; x_2 — объёмы инвестиций в основной капитал в фактических ценах, млрд. грн.; x_3 — реальный прирост/снижение инвестиций в основной капитал, процент от предыдущего года; x_4 — степень износа основных фондов, процент от предыдущего года; x_5 — чистый прирост прямых внешних инвестиций (ПВН), млн. дол. США; x_6 — чистый прирост ПВН, млрд. грн.; x_7 — часть иностранных инвестиций относительно объема инвестиций в основной капитал, процент; x_8 — общий объём инвестиций, млрд. грн.

Пренебрежение инвестиционной безопасностью может привести к катастрофическим последствиям: остановке отраслей народного хозяйства, банкротству предприятий и, в конечном итоге, к подрыву системы жизнеобеспечения нации вплоть до потери суверенитета. Поэтому важной научно-прикладной задачей является оценка системы показателей инвестиционной безопасности Украины в будущем.

Задача моделирования состоит в построении прогнозных моделей на один шаг вперёд для каждого из вышеперечисленных индикаторов. С целью получения наиболее точных результатов предполагается построение системы моделей вида:

$$\begin{cases} x_{1,t+1} = f(x_{1,t}, \dots, x_{1,t-l}, x_{2,t}, \dots, x_{2,t-l}, \dots, x_{8,t}, \dots, x_{8,t-l}) \\ \dots \\ x_{8,t+1} = f(x_{1,t}, \dots, x_{1,t-l}, x_{2,t}, \dots, x_{2,t-l}, \dots, x_{8,t}, \dots, x_{8,t-l}) \end{cases}, \quad (1)$$

где $x_i, i = \overline{1,8}$ — показатели, l — количество запаздываний, t — время.

Алгоритм построения моделей САСПИ строит модели линейные по параметрам и линейные/нелинейные по входным переменным. Функция f имеет вид:

$$f(z_1, \dots, z_m, \theta_1, \dots, \theta_m) = \sum_{i=1}^m \theta_i z_i,$$

где в качестве z_i могут выступать как входные переменные x_i , так и их

нелинейные преобразования вида $\prod_{j=1}^k x_j$ (степень k многочлена задаётся

заранее), θ_i — параметры, определяемые по рекуррентному методу с использованием метода наименьших квадратов (МНК) [2]. В таблице 1 представлены исследуемые показатели.

Таблица 1.

Показатели инвестиционной безопасности Украины (1996-2012 гг.)

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
x_1	81.52	93.37	102.59	130.44	170.07	204.19	225.81	267.34	345.11	441.45	544.15	720.73	948.06	913.35	1082.57	1302.08	1408.89
x_2	12.56	12.4	13.96	17.55	23.63	32.57	37.18	51.01	75.71	93.1	125.25	188.49	233.08	151.78	171.09	238.17	263.73
x_3	-22	-8.8	-6.1	0.4	14.4	20.8	8.9	31.3	28	1.9	19	29.8	-2.6	-41.5	-0.6	22.4	8.3
x_4	40	38	40.4	42.3	43.7	45	47.8	48	49.3	49	51.5	52.6	61.2	60	74.9	74.9	74.9
x_5	541.3	625.4	747.1	471.1	593.2	680.3	916.51	1318.61	1696.31	8021.3	4810.8	8303.4	6234	4303.4	4681.2	4654.3	5100.1
x_6	0.99	1.16	1.83	1.95	3.23	3.65	4.89	7.03	9.02	41.11	24.29	41.93	32.84	33.53	37.15	37.04	40.75
x_7	7.89	9.39	13.11	11.09	12.02	10.09	11.62	12.11	10.65	30.63	16.25	18.2	12.35	18.09	17.84	13.47	13.38
x_8	13.55	13.56	15.79	19.5	26.86	36.23	42.07	58.04	84.74	134.2	149.55	230.42	265.92	185.31	208.24	275.26	304.48

Процесс моделирования. Загруженные данные отображаются в таблице «Содержимое файла» (рис. 6). Загруженные столбцы таблицы отображаются в списке «Переменные». Поскольку выборка совсем небольшая ($n = 17$), ограничимся количеством лагов, $l = 1$.

В большинстве случаев, линейная по переменным модель была не в состоянии дать достаточно точный прогноз, поэтому ко входным переменным, кроме исходных $x_i, i = \overline{1,8}$, были добавлены обратные их величины $1/x_i, i = \overline{1,8}$, а также одночлены второй, и в редких случаях третьей степени.

Поскольку стоит задача построения прогноза на один шаг вперед, экзамен (выборка D) составил последнее наблюдение. Оно будет использовано для проверки качества и выбора наиболее адекватной модели. Рабочую выборку C разделим на подвыборки обучения A и проверки B в соотношении 2/1 соответственно. Критерием отбора лучших моделей алгоритма будет нормированный внешний критерий регулярности NAR_B (NRSS проверки на рис. 6).

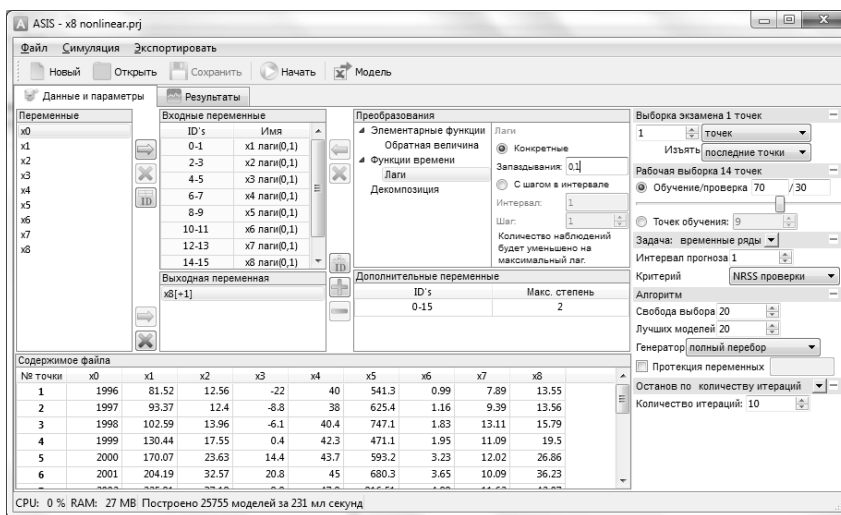


Рис. 6. Настройка параметров моделирования

© А.В. Павлов, 2015

Количество моделей, передаваемых от итерации к итерации — свобода выбора — равна 20 моделям; количество лучших моделей равно двадцати.

Остановку алгоритма будем осуществлять по количеству итераций, а лучшую из 20 моделей определим по следующим трём критериям: 1) точность на экзамене; 2) простота структуры модели; 3) величина коэффициентов при одночленах модели. Установим максимальное количество итераций алгоритма равным десяти.

Результаты моделирования. Результаты моделирования для показателя реального прирост/снижение инвестиций в основной капитал (x_3) представлены на рис. 7. Они отображаются как в графическом, так и числовом виде. Дополнительно рассчитываются остатки моделей и доверительный интервал, равный $[-2\sigma; +2\sigma]$.

Аналитический вид полученных моделей следующий:

$$x_{1,t+1} = 1.29116x_{1,t} - 1.04917 \cdot 10^{-5} x_{8,t} x_{8,t-1} x_{8,t-1} - 12.2856,$$

$$x_{2,t+1} = 0.008213x_{5,t} + 4.41185x_{7,t-1} + 5.17354x_{4,t} - 0.0847269x_{3,t-1} - 240.877,$$

$$x_{3,t+1} = -624.88 \frac{1}{x_{7,t}} - 4.3247 \cdot 10^{-5} x_{5,t} x_{8,t-1} + 0.005263x_{7,t-1} x_{8,t} -$$

$$-0.0008958x_{3,t-1} x_{8,t} + 67.68$$

$$x_{4,t+1} = 0.083198x_{1,t} + 0.00319x_{4,t} x_{8,t} + 35.2331,$$

$$x_{5,t+1} = -500580 \frac{1}{x_{1,t-1}} + 5601, \quad x_{7,t+1} = 0.0668215x_{7,t-1} + 13.2782$$

$$x_{6,t+1} = 0.001549x_{4,t-1} x_{7,t} x_{7,t-1} + 4.84545 \cdot 10^{-8} x_{5,t} x_{8,t-1}^2 -$$

$$-2.06235 \frac{1}{x_{3,t-1}} - 2.5223 \cdot 10^{-5} x_{4,t} x_{4,t-1} x_{7,t} + 1.5618$$

$$x_{8,t+1} = 0.0408775x_{7,t} x_{8,t-1} + 0.11994x_{3,t} x_{7,t-1} + 0.0336721x_4^2 - 42.6608.$$

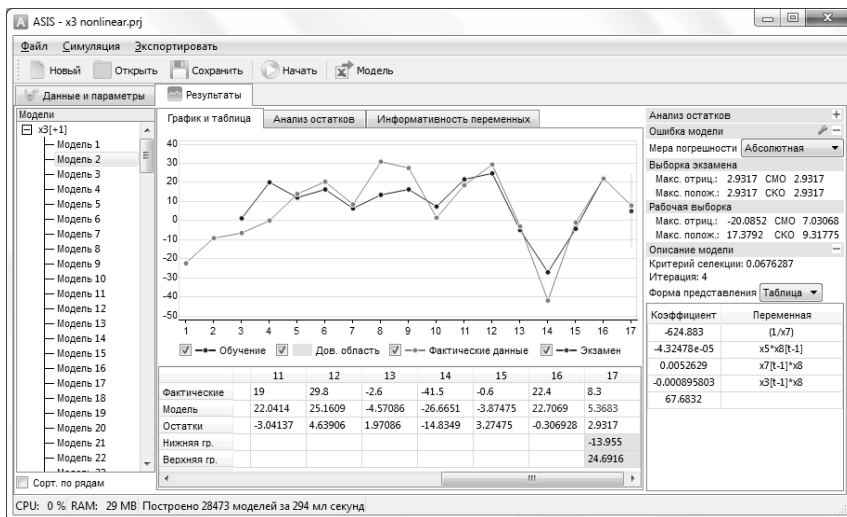


Рис. 7. Результаты моделирования для показателя «Реальный прирост/снижение инвестиций в основной капитал»

Расчёт каждой модели занял меньше секунды. САСПИ также рассчитывает ряд ошибок модели. Значения средней абсолютной ошибки в процентах (САОП) и средней абсолютной ошибки (САО) представлены в таблице 2.

Таблица 2.

Точность моделей на рабочей и экзаменационной выборках

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Рабочая выборка	САОП	4.92	26.52	338	2.39	71.4	74.7	21.78	21.41
	САО	24.11	17.95	7.03	1.29	38.47	4.53	3.53	14.61
Выборка экзамена	САОП	6.58	0.047	35.32	3.86	0.754	0.024	5.96	0.911
	САО	92.76	0.124	2.93	2.89	1408.9	0.01	0.798	2.77

Ошибки прогноза полученных моделей на экзамене укладывается в интервал $[-7\%; +7\%]$ почти для всех моделей (кроме модели для x_3) что говорит об отличных прогностических способностях. Известно, что меру САОП некорректно использовать при малых значений показателя, что наблюдается для показателя x_3 . В таких случаях точность модели лучше оценить по САО или графику (рис. 7). Чистый прогноз по моделям представлен в таблице 3.

Таблица 3.

Прогноз на 2013 год по всем показателям

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
2012	1408.89	263.73	8.3	74.9	5100.1	40.75	13.38	304.48
2013	1538.84	246.03	-24.25	79.70	5216.5	39.21	14.17	309.7

По результатам прогнозирования предполагается увеличение валового внутреннего продукта, степени износа основных фондов, чистого прироста ПВН (дол. США), части иностранных инвестиций относительно объема инвестиций в основной капитал, общего объема инвестиций; уменьшении объемов инвестиций в основной капитал в фактических ценах, реального прироста/снижения инвестиций в основной капитал, чистого прироста ПВН в грн.

Выводы

Предложенный подход к организации рекуррентно-параллельных вычислений в системе автоматизированной структурно-параметрической идентификации (САСПИ) позволил интегрировать эффективные методы решения задач моделирования по экспериментальным данным в удобной, практичной и целостной системе, позволяющей пользователю разного уровня решать задачи прогнозирования, экстраполяции и аппроксимации с высокой эффективностью.

Работа системы продемонстрирована на примере решения задачи прогнозирования инвестиционной безопасности Украины. Полученные модели имеют погрешность на экзамене в интервале $[-7\%; +7\%]$, что говорит о хорошей точности и отличных прогностических свойствах.

Согласно экономической интерпретации прогнозов, полученных по моделям, в 2013 году Украина получит больше инвестиций из-за рубежа, нежели от внутренних инвесторов. Производство Украины будет понемногу наращивать темпы. На фоне роста общего объема инвестиций в страну, инвестиции в основной её капитал будут падать.

1. Ефименко С. Н., Степашко В.С. Основы рекуррентно-параллельных вычислений в комбинаторном алгоритме COMBI МГУА // Управляющие системы и машины. — 2014. — № 6. — С. 27–33.
2. Єфіменко С. М. Комбінаторний алгоритм МГУА з послідовним ускладненням структур моделей на основі рекуррентно-паралельних обчислень // Індуктивне моделювання складних систем. — 2014. — Вип. 6. — С. 81–89.
3. Pavlov A.V. Parallel relaxational iterative algorithm of GMDH. // Індуктивне моделювання складних систем. Зб. наук. праць. — Вип. 6. — К.: МННЦІТС НАНУ. — 2014. — С. 33–40.
4. Павлов А.В. Проектирование системы автоматизированной структурно-параметрической идентификации // Індуктивне моделювання складних систем. Зб. наук. праць. — Вип. 7. — К.: МННЦІТС НАНУ. — 2015. — С. 95–108.
5. Интернет ресурс <https://ru.wikipedia.org/wiki/Qt>

UDC 681.513.8

APPROACH TO ORGANIZATION OF RECURRENT-AND-PARALLEL COMPUTATIONS IN AUTOMIZED STRUCTURE-PARAMETRIC IDENTIFICATION SYSTEM

A.V. Pavlov

International Research and Training Center of Information Technologies and Systems of National Academy of Sciences of Ukraine and Ministry of Education and Science of Ukraine (Kiev)

Introduction. Development and optimization of methods and algorithms for solving statistical modeling problems are the basic directions in science undoubtedly. Although even the most efficient methods and technologies lose their value if they stay just program modules which can be used only by programmers. To be used in practice they should be integrated in some software that has intuitive user-friendly interface. Such software helps to discover a real value of the methods behind it.

The purpose of the paper is increasing the usage effectiveness of a recurrent-and-parallel iterative algorithm by developing a full-fledged modern software based on it for solving forecasting, extrapolation and approximation problems.

Results. The main task of integration of GUI and computational module is a union of two mechanisms of parallization: Threading Building Blocks (TBB) parallization and Qt-parallization. The main idea of proposed approach is that every operation (including the model building operation that create own TBB-threads) initiated by a user should perform in additional Qt-thread. A design pattern that solve this task was developed. The pattern was used to finally implement the ASIS.

The system applied to forecast investments security of Ukraine. A system of forecasting models that describe the state of Ukrainian investments security was

build. Mean absolute percentage error of the models hit the [-7; +7] interval on independent dataset, that indicate a good forecasting ability of the models.

Conclusion. The work suggests an approach to organization of recurrent-and-parallel computations in ASIS that allow integrating the most effective methods for solving statistical modeling problems in user-friendly intuitive full-fledge system that allows any user to solve forecasting, regression and approximation problems with high efficiency. The system was applied to forecast Ukrainian investments security. The economic interpretation of the obtained forecasts says that in 2013 year Ukraine will gain more investments from abroad than from inner investors Ukrainian manufacturing will increase rates; on the background rise of overall country investments, the investments in basic capital will decline.

Keywords: Multithreaded parallelization, design patterns, group method of data handling, recurrent-and-parallel computations, Qt, TBB.

1. Efimenko S., Stepashko V. S, Basics of recurrent-and-parallel computations in combinatorial algorithm COMBI of GMDH // *Controll systems and machines* — 2014. — № 6. — P. 27–33. (in Russian)
2. Efimenko S.M. Combinatorial algorithm of GMDH with sequential complication of structures based on recurrent-and-parallel computations // *Inductive modeling of complex systems* — Vol. 6. — 2014. — P. 81–89. (in Ukrainian)
3. Pavlov A.V. Parallel relaxational iterative algorithm of GMDH // *Inductive modeling of complex systems* — Vol. 6. — 2014. — P. 33–40.
4. Pavlov A.V. Design of automated structure-parametric identification system // *Inductive modeling of complex systems* — Vol.7— 2015. — P. 33–40.
5. Internet resource <https://ru.wikipedia.org/wiki/Qt>.

Получено 28.10.2015

Интеллектуальное управление и системы

УДК 517.977

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ КОЛЕБАНИЯМИ БАЛКИ ПЕРЕМЕННОГО ПОПЕРЕЧНОГО СЕЧЕНИЯ

М.М. Копец

Национальный технический университет Украины «Киевский политехнический институт»

Статья посвящена исследованию линейно-квадратической задачи оптимального управления процессом колебаний балки переменного поперечного сечения в случае свободных концов балки. Для рассматриваемой задачи оптимизации получены необходимые условия оптимальности. Анализ этих условий дал возможность вывести систему интегро-дифференциальных уравнений Риккати с частными производными. Решение этой системы использовано при построении явной формулы для вычисления оптимального управления.

Ключевые слова: линейно-квадратичная задача, оптимальное управление, метод множителей Лагранжа, необходимые условия оптимальности, колебания балки, частные производные, система интегро-дифференциальных уравнений с частными производными.

Статтю присвячено дослідженню лінійно-квадратичної задачі оптимального керування процесом коливань бруса змінного поперечного перерізу у випадку вільних кінців бруса. Для розглядуваної задачі оптимізації одержано необхідні умови оптимальності. Аналіз цих умов дав можливість вивести систему інтегро-диференціальних рівнянь Ріккати з частинними похідними. Розв'язок цієї системи використано при побудові явної формули для обчислення оптимального керування.

Ключові слова: лінійно-квадратична задача, оптимальне керування, метод множників Лагранжа, необхідні умови оптимальності, коливання бруса, частинні похідні, система інтегро-диференціальних рівнянь з частинними похідними.

ВВЕДЕНИЕ

Последнее пятидесятилетие характеризуется бурным развитием техники. Существенный прогресс наблюдается в ракетостроении, самолетостроении, судостроении, космической технике и т.п. В каждой из перечисленных отраслей имеют место быть колебательные процессы. В одних случаях их можно с пользой учитывать для повышения качества процесса, в других, наоборот, надо подавлять в силу их отрицательного влияния на окончательный процесс. Это означает, что колебательные процессы необходимо не только изучать, но и уметь эффективно ими управлять. Подобные задачи эффективного управления механическими процессами как раз и изучает теория оптимального управления. Одной из наиболее фундаментальных проблем в теории оптимального управления является

линейно-квадратичная задача. Под этим термином подразумевается проблема нахождения минимума квадратичного функционала на множестве решений системы линейных дифференциальных уравнений, правые части которых включают в себя управляющие параметры (управления). Если рассматривается система линейных дифференциальных уравнений, то в этом случае линейно-квадратичная задача исследована достаточно подробно [1, 2]. В практических приложениях часто встречаются математические модели, которые включают в себя линейные дифференциальные уравнения с частными производными или же системы таких уравнений. Понятно, что такие задачи более сложны и поэтому менее исследованы. В монографии [3, с. 56] предложена идея трактовать движение летательного аппарата подобно колебаниям упругой балки переменного поперечного сечения со свободными концами.

В данной статье эта идея получила дальнейшее развитие в форме задачи оптимального гашения колебательных процессов рассматриваемого процесса.

Цель работы — для решения задачи оптимального управления процессом колебаний балки переменного поперечного сечения в случае свободных концов балки определить необходимые условия оптимальности и вывести решение системы интегро-дифференциальных уравнений Риккати с частными для вычисления оптимального управления.

ПОСТАНОВКА ЗАДАЧИ

Вынужденные колебания балки переменного поперечного сечения можно описать следующим линейным дифференциальным уравнением с частными производными:

$$m(x) \frac{\partial^2 z(t, x)}{\partial t^2} + \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 z(t, x)}{\partial x^2} \right) = u(t, x), \quad (1)$$

где $0 \leq x \leq l$, $t_0 \leq t \leq t_1$, действительные числа $l > 0$, $t_0 \geq 0$, $t_1 > t_0$ известны, $EJ(x)$ — жесткость балки, $m(x)$ — закон распределения массы вдоль продольной оси балки, функция $z(t, x)$ описывает прогиб балки в точке x в момент времени t . Начальные условия для уравнения (1) имеют вид:

$$z(t_0, x) = h(x), \quad \frac{\partial z(t_0, x)}{\partial t} = j(x), \quad (2)$$

где функции $h(x) \in W_2^{1,0}(0, l)$, $j(x) \in L_2(0, l)$ предполагаются известными. Краевые условия для уравнения (1) заданы следующим образом:

$$\begin{cases} EJ(0) \frac{\partial^2 z(t,0)}{\partial x^2} = 0, & \frac{\partial}{\partial x} \left(EJ(0) \frac{\partial^2 z(t,0)}{\partial x^2} \right) = 0, \\ EJ(l) \frac{\partial^2 z(t,l)}{\partial x^2} = 0, & \frac{\partial}{\partial x} \left(EJ(l) \frac{\partial^2 z(t,l)}{\partial x^2} \right) = 0. \end{cases} \quad (3)$$

Условия (3) означают, что концы балки свободны. Такая ситуация возникает, например, при исследовании динамики полета самолета. Если через Ω обозначить множество $\Omega = \{(t,x) : t \in [t_0, t_1], x \in [0, l]\}$, то функция $u(t,x) \in L_2(\Omega)$ называется допустимым управлением. Для фиксированного допустимого управления $u(t,x)$ под решением $z(t,x)$ уравнения (1) подразумеваем ее обобщенное решение $z(t,x) \in W_2^{1,0}(\Omega)$. Качество управляемого процесса оценивается с помощью следующего функционала:

$$\begin{aligned} I(u, z) = & \frac{1}{2} \int_0^l m(x) z^2(t_1, x) dx + \frac{1}{2} \int_0^l m(x) \left[\frac{\partial z(t_1, x)}{\partial t} \right]^2 dx + \\ & + \frac{1}{2} \int_{t_0}^{t_1} \int_0^l [m(x) z^2(t, x) + u^2(t, x)] dx dt. \end{aligned} \quad (4)$$

Требуется найти допустимое управление $u(t,x)$ и соответствующее ему решение $z(t,x)$ задачи (1)–(3), на которых реализуется минимальное значение функционала (4). Такое управление называется оптимальным.

НЕОБХОДИМЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ

Решение сформулированной выше задачи оптимального управления (1)–(4) может быть найдено с помощью метода множителей Лагранжа [4]. Для этого следует рассмотреть функционал

$$\begin{aligned} \Psi(p, u, z) = & \frac{1}{2} \int_0^l m(x) z^2(t_1, x) dx + \frac{1}{2} \int_0^l m(x) \left[\frac{\partial z(t_1, x)}{\partial t} \right]^2 dx + \\ & + \frac{1}{2} \int_{t_0}^{t_1} \int_0^l [m(x) z^2(t, x) + u^2(t, x)] dx dt + \int_{t_0}^{t_1} \int_0^l p(t, x) [u(t, x) - \\ & - \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 z(t, x)}{\partial x^2} \right) - m(x) \frac{\partial^2 z(t, x)}{\partial t^2}] dx dt. \end{aligned} \quad (5)$$

Такой подход позволяет вместо задачи на условный экстремум (1)–(4) исследовать задачу минимизации функционала (5) с учетом условий (2) и (3). Следующий шаг состоит в нахождении приращения $\Delta \Psi$ функционала (5). С учетом соотношения (5) выражение для $\Delta \Psi$ примет следующий вид:

$$\begin{aligned}
\Delta \Psi = & \frac{1}{2} \int_0^l m(x) [z(t_1, x) + \varepsilon \delta z(t_1, x)]^2 dx + \\
& + \frac{1}{2} \int_0^l m(x) \left[\frac{\partial z(t_1, x)}{\partial t} + \varepsilon \frac{\partial \delta z(t_1, x)}{\partial t} \right]^2 dx + \\
& + \frac{1}{2} \int_{t_0}^{t_1} \int_0^l [m(x) [z(t, x) + \varepsilon \delta z(t, x)]^2 + [u(t, x) + \varepsilon \delta u(t, x)]^2] dx dt + \\
& + \int_{t_0}^{t_1} \int_0^l [p(t, x) + \varepsilon \delta p(t, x)] [u(t, x) + \varepsilon u(t, x) - \\
& - \frac{\partial^2}{\partial x^2} \left(EJ(x) \left(\frac{\partial^2 z(t, x)}{\partial x^2} + \varepsilon \frac{\partial^2 \delta z(t, x)}{\partial x^2} \right) \right) - \\
& - m(x) \left(\frac{\partial^2 z(t, x)}{\partial t^2} + \varepsilon \frac{\partial^2 \delta z(t, x)}{\partial t^2} \right)] dx dt - \\
& - \frac{1}{2} \int_0^l m(x) z^2(t_1, x) dx - \frac{1}{2} \int_0^l m(x) \left[\frac{\partial z(t_1, x)}{\partial t} \right]^2 dx - \\
& - \int_{t_0}^{t_1} \int_0^l p(t, x) [u(t, x) - \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 z(t, x)}{\partial x^2} \right) - m(x) \frac{\partial^2 z(t, x)}{\partial t^2}] dx dt .
\end{aligned} \tag{6}$$

Если предположить выполнение краевых условий

$$\begin{cases} EJ(0) \frac{\partial^2 p(t, 0)}{\partial x^2} = 0, \frac{\partial}{\partial x} \left(EJ(0) \frac{\partial^2 p(t, 0)}{\partial x^2} \right) = 0, \\ EJ(l) \frac{\partial^2 p(t, l)}{\partial x^2} = 0, \frac{\partial}{\partial x} \left(EJ(l) \frac{\partial^2 p(t, l)}{\partial x^2} \right) = 0, \end{cases} \tag{7}$$

то после преобразований, подобных использованным в [4], приходим к такому соотношению:

$$\begin{aligned}
\Delta \Psi = & \varepsilon \int_0^l m(x) \left(z(t_1, x) + \frac{\partial p(t_1, x)}{\partial t} \right) \delta z(t_1, x) dx + \\
& + \varepsilon \int_0^l m(x) \left(\frac{\partial z(t_1, x)}{\partial t} - p(t_1, x) \right) \frac{\partial \delta z(t_1, x)}{\partial t} dx + \\
& + \varepsilon \int_{t_0}^{t_1} \int_0^l \left[m(x) z(t, x) - \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 p(t, x)}{\partial x^2} \right) \right] -
\end{aligned} \tag{8}$$

$$\begin{aligned}
& -m(x)\frac{\partial^2 p(t,x)}{\partial t^2} \Big) + [u(t,x) + p(t,x)]\delta u(t,x) \Big] dxdt + \\
& + \varepsilon \int_{t_0}^{t_1} \int_0^l \delta p(t,x) \left[u(t,x) - \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 z(t,x)}{\partial x^2} \right) - \right. \\
& \left. - m(x) \frac{\partial^2 z(t,x)}{\partial t^2} \right] dxdt + \frac{\varepsilon^2}{2} \int_0^l m(x) [\delta z(t_1,x)]^2 dx + \\
& + \frac{\varepsilon^2}{2} \int_0^l m(x) \left[\frac{\partial \delta z(t_1,x)}{\partial t} \right]^2 dx + \\
& + \frac{\varepsilon^2}{2} \int_{t_0}^{t_1} \int_0^l [m(x) [\delta z(t,x)]^2 + [\delta u(t,x)]^2] dxdt .
\end{aligned}$$

Соотношения (7) и (8) позволяют сделать следующий вывод.

Теорема 1. Для нахождения единственного оптимального управления $u(t, x)$ имеем систему соотношений

$$\left\{ \begin{aligned}
& m(x) \frac{\partial^2 z(t,x)}{\partial t^2} + \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 z(t,x)}{\partial x^2} \right) - u(t,x), \\
& z(t_0, x) = f(x), \quad \frac{\partial z(t_0, x)}{\partial t} = g(x), \\
& EJ(0) \frac{\partial^2 z(t,0)}{\partial x^2} = 0, \quad \frac{\partial}{\partial x} \left(EJ(0) \frac{\partial^2 z(t,0)}{\partial x^2} \right) = 0, \\
& EJ(l) \frac{\partial^2 z(t,l)}{\partial x^2} = 0, \quad \frac{\partial}{\partial x} \left(EJ(l) \frac{\partial^2 z(t,l)}{\partial x^2} \right) = 0, \\
& m(x) \frac{\partial^2 p(t,x)}{\partial t^2} + \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 p(t,x)}{\partial x^2} \right) = m(x)z(t,x), \\
& p(t_1, x) = \frac{\partial z(t_1, x)}{\partial t}, \quad \frac{\partial p(t_1, x)}{\partial t} = -z(t_1, x), \\
& EJ(0) \frac{\partial^2 p(t,0)}{\partial x^2} = 0, \quad \frac{\partial}{\partial x} \left(EJ(0) \frac{\partial^2 p(t,0)}{\partial x^2} \right) = 0, \\
& EJ(l) \frac{\partial^2 p(t,l)}{\partial x^2} = 0, \quad \frac{\partial}{\partial x} \left(EJ(l) \frac{\partial^2 p(t,l)}{\partial x^2} \right) = 0, \\
& u(t, x) + p(t, x) = 0.
\end{aligned} \right. \quad (9)$$

Доказательство. В случае выполнения системы соотношений (9) первая вариация функционала (5) равна нулю, что является необходимым условием экстремума этого функционала. Если имеют место соотношения (9), то выражение (8) станет следующим:

$$\begin{aligned} \Delta \Psi = & \frac{\varepsilon^2}{2} \int_0^l m(x) [\delta z(t_1, x)]^2 dx + \frac{\varepsilon^2}{2} \int_0^l m(x) \left[\frac{\partial \delta z(t_1, x)}{\partial t} \right]^2 dx + \\ & + \frac{\varepsilon^2}{2} \int_{t_0}^{t_1} \int_0^l [m(x) [\delta z(t, x)]^2 + [\delta u(t, x)]^2] dx dt. \end{aligned} \quad (10)$$

Из соотношения (10) следует, что: во-первых, при управлении $u(t, x)$ реализуется минимум функционала (4); во-вторых, оптимальное управление $u(t, x)$ единственно. Таким образом, теорема 1 полностью доказана.

СИСТЕМА ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ РИККАТИ

Далее все равенства, связанные с дельта-функцией Дирака, следует понимать как равенства в смысле теории обобщенных функций. Поскольку система соотношений (9) линейна относительно неизвестных функций $z(t, x)$, $p(t, x)$ и $u(t, x)$, то с учетом условий трансверсальности

$$p(t_1, x) = \frac{\partial z(t_1, x)}{\partial t} \text{ и } \frac{\partial p(t_1, x)}{\partial t} = -z(t_1, x) \text{ рассмотрим такие два выражения}$$

$$\frac{\partial p(t, x)}{\partial t} = - \int_0^l m(y) \left[R_{11}(t, x, y) z(t, y) + R_{12}(t, x, y) \frac{\partial z(t, y)}{\partial t} \right] dy,$$

$$p(t, x) = \int_0^l m(y) \left[R_{21}(t, x, y) z(t, y) + R_{22}(t, x, y) \frac{\partial z(t, y)}{\partial t} \right] dy.$$

Подобно тому, как это было сделано в статье [4], для нахождения функций $R_{11}(t, x, y)$, $R_{12}(t, x, y)$, $R_{21}(t, x, y)$, $R_{22}(t, x, y)$ получим такую систему уравнений с частными производными:

$$\begin{aligned} m(x)m(y) \frac{\partial R_{11}(t, x, y)}{\partial t} - m(x) \frac{\partial^2}{\partial y^2} \left(EJ(y) \frac{\partial^2 R_{12}(t, x, y)}{\partial y^2} \right) - \\ - m(y) \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 R_{21}(t, x, y)}{\partial x^2} \right) + m(y) \delta(x-y) - \end{aligned} \quad (11)$$

$$- m(x)m(y) \int_0^l R_{12}(t, x, s) R_{21}(t, s, y) ds = 0,$$

$$\begin{aligned} m(x) \frac{\partial R_{12}(t, x, y)}{\partial t} - \frac{\partial^2}{\partial x^2} \left(EJ(x) \frac{\partial^2 R_{22}(t, x, y)}{\partial x^2} \right) + \\ + m(x) R_{11}(t, x, y) - m(x) \int_0^l R_{12}(t, x, s) R_{22}(t, s, y) ds = 0, \end{aligned} \quad (12)$$

$$m(y) \frac{\partial R_{21}(t, x, y)}{\partial t} - \frac{\partial^2}{\partial y^2} \left(EJ(y) \frac{\partial^2 R_{22}(t, x, y)}{\partial y^2} \right) +$$

$$+ m(y) R_{11}(t, x, y) - m(y) \int_0^l R_{22}(t, x, s) R_{21}(t, s, y) ds = 0, \quad (13)$$

$$\frac{\partial R_{22}(t, x, y)}{\partial t} + R_{12}(t, x, y) + R_{21}(t, x, y) - \int_0^l R_{22}(t, x, s) R_{22}(t, s, y) ds = 0, \quad (14)$$

где символом $\delta(x)$ обозначена дельта-функция Дирака. При этом должны выполняться условия трансверсальности

$$\begin{cases} m(y) R_{11}(t_1, x, 0) = \delta(x - y), m(y) R_{12}(t_1, x, 0) = 0, \\ m(y) R_{21}(t_1, x, 0) = 0, m(y) R_{22}(t_1, x, 0) = \delta(x - y). \end{cases} \quad (15)$$

и краевые условия

$$\begin{cases} EJ(0) \frac{\partial^2 R_{12}(t, x, 0)}{\partial y^2} = 0, \quad \frac{\partial}{\partial y} \left(EJ(0) \frac{\partial^2 R_{12}(t, x, 0)}{\partial y^2} \right) = 0, \\ EJ(l) \frac{\partial^2 R_{12}(t, x, l)}{\partial y^2} = 0, \quad \frac{\partial}{\partial y} \left(EJ(l) \frac{\partial^2 R_{12}(t, x, l)}{\partial y^2} \right) = 0, \\ EJ(0) \frac{\partial^2 R_{22}(t, x, 0)}{\partial y^2} = 0, \quad \frac{\partial}{\partial y} \left(EJ(0) \frac{\partial^2 R_{22}(t, x, 0)}{\partial y^2} \right) = 0, \\ EJ(l) \frac{\partial^2 R_{22}(t, x, l)}{\partial y^2} = 0, \quad \frac{\partial}{\partial y} \left(EJ(l) \frac{\partial^2 R_{22}(t, x, l)}{\partial y^2} \right) = 0 \end{cases} \quad (16)$$

Принимая во внимание вышеизложенные рассуждения, приходим к следующему выводу.

Теорема 2. Для нахождения функций $R_{11}(t, x, y)$, $R_{12}(t, x, y)$, $R_{21}(t, x, y)$, $R_{22}(t, x, y)$ требуется решить систему уравнений (11)–(14) с учетом краевых условий (16) и условий трансверсальности (15).

Зная функции $R_{21}(t, x, y)$ и $R_{22}(t, x, y)$, можно найти оптимальное управление $u(t, x)$.

Теорема 3. Для оптимального управления $u(t, x)$ справедлива формула

$$u(t, x) = - \int_0^l m(y) \left[R_{21}(t, x, y) z(t, y) + R_{22}(t, x, y) \frac{\partial z(t, y)}{\partial t} \right] dy.$$

Дальнейшее развитие полученных в статье результатов состоит в исследовании случая, когда время регулирования стремится к бесконечности. В теории оптимального управления эта проблема именуется задачей аналитического конструирования регулятора. Перспективным исследованием также является игровая постановка задачи, чтобы воспользоваться методом разрешающих функций.

Выводы.

Проведено исследование линейно-квадратической задачи оптимального управления процессом колебаний балки переменного поперечного сечения в случае свободных концов балки. Для рассматриваемой задачи оптимизации получены необходимые условия оптимальности. Анализ этих условий дал возможность вывести систему интегро-дифференциальных уравнений Риккати с частными производными. Решение этой системы использовано при построении формулы для вычисления оптимального управления.

1. Бублик Б.Н., Кириченко Н.Ф. Основы теории управления. — К.: Вища школа. — 1975. — 328 с.
2. Naidu D.S. Optimal control systems//Electrical engineering textbook series. — CRC PRESS — Boca Raton London — New York — Washington, D. C. — 2003. — 433 p.
3. Боднер В.А. Теория автоматического управления полетом. — М.: Наука. — 1964. — 700 с.
4. Копец М.М. Оптимальное управление колебаниями прямоугольной мембраны//Кибернетика и вычисл. техника. — 2014. — Вып. 177. — С. 28–42.

UDC 517.977

OPTIMAL CONTROL BY VIBRATIONS OF THE BEAM WITH VARIABLE CROSS-SECTION

M.M. Kopets

National Technical University of Ukraine «Kyiv Polytechnic Institute» (Kiev)

Introduction. The last half-century is characterized by the rapid development of technology. Significant progress has been made in the rocket, aircraft, shipbuilding and space technology, etc. All sectors have oscillatory processes. In some cases, they can usefully be taken into account to improve the quality of the process, while others, on the contrary, it is necessary to suppress because of their negative impact on the final process. This means that the oscillatory processes must not only be learned, but also be able to manage them effectively. Similar problems effectively manage mechanical processes just studying optimal control theory. The purpose of this article is to study the linear-quadratic problem of optimal control by oscillations of the beam with variable cross-section in the case of the free ends of the beam.

Statement of the Problem. The state equation is linear partial differential equation of the fourth order of hyperbolic type with given initial conditions and homogeneous boundary conditions. Quality of the process is estimated by quadratic functional. The admissible control is such a function which belongs to the class of square Lebesgue integrable functions. Optimal control is admissible control which is implemented at least the cost functional.

The purpose of the paper is to determine the necessary conditions for optimal control of process vibrations of a beam of variable cross-section in the case of the

free ends of the beam and to give solution of integral-differential Riccati equations for the optimal control.

The main results. Necessary optimality conditions for the considered optimization problem are obtained. Analysis of these conditions made it possible to bring the system of integro-differential Riccati equations with partial derivatives. The solution of this system is used in the construction of an explicit formula for the calculation of optimal control.

Conclusions. The article investigates the linear-quadratic optimal control process vibrations of a beam of variable cross-section in the case of the free ends of the beam. Necessary optimality conditions for the considered optimization problem are obtained. Analysis of these conditions made it possible to bring the system of integro-differential Riccati equations with partial derivatives. The solution of this system is used in the construction of an explicit formula for the calculation of optimal control. Further development of the obtained results is to study the case where the control time tends to infinity. In the theory of optimal control, this problem is called the problem of analytical construction of the regulator.

Keywords: linear quadratic optimal control problem, method of Lagrange multipliers, necessary optimality conditions, oscillations of the beam, partial derivatives, system of integro-differential equations.

1. Bublik B.N., Kirichenko N.F. Fundamentals of control theory. — K. : Higher School — 1975. — 328 p.
2. Naidu D.S. Optimal control systems // *Electrical engineering textbook series.* — CRC PRESS — Boca Raton London — New York — Washington, D. C. — 2003. — 433 p.
3. Bodner V.A. Theory of automatic flight control. — M. : Nauka — 1964. — 700 p.
4. Kopets M.M. Optimal control of vibrations of a rectangular membrane // *Cybernetics and computer engineering.* — 2014. — Vol. 177. — P. 28–42.

Получено 10.07.2015

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ КРАТКОСРОЧНОЙ И СРЕДНЕСРОЧНОЙ АДАПТАЦИИ СИСТЕМЫ ДЫХАНИЯ ЛИЦ, РАБОТАЮЩИХ В ЭКСТРЕМАЛЬНЫХ УСЛОВИЯХ ВЫСОКОГОРЬЯ

Н.И. Аралова

Институт кибернетики им. В.М.Глушкова НАН Украины

Предлагается математическая модель краткосрочной и среднесрочной адаптации системы дыхания лиц, находящихся под сочетанным воздействием гипобарической и гиперметаболической гипоксии. Приведены результаты прогнозирования состояния функционирования системы дыхания по напряжению кислорода в крови и тканях работающих органов у спасателей при их адаптации к условиям горных метеофакторов.

Ключевые слова: краткосрочная адаптация, среднесрочная адаптация, система дыхания, гипобарическая (гипоксическая) гипоксия, надежность, математическая модель системы дыхания.

Запропоновано математичну модель короткотермінової та середньотермінової адаптації системи дихання осіб, які перебувають під сумісним впливом гіпобаричної та гіперметаболічної гіпоксії. Наведено результати прогнозування стану функціонування системи дихання за напруженням кисню у крові та тканинах працюючих органів у рятувальників при їх адаптації до умов гірських метеофакторів.

Ключові слова: короткотермінова адаптація, середньотермінова адаптація, система дихання, гіпобарична (гіпоксична) гіпоксія, надійність, математична модель системи дихання.

ВВЕДЕНИЕ

Успешное выполнение человеком тяжелой работы в условиях гипобарической гипоксии во многом зависит от функционального состояния, подготовленности, особенностей физиологического и психофизиологического статусов. Объективную оценку степени подготовленности человека к выполнению высоких нагрузок в экстремальных условиях высокогорья дают данные медицинских и физиологических обследований. Наряду с экспериментальными исследованиями в последнее время широко используются методы математического моделирования отдельных функциональных систем и целостного организма в тех или иных жизненных ситуациях, результаты которого дополняют систему экспериментальных данных и позволяют сделать более полную оценку функционального состояния организма.

Известно, что в условиях гипобарической гипоксии при выполнении нагрузки большой мощности человек тратит значительное количество энергии (механической и нервной), что приводит к развитию гипоксических состояний — кислородной недостаточности в работающих мышцах и тканях (скелетных, сердечной и нервной, тканях мозга), которые могут существенно ограничить его работоспособность. Одним из подходов для оценки степени развития гипоксии в отдельных группах тканей и в целостном организме

является использование математической модели функционирования системы дыхания (ФСД) [1].

Эта модель описывает транспорт и массообмен респираторных газов в различных звеньях системы дыхания — дыхательных путях, альвеолярном пространстве легких, крови легочных и тканевых капилляров, артериальной и смешанной венозной крови, тканевых резервуарах и органах (мозг, сердце, дыхательные и скелетные мышцы, другие ткани и органы), и компенсирующие воздействия механизмов саморегуляции (величина вентиляции легких, минутный объем системного и органных (тканевых) кровотоков, стабилизирующих функциональные состояния организма при заданном уровне его функциональной активности).

По существу, модель ФСД представляет собой управляемую динамическую систему, состояние которой определяется в каждый момент времени напряжениями кислорода и углекислоты в каждом структурном звене системы дыхания (альвеолах, крови и тканях). Управление (саморегуляция) состоянием при постоянно или на заданном временном отрезке действующем возмущении (высокая функциональная активность отдельных групп тканей) осуществляется исполнительными органами саморегуляции — дыхательными мышцами, формирующими необходимый уровень вентиляции для компенсации возникающих гипоксических состояний, сердечной мышцей, обеспечивающей минутный объем крови (МОК), и гладкими мышцами тканевых сосудов, вазодилатация и вазоконстрикция которых способствует распределению системного кровотока по органам и тканям.

Кроме этих активных механизмов саморегуляции в модели учтены пассивные механизмы: концентрация гемоглобина в крови, миоглобина в скелетных и сердечной мышцах, их возможности к оксигенации, концентрация буферных оснований в крови и др. Предполагается, что решение о выборе величин компенсирующих воздействий принимается центром принятия решения на основании информации об уровне функциональной активности и степени кислородной недостаточности, избыточности накопления углекислоты во всех тканевых регионах организма, передается на исполнительные органы саморегуляции, повышает их функциональную активность, чем обеспечивается выполнение основной функции дыхания.

Цель работы – на математической модели функционирования системы дыхания исследовать механизмы адаптации этой системы к условиям горных метеофакторов у лиц, выполняющих тяжелую физическую нагрузку в условиях гипобарической гипоксии.

ПОСТАНОВКА ЗАДАЧИ И МЕТОДЫ

Необходимой предпосылкой высокой надежности функционирования организма в различных условиях жизнедеятельности является адаптация к гипоксии и обеспечивающие ее механизмы [2]. Для математического анализа адаптационных возможностей организма при гипоксии различной этиологии используется модель ФСД [1], описывающая транспорт и массообмен

респираторных газов в дыхательных путях, альвеолярном пространстве, крови и тканях обыкновенными нелинейными дифференциальными уравнениями, регуляция осуществляется на основе компромиссного разрешения конфликтной ситуации, возникающей между тканями и органами при борьбе за кислород в условиях его дефицита [3].

Анализ математической модели ФСД показал, что кратковременные возмущения системы, приводящие к возникновению гипоксии, могут быть компенсированы реакцией механизмов саморегуляции, сложившихся в процессе эволюции — интенсификацией работы системы внешнего дыхания, сердечной мышцы и гладких мышц сосудов.

При среднесрочной адаптации к гипоксии формируются функциональные механизмы, меняющие чувствительность организма к гипоксии, повышающие эффективность тканевого массообмена, стимулирующие эритропоэз.

Долгосрочная адаптация организма к действию гипоксии заключается в структурных изменениях в организме, например, перераспределение массы тканей, гипертрофия исполнительных органов саморегуляции процесса дыхания, в первую очередь гипертрофия левого желудочка.

ПОСТРОЕНИЕ МОДЕЛИ КРАТКОСРОЧНОЙ АДАПТАЦИИ

Математическая модель ФСД дает исследователю возможность: проанализировать кислородные и углекислотные режимы организма в динамике при различных уровнях функциональной нагрузки и при различных условиях окружающей среды; сформировать такие режимы системы внешнего дыхания, которые способствуют увеличению запасов кислорода в организме и тем самым повышают ресурс сердечной мышцы при регуляции гипоксических состояний, возникающих при сочетанном воздействии гипобарической гипоксии и гиперметаболической гипоксии; прогнозировать состояние организма в различных физических усилиях и оценить эффективность процесса подготовки; планировать и распределить тяжелые нагрузки с учетом функциональных возможностей данного индивидуума и в зависимости от складывающихся ситуаций.

Процесс дыхания, при котором происходит транспорт и массообмен респираторных газов, рассматривается как управляемая динамическая система, которая описывается системой дифференциальных уравнений и алгебраических соотношений. Управляемыми параметрами являются вентиляция \dot{V} , системный кровоток Q и локальные кровотоки Q_{t_j} , $j = \overline{1, m}$, т.е. исполнительными органами регуляции являются сердечная и дыхательные мышцы, гладкие мышцы сосудов.

Действующие на систему дыхания возмущения подразделяются на внутренние и внешние. К внешним относятся изменение состава вдыхаемого воздуха, барометрического давления окружающей среды и т.д., к внутренним — изменение интенсивности обменных процессов в органах и тканях. Количественно вторые характеризуются скоростью потребления кислорода

$q_{t_j}O_2$, $j = \overline{1, m}$ и выделения углекислого газа $q_{t_j}CO_2$, $j = \overline{1, m}$. Роль краткосрочной адаптации состоит в выведении возмущенной динамической системы транспорта и массообмена респираторных газов в некоторое устойчивое для сформированных условий жизнедеятельности организма стационарное состояние [4].

Зададим:

— начальное состояние системы: $p_{RP}O_2$, $p_{RP}CO_2$, p_AO_2 , p_ACO_2 , характеризующие парциальные давления кислорода и углекислого газа в дыхательных путях и альвеолярном пространстве, $p_{LC}O_2$, $p_{LC}CO_2$, p_aO_2 , p_aCO_2 , $p_{c_j}O_2$, $p_{c_j}CO_2$, $p_{t_j}O_2$, $p_{t_j}CO_2$, $p_{\bar{v}}O_2$, $p_{\bar{v}}CO_2$, $j = \overline{1, m}$, характеризующие напряжения кислорода и углекислого газа в крови легочных капилляров, артериальной крови, тканях, смешанной венозной крови в момент времени τ_0 начала действия возмущения;

— области изменения параметров управления:

$$\left. \begin{aligned} \dot{V}_{\min} \leq \dot{V} \leq \dot{V}_{\max} \\ Q_{\min} \leq Q \leq Q_{\max} \\ Q_{c_j \min} \leq Q_{c_j} \leq Q_{c_j \max}, j = \overline{1, m} \\ \sum_{j=1}^m Q_{c_j} \end{aligned} \right\}, \quad (1)$$

— терминальное множество состояний, обусловленное соотношениями:

$$\left. \begin{aligned} |G_{t_j}O_2 - q_{t_j}O_2| \leq \varepsilon_{t_j}, j = \overline{1, m} \\ |G_{t_j}CO_2 + q_{t_j}CO_2| \leq \varepsilon_{t_j}, j = \overline{1, m} \end{aligned} \right\}, \quad (2)$$

где $\varepsilon_{t_j}O_2$, $\varepsilon_{t_j}CO_2$, $j = \overline{1, m}$ достаточно малые положительные величины.

Решением задачи краткосрочной адаптации, сформулированной таким образом, будет любой набор значений управляющих параметров \dot{V} , Q , Q_{c_j}

$j = \overline{1, m}$ из уравнения (1), т.к. именно эти параметры через некоторое время переведут возмущенную систему в состояние, характеризующееся условиями (2). При этом степень недостатка кислорода или накопления углекислого газа будут достоверными. Представим задачу краткосрочной адаптации как задачу оптимальной саморегуляции. Предположим, что оптимальным является набор параметров управления \dot{V} , Q , Q_{c_j} , $j = \overline{1, m}$ из уравнения (1), обеспечивающих на траекториях движения возмущенной динамической системы минимум функционала

$$I = \int_{\tau}^{\tau_0+T} \left[\rho_1 \sum_{i=1}^m \lambda_{t_j} (G_{t_j}O_2 - q_{t_j}O_2)^2 + \rho_2 \sum_{i=1}^m \lambda_{t_j} (G_{t_j}CO_2 + q_{t_j}CO_2)^2 \right] d\tau,$$

где $G_i O_2, G_i CO_2$ — соответственно потоки кислорода и углекислого газа через капиллярно–тканевый барьер; $q_i O_2, q_i CO_2$ — скорость утилизации кислорода и образования углекислого газа в i -том тканевом регионе; ρ_1 и ρ_2 — коэффициенты чувствительности организма к недостатку кислорода и избытку углекислого газа соответственно; λ_{t_i} — коэффициенты, характеризующие жизненную значимость каждого органа или региона, их морфо-функциональные особенности. При расчетах было принято, что

$$\lambda_{t_j} = \varphi \left(\frac{V_{c_j}}{V_{t_j}} \right), j = \overline{1, m}.$$

Квадратичная функция φ характеризует степень кровенаполнения единичного объема тканевого резервуара.

ПОСТРОЕНИЕ МОДЕЛИ СРЕДНЕСРОЧНОЙ АДАПТАЦИИ

Саморегуляция системы дыхания осуществляется не только при краткосрочной адаптации, но и на этапах среднесрочной и долгосрочной адаптации, когда возмущения действуют на систему на протяжении длительного времени или периодически повторяются. Это приводит к развитию добавочных адаптационных механизмов, которые позволяют в ответ на возмущение осуществить более эффективную организацию обменных функций в тканях.

При утилизации кислорода в тканях высвобождается энергия, необходимая для работы мышц, опорно-двигательного аппарата, поддержку основных функций органов человека. Часть энергии выделяется в виде тепла. Скорость потребления кислорода тканями можно представить в виде [5]:

$$q_{t_j} O_2 = q_{t_j}^f O_2 + q_{t_j}^T O_2, j = \overline{1, m},$$

где $q_{t_j}^f O_2$ — скорость утилизации кислорода, необходимого для выполнения на заданном уровне функций органов и тканей; $q_{t_j}^T O_2$ — составляющая скорости потребления кислорода, обеспечивающая выделение тепловой и других видов энергии.

На среднесрочном этапе адаптации $q_{t_j}^f O_2 = const, j = \overline{1, m}$ для заданного уровня нагрузки, а $q_{t_j}^T O_2, j = \overline{1, m}$ может быть уменьшена в процессе адаптации за счет лучшей организации обменных процессов.

Предполагается, что

$$q_{t_j}^T O_{2\text{адапт.}} = \left(q_{t_j}^T O_{2\text{неадапт.}} - q_{t_j}^T O_{2\text{крит.}} \right) e^{-k\tau} + q_{t_j}^T O_{2\text{крит.}}, j = \overline{1, m},$$

где $q_{t_j}^T O_{2\text{адапт.}}$, $q_{t_j}^T O_{2\text{неадапт.}}$ — тепловые составляющие скорости потребления кислорода в адаптированном и неадаптированном организме соответственно; $q_{t_j}^T O_{2\text{крит.}}$ — скорость потребления кислорода, необходимая для выделения минимального количества энергии для поддержания теплового баланса организма при адаптации; k — заданный коэффициент; τ — скорость процесса адаптации.

Аналогично можно записать выражения для углекислого газа.

При среднесрочной адаптации изменяются коэффициенты чувствительности к гипоксии и гиперкапнии:

$$\rho_{1\text{адапт.}} = (\rho_{1\text{неадапт.}} - \rho_{1\text{крит.}})e^{-k\tau} + \rho_{1\text{крит.}},$$

$$\rho_{2\text{адапт.}} = (\rho_{2\text{неадапт.}} - \rho_{2\text{крит.}})e^{-k\tau} + \rho_{2\text{крит.}},$$

где $\rho_{1\text{крит.}}$, $\rho_{2\text{крит.}}$ — минимальные коэффициенты чувствительности, обеспечивающие гипоксическую и гиперкапническую стимуляцию при работе механизмов краткосрочной адаптации.

Состояние динамической системы, которая представлена в модели, определяется уровнем напряжений кислорода (pO_2) и углекислоты (pCO_2) в крови и тканевых регионах. Таким образом, в процессе моделирования формируются кислородные и углекислотные портреты организма при различной интенсивности функциональной деятельности мышц.

Предусмотрена индивидуализация модели. Учитывается вес, рост, структура мышц. В модель введены коэффициенты чувствительности организма к гипоксии и избытку углекислоты. Для каждого индивидуума они различны и зависят от степени адаптации организма к физическим нагрузкам и состояния его психофизиологического статуса. Опыт использования модели показывает, что уменьшение этих коэффициентов в тренировочном процессе свидетельствует об его эффективности. Однако эти коэффициенты должны быть больше некоторых пороговых значений, ведущих к разрегулируемости (неуправляемости) динамической системы и, как следствие, к развитию патологических процессов. В покое и при нагрузке определяются уровень вентиляции, МОК, напряжения кислорода в артериализированной крови, общее потребление кислорода и выделение углекислоты. Выбор коэффициентов чувствительности осуществляется таким образом, чтобы вентиляция и МОК при данном выборе, определенные в результате моделирования проведенного эксперимента, совпадали с данными, полученными в эксперименте.

ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Для расчетов использовались данные, полученные сотрудниками Эльбрусской медико-биологической станции при обследовании группы спасателей, работающих в Приэльбрусье. Обследование проводилось в покое

на велоэргометре на высоте 2100 м и в барокамере на «высоте» 5600 м над уровнем моря дважды на 2-е и 10-е сутки адаптации, т.е. исследовалась краткосрочная и среднесрочная адаптация тренированных лиц в условиях гипобарической гипоксии. При этом, между обследованиями, спасатели совершили восхождение на Эльбрус. Определялись минутный объем дыхания и кровообращения, содержание респираторных газов в выдыхаемом и альвеолярном воздухе, частота сердечных сокращений, частота дыхания, количество гемоглобина и насыщение крови кислородом. Полученные данные использовались для расчета параметров кислородных режимов организма и построения модельной характеристики системы дыхания [5]. Далее данные, полученные в результате физиологического обследования, и данные, полученные в результате моделирования в статике, вводились в качестве исходных в имитационную модель функционирования системы дыхания и осуществлялись расчеты показателей, характеризующих функциональное состояние системы дыхания в динамике дыхательного цикла. Расчеты выполнялись на модели функциональной системы дыхания с четырьмя тканями (Рис. 1). Результаты расчетов приведены в таблицах 1 и 2.

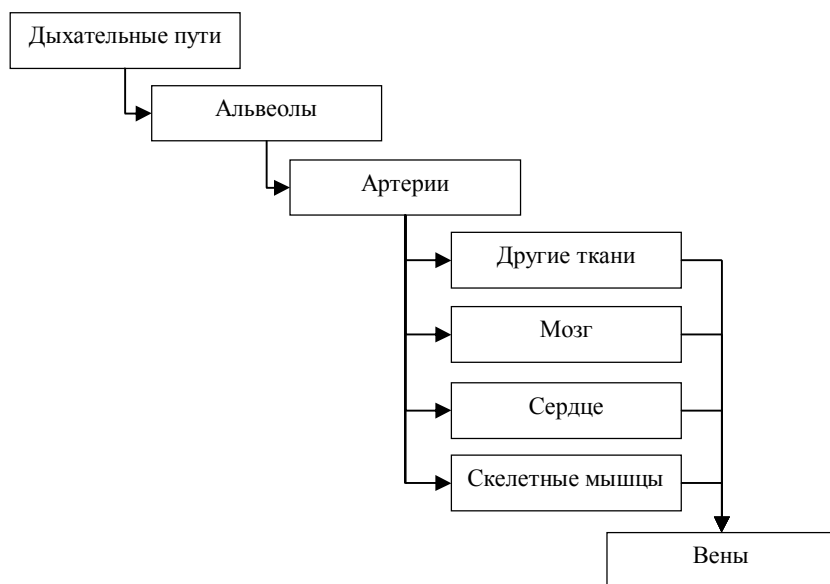


Рис.1. Структурно-функциональная схема подсистем, участвующих в процессе адаптации системы дыхания к условиям высокогорья

Выше отмечалось, что исполнительными органами саморегуляции основной функции дыхания являются дыхательные мышцы, сердечная и гладкие мышцы тканевых сосудов. Человек сознательно способен управлять только работой дыхательных мышц, формировать необходимый уровень вентиляции. Центр принятия решений при регуляции системы дыхания как бы раскладывает ресурс регуляции на все три исполнительных органа. Поэтому, задавая более интенсивный режим внешнего дыхания, снимается нагрузка в сердечной мышце и тем самым увеличивается ее регуляторный ресурс. Для увеличения кислородного запаса тканей и регуляторного ресурса сердца желательно формировать соответствующие режимы системы

внешнего дыхания, определяя оптимальный дыхательный объем, продолжительность фаз вдоха и выдоха для каждого возможного уровня функциональной активности. Особенно это важно делать при выполнении работы в экстремальных условиях высокогорья, т.е. при сочетанном воздействии гипобарической и гиперметаболической гипоксии.

РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ЭКСПЕРИМЕНТОВ И ИХ ОБСУЖДЕНИЕ

Расчёты показали, что даже в начальный период адаптации напряжение кислорода в артериальной крови у спасателей составляло $79,7 \pm 3$ мм рт.ст., а, следовательно, уровень тканевой гипоксии не был глубоким: в мозге $30,91 \pm 2$ мм рт.ст., в тканях сердца — $25,15 \pm 2$ мм рт.ст., в скелетных мышцах — $25,65 \pm 2,3$ мм рт.ст., в тканях других органов в среднем — $39,23 \pm 2,8$ мм рт.ст., что обусловило, в свою очередь, не очень высокий уровень венозной гипоксемии — $34,5 \pm 2,31$ мм рт.ст.

Имитация покоя на высоте 5600 м в период краткосрочной адаптации (первое обследование) выявила, что, несмотря на развитие артериальной гипоксемии, не произошло углубления тканевой гипоксии: pO_2 в тканях мозга составляло $30,28 \pm 2,28$ мм рт.ст., в сердце — $26,88 \pm 3,31$ мм рт.ст., в скелетных мышцах — $27,8 \pm 1,94$ мм рт.ст., в тканях других органов в среднем $38,06 \pm 4,27$ мм рт.ст., в венозной крови напряжение кислорода составило $34,04 \pm 3,85$ мм рт.ст.

Второе обследование было проведено после тренировочного восхождения на Эльбрус, т.е. в процессе среднесрочной адаптации обследуемые испытали сочетанное воздействие гипобарической гипоксии и гиперметаболической гипоксии. Расчёты показали, что на высоте 2100 м снизился уровень артериальной гипоксемии с $80,04 \pm 3$ мм рт.ст. до $84,18 \pm 2,54$ мм рт.ст. Также повысилось напряжение кислорода в тканях головного мозга до $33,3 \pm 0,8$ мм рт.ст., что практически составляет нормальный уровень pO_2 для условий нормоксии. В тканях сердца pO_2 возросло примерно на $2,75 \pm 0,4$ мм рт.ст. и составило $27,9 \pm 1,63$ мм рт.ст., в скелетных мышцах pO_2 выросло на $1,64 \pm 0,23$ мм рт.ст. и достигло уровня $27,29 \pm 1,46$ мм рт.ст., в тканях других органов произошло увеличение pO_2 с $39,23 \pm 2,8$ мм рт.ст. до $41,5 \pm 3,06$ мм рт.ст. Соответственно, несколько повысилось напряжение кислорода в смешанной венозной крови с $34,5 \pm 2,31$ мм рт.ст. до $36,46 \pm 1,97$ мм рт.ст.

Обследование на высоте 5600 м выявило, что, благодаря адаптационным процессам, немного снизился уровень артериальной гипоксемии по сравнению с первым обследованием: pO_2 в артериальной крови выросло до $57,88 \pm 3,42$ мм рт.ст., или на $3,28 \pm 2,6$ мм рт.ст. Увеличение напряжения кислорода в тканях головного мозга было недостоверным (на $0,63 \pm 0,04$ мм рт.ст.), в сердечной мышце произошло снижение pO_2 на $1,38 \pm 0,24$ мм рт.ст., в скелетных мышцах pO_2 практически не изменилось (увеличилось на $0,18 \pm 0,03$ мм рт.ст.), в тканях других органов напряжение кислорода увеличилось на $1,66 \pm 0,32$ мм рт.ст., в смешанной венозной

крови — возросло на $1,24 \pm 0,36$ мм рт.ст.

Таблица 1.

Результаты расчетов парциального давления и напряжения кислорода в крови и тканях работающих органов у лиц, выполняющих тяжелую физическую работу в условиях среднегорья на 2-е сутки адаптации

	Высота	Легкие	Аль- веолы	Арте- рия	Ткани				Вена
					мозг	сердце	Скелет- ные мышцы	Другие	
1	2100	87,53	82,97	82,94	28,39	28,26	23,74	36,93	32,45
	5600	61,78	60,49	55,49	25,4	29,13	32,31	43,2	38,87
2	2100	88,9	84,7	84,24	33,87	25,39	27,41	41,59	36,61
	5600	69,16	69,09	64,9	27,54	34,84	22,96	34,67	30,43
3	2100	85,43	80,38	79,49	30,76	23,41	25,6	39,15	34,37
	5600	68,62	68,31	61,84	34,21	26,41	29,88	42,15	37,39
4	2100	84,23	79,5	79,97	32,91	26,01	26,93	40,81	36,03
	5600	62,88	59,82	44,54	22,9	18,34	23,51	32,31	28,94
5	2100	77,96	72,67	71,85	29,42	22,67	24,59	37,67	33,1
	5600	61,44	58,9	46,27	30,74	25,7	30,36	37,99	34,52
6	2100	97,72	97,58	85,11	25,67	29,31	22,7	35,19	30,85
	5600	66,65	64,95	60,15	38,37	30,58	32,84	45,56	40,6
Ср.	2100	86,96	82,97	80,6	30,91	27,15	28,64	39,23	34,5
	5600	65,09	63,59	55,53	31,63	27,5	28,64	39,31	35,13

Таблица 2.

Результаты расчетов парциального давления и напряжения кислорода в крови и тканях работающих органов у лиц, выполняющих тяжелую физическую работу в условиях среднегорья на 10-е сутки адаптации

	Высота	Легкие	Аль- веолы	Арте- рия	Ткани				Вена
					мозг	сердце	Скелет- ные мышцы	Другие	
1	2100	90,12	87,97	86,05	33,26	26,51	26,81	40,96	36,11
	5600	66,07	64,6	58,21	25,4	20,88	23,18	34,5	30,52
2	2100	94,52	93,47	89,41	36,6	27,94	29,04	44,05	38,88
	5600	66,2	64,48	58,33	24,06	22,71	23,04	34,51	30,33
3	2100	90,0	85,36	84,75	36,54	27,02	29,13	43,98	38,73
	5600	67,88	66,55	60,58	31,63	24,23	27,82	41,77	36,64
4	2100	89,95	86,67	82,76	25,05	31,98	22,98	36,86	30,95
	5600	65,07	63,58	57,83	36,76	29,65	32,43	44,24	39,63
5	2100	81,83	78,09	77,92	35,03	26,07	28,45	42,68	37,63
	5600	62,07	61,22	54,55	36,69	30,06	33,41	43,52	39,31
6	2100	81,13	81,58	81,54	31,0	25,42	25,74	39,44	34,71
	5600	65,74	65,25	54,0	23,54	19,07	22,63	33,32	29,45
Ср.	2100	87,93	85,44	83,74	32,92	27,49	27,03	41,16	36,17
	5600	65,5	64,28	57,23	29,68	24,43	27,09	38,65	34,31

Результаты данного обследования свидетельствуют о том, что практически у всех обследуемых при гипобарической гипоксии имеются резервы физической мощности и они могут выполнить значительную физическую нагрузку, т.к. напряжение кислорода в скелетных мышцах и мышце сердца еще весьма далеко от критических уровней. Гораздо ниже запас прочности у тканей мозга, что может привести, в частности, к нарушению координации и затруднить принятие адекватных решений. Следовательно, лицам, подвергающимся совместному воздействию гипобарической гипоксии, гиперметаболической гипоксии и работающим в

условиях повышенной ситуационной напряженности, необходимо обратить особое внимание на способы обеспечения психофизиологической работоспособности и быстрого восстановления после физического утомления.

Выводы

Представлен срез функционального состояния лиц, подвергающихся сочетанному воздействию гипобарической гипоксии и гиперметаболической гипоксии, характеризующий слабые звенья их организма.

Данная работа позволила наметить ряд задач, которые необходимо разрешить для оценки функционального состояния организма человека в процессе его краткосрочной и среднесрочной адаптации. Для того, чтобы оценить состояние функциональных систем организма, необходимо провести обследование системы дыхания при физической нагрузке и при нагрузке в условиях повышенного ситуационного напряжения при имитации в барокамере различных внешних возмущающих воздействий. При этом исследование психофизиологических функций целесообразно проводить дважды — до выполнения стандартной нагрузки и непосредственно после.

Формирование соответствующих режимов системы внешнего дыхания с учетом оптимального дыхательного объема, продолжительности фаз вдоха и выдоха для каждого возможного уровня функциональной активности содействует увеличению кислородного запаса тканей и регуляторного ресурса сердца, что особенно значимо при выполнении работы в экстремальных условиях высокогорья.

1. Онопчук Ю.Н. Гомеостаз функциональной системы дыхания как результат внутрисистемного и системно-средового информационного взаимодействия //Биоэкология. Единое информационное пространство — К. —2001. — С. 59–81.
2. Онопчук Ю.Н., Белошицкий П.В., Аралова Н.И. К вопросу о надежности функциональных систем организма // Кибернетика и вычислительная техника. — 1999. — Вып. 122. — С. 72–89.
3. Полинкевич К.Б., Онопчук Ю.Н. Конфликтные ситуации при регулировании основной функции системы дыхания организма и математические модели их разрешения// Кибернетика. — 1986. — № 3. — С. 100–104.
4. Білошицький П.В., Ключко О.М., Онопчук Ю.М. Результати дослідження проблем адаптації українськими вченими в Приельбруссі//Вісн. НАУ. — 2008. — № 1. — С. 102–108.
5. Онопчук Ю.М., Білошицький П.В., Ключко О.М. Створення математичних моделей за результатами досліджень українських вчених на Ельбрусі//Вісн. НАУ. — 2008. — № 3. — С. 146–155.

MATHEMATICAL MODEL OF THE SHORT- AND MEDIUM-TERM ADAPTATION OF RESPIRATORY SYSTEM OF THE PERSONS WORKING IN EXTREME CONDITIONS OF HIGH MOUNTAINS

N.I. Aralova

*V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences
of Ukraine (Kiev)*

Introduction. In addition to experimental studies in recent years the methods of mathematical modeling of individual functional systems and the whole organism in certain situations are widely used, the results of which complement the system of experimental data and allow to make a more complete assessment of the functional state of the organism.

Purpose. To explore on a mathematical model of the respiratory system the functional mechanisms of adaptation of the respiratory system to the conditions of mountain meteorological factors for persons performing heavy exercise in a hypobaric hypoxia.

Results. The model, that describes transport and mass exchange of respiratory gases in the respiratory tract, the alveolar space, blood and tissues with use of ordinary nonlinear differential equations, for the mathematical analysis of the adaptive capacity of the organism hypoxia of various etiologies is used. The regulation is based on a compromise resolution of conflicts arising between the tissues and organs in the struggle for oxygen in a deficit. On the basis of this model, the models of short time and medium adaptation persons performing heavy physical activity in a midlands are created. Results of simulation experiment are presented.

Conclusion. The article presents a mathematical model of short-term and medium term adaptation FRS for rescuers and the results of the numerical analysis of this model. On this basis, the practical recommendations for the selection of the persons, that are exposed to the combined effects of hypobaric hypoxia and the hypermetabolic hypoxia, are given.

Keywords: short-term adaptation, medium term adaptation, respiration system, hypobaric hypoxia, hypermetabolic hypoxia, reliability, mathematical model of respiratory system.

1. Onopchuk Yu.N. Homeostasis function of the respiratory system as a result of in-system and system-environment interaction // *Bioekomedicine. Unified Information Space* — Kiev — 2001. — P.59–81(in Russian)
2. Onopchuk Y.N., Beloshitsky P.V. Aralova N.I. On the question of reliability of functional systems // *Cybernetics and computing tehnics*. 1999. — Vol. 122. — P. 72–89 (in Russian)
3. Polinkevich K.B., Onopchuk Y.N. Conflicts in the regulation of the main function of the respiratory system of the body, and mathematical models of their solution // *Cybernetics*. — 1986. — № 3. — P. 100–104.
4. Biloshitsky P.V. Klyuchko O.M., Onopchuk Y.N. Research results of the problems of adaptation by Ukrainian scientists on Elbrus // *Visn. NAU*. — 2008. — № 1. — P. 102–108.
5. Onopchuk Y.M., Biloshitsky P.V. Klyuchko O.M. Creation of mathematical models for the research by Ukrainian scientists on Elbrus. // *Visn. NAU*. — 2008. — № 3. — P. 146–155.

Получено 07.09.2015

Медицинская и биологическая кибернетика

УДК 519.254:616-036.22

МЕТОД ОПРЕДЕЛЕНИЯ ЭФФЕКТОВ СОВМЕСТНОГО ДЕЙСТВИЯ ФАКТОРОВ СРЕДЫ В ЭПИДЕМИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЯХ ДЛЯ БИНАРНЫХ ПЕРЕМЕННЫХ

М.Ю. Антомонов

*ГУ «Институт гигиены и медицинской экологии им. А.Н. Марзеева
НАМН Украины»*

Рассмотрены существующие подходы расчета эффектов совместного действия факторов, использующиеся в медико-экологических исследованиях. Оценены недостатки этих подходов, реализованных в теории вероятностей и математической статистике. Предложена оригинальная методика расчета совместного действия факторов, которая позволяет оперировать с качественной информацией, выраженной в бинарном виде. При конструировании итогового выражения использованы подходы формальной логики и теории вероятностей. Полученная функция представляет собой обобщение формулы расчета вероятности независимых событий с учетом их возможной взаимозависимости. Представлен пример использования этого метода.

Ключевые слова: совместное действие факторов, вероятности независимых и взаимозависимых событий, качественные данные, бинарные переменные.

Розглянуто сучасні підходи розрахунку ефектів сумісної дії факторів, які використовуються в медико-екологічних дослідженнях. Оцінено недоліки цих підходів, реалізованих в теорії ймовірностей і математичній статистиці. Запропоновано оригінальну методіку розрахунку спільної дії факторів, яка дозволяє оперувати з якісною інформацією, що виражена у бінарному вигляді. При конструюванні остаточного виразу використано підходи формальної логіки і теорії ймовірностей. Одержана функція є узагальненням формули розрахунку ймовірності незалежних подій з урахуванням їх можливої взаємозалежності. Надано приклад використання цього методу.

Ключові слова: сумісна дія факторів, ймовірності незалежних і взаємозалежних подій, якісні дані, бінарні змінні.

ВВЕДЕНИЕ

В последнее время в прикладных направлениях математического анализа данных все больший приоритет приобретает практическая направленность методов. Например, такое мощное течение как *Data Mining* гораздо в большей степени ориентировано на практическое приложение полученных результатов, чем на следование математическим канонам [1]. Технология *Data Mining* не зависит от строгого соответствия массивов данных

статистическим критериям или математическому описанию функциональных связей между переменными. Объединяя количественный и качественный анализ данных, технология нацелена на получение новых знаний и принятие решений в условиях многомерности данных и неопределенности их структуры.

В дополнение к классическому регрессионному анализу, предполагающему количественный характер переменных и «нормальность» их распределений, в прикладных исследованиях активно применяются логистические регрессионные модели, являющиеся синтезом функционального и вероятностного описания. Для этих моделей могут использоваться переменные любой шкалы измерения: количественные, ранговые, бинарные и даже оцифрованные метки [2, 3].

Корреспондентский анализ, являющийся развитием корреляционного анализа, позволяет устанавливать связи между неколичественными переменными [4].

Метод Multifactor Dimensionality Reduction (MDR) позволяет описывать связи и оценивать тип взаимодействия для бинарных переменных и строить диаграммы по аналогии с кластерным анализом (для которого, как известно, переменные должны быть только количественными) [5, 6].

Таким образом, вполне оправданным и соответствующим современным тенденциям в математической статистике является разработка таких подходов для анализа данных, которые не столько отличались бы строгостью математического описания, сколько являлись бы синтезом классических приемов и были бы направлены на их практическое использование.

ПОСТАНОВКА ЗАДАЧИ

В медико-экологических исследованиях традиционной является задача оценки действия совместного действия совокупности неблагоприятных факторов окружающей среды на те или иные показатели здоровья. Особенно часто такая проблема возникает в эпидемиологических исследованиях при натурной регистрации выраженности тех или иных экологических и социальных характеристик.

Результаты таких исследований в значительной мере основываются на соответствующей математической (статистической) обработке. В свою очередь, математическая обработка определяется типом полученных данных. В прикладной математике данные разграничиваются на четыре основных типа: количественные, ранговые, бинарные и качественные. От принадлежности к этим типам зависит выбор адекватного математического аппарата [7].

Если и воздействующие факторы $\{x_i\}$ и показатели здоровья $\{y_j\}$ относятся к количественным переменным, наиболее популярными математическими методами являются корреляционный и (или) регрессионный анализ.

В рамках корреляционного анализа, как правило, рассчитываются коэффициенты парной корреляции $r(x_i y_j)$. Их достоверные значения

свидетельствуют в пользу причинно-следственных связей для некоторой (устраивающей исследователя) совокупности пар: отдельный i -ый фактор среды — отдельный j -ый показатель здоровья. Значительно реже применяется расчет множественных коэффициентов корреляции, что более правильно для выявления эффектов совместного действия факторов. Но, в любом случае, выделение с помощью корреляционного анализа в результирующей реакции биосистемы составляющих, связанных именно с совместным действием факторов, крайне затруднено.

Более адекватным приемом является построение регрессионных функций, описывающих зависимость исследуемых характеристик здоровья от совокупности всех учитываемых (регистрируемых) факторов окружающей среды. Эффекты совместного действия в уравнениях множественной регрессии, как правило, связываются с коэффициентами при некоторых искусственных свертках исходных переменных (чаще всего для свертки используется произведение переменных). При этом появляется возможность расчета прогнозных значений здоровья и относительного вклада каждого из исследуемых факторов.

В том случае, если и воздействующие факторы, и показатели здоровья получены в ранговой шкале, адекватными методами являются ранговый корреляционный анализ или составление таблиц сопряженности.

Если среда оценивается в бинарном виде, например, как «опытный» («загрязненный») и «контрольный» («условно чистый») районы, а показатели здоровья относятся к количественной шкале измерений, то традиционно применяется критерий Стьюдента. С его помощью устанавливаются различия между «чистой» и «грязной» территорией для показателей здоровья, на основании чего делается вывод о влиянии на здоровье одного фактора, обуславливающего «загрязнение», или их совокупности.

Часто эффект изменения (ухудшения) здоровья регистрируется в бинарном виде («изменение есть / изменения нет»), а воздействующие факторы — в количественной шкале. В таком случае целесообразно применение дискриминантного анализа, с помощью которого строятся классификационные функции и рассчитывается вероятность ухудшения здоровья [8, 9].

Однако в практике медико-экологических исследований не всегда есть возможность получения количественной информации о загрязнении окружающей среды. Иногда информация носит качественный характер констатации самого факта загрязнения объекта исследования типа «да» — «нет». Кроме того, в ходе математического анализа данных часто возникает потребность преобразования количественных данных (если они есть) в бинарную качественную шкалу. Бинарные переменные легко преобразуются в вероятности (проценты). Иногда и сами показатели статистической отчетности также фиксируются в процентах фактов загрязнения («процент проб, превышающих гигиенический норматив»).

Таким образом, задача описания результатов очень часто попадает в область применения теории вероятностей.

Цель — разработка методики определения совместного действия факторов, позволяющей оперировать с качественной (бинарной) информацией и использующей подходы и методы теории вероятностей.

СОСТОЯНИЕ ПРОБЛЕМЫ

Рассмотрим ситуацию, когда данные и о среде, и о здоровье получены либо в качественном виде («чистый» и «грязный» район и «хорошее» и «плохое» здоровье), либо преобразованы к такому виду, исходя из целей исследования.

Традиционным и широко используемым приемом обработки такого типа данных является составление четырехпольных таблиц сопряженности с последующим подсчетом соответствующих статистических критериев различия (связности), например, χ^2 Пирсона.

В последнее время популярным методом обработки четырехпольных таблиц в эпидемиологии становится расчет относительного риска (RR), позволяющий оценивать «степень вредности воздействия» в относительных единицах.

Однако и в первом, и во втором варианте корректно рассчитываются характеристики связности только для пар «воздействие – эффект». Для оценки совместного действия совокупности факторов приходится прибегать к вспомогательным расчетам, являющимся не всегда корректными.

МЕТОД ОПРЕДЕЛЕНИЯ СОВМЕСТНОГО ДЕЙСТВИЯ ФАКТОРОВ

Рассмотрим возможность использования имеющихся в классической теории вероятностей методов и формул, с помощью которых можно рассчитать совместные эффекты действия [10].

Для согласования строгих подходов теории вероятностей и предлагаемых прикладных приемов сделаем следующие допущения. Во-первых, мы не всегда будем соблюдать аксиоматику этой математической дисциплины, делая упор на утилитарном использовании основных математических формул. Во-вторых, определимся с понятием «событие» — ключевым понятием для расчета соответствующих вероятностей.

Будем считать, что событием x_{jy} является факт одновременного наличия и действующего i -го фактора, и эффекта от его действия (т.е. заболевания $-y$) при j -ом измерении (регистрации медико-экологической ситуации в j -ом пункте). Если такое событие происходит, $x_{jy} = 1$, в противном случае $x_{jy} = 0$.

Тогда вероятность действия i -го фактора на показатель здоровья y будет равна:

$$P_{iy} = \frac{1}{m} \sum_{j=1}^m x_{jy} ,$$

где m — число точек регистрации.

Рассмотрим несколько вариантов расчета вероятности совместного действия разных факторов.

1. Наиболее простой для нашей цели в теории вероятностей является формула расчета *вероятности совместного действия независимых событий*. В соответствии с ней вероятность P_{Xy} совместного действия независимых событий $\{x_i\} = X$ на произвольную характеристику здоровья (y) будет выглядеть следующим образом:

$$P_{Xy} = 1 - (1 - p_{1y})(1 - p_{2y}) \dots (1 - p_{ny}) = 1 - \prod_{i=1}^n (1 - p_{iy}), \quad (1)$$

где n — количество действующих факторов.

Главные недостатки применения этой формулы в медико-экологических исследованиях следующие.

Само предположение о независимости действующих факторов вряд ли оправдано, поскольку, как правило, объекты исследования, значительно «загрязненные» одним из факторов, оказываются «загрязненными» и другими факторами. Это приводит к их достоверной множественной корреляции, что неоднократно выявлялось на анализируемых массивах данных.

Формула очень чувствительна к количеству факторов. Если имеются незначительные частные вероятности возникновения эффектов, только слегка (и недостоверно) превышающие неопределенность (например, 51%), то уже для $n = 3$ итоговая вероятность приближается к 90%, а при $n = 5$ достигает 97%. При больших частных вероятностях, даже при учете двух-трех факторов, вероятность эффекта становится практически стопроцентной.

2. Другим вариантом расчета совместного действия является привлечение понятия «условной вероятности». В таком случае событием — причиной, можно считать одновременное действие всех n факторов, а возникновение результата (например, заболевания) считать зависимым от этой причины.

Вероятность одновременного действия n факторов подсчитывается как отношение суммы событий наличия всех действующих факторов x_n к числу измерений m :

$$P(X) = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n x_{ji} = \frac{x_n}{m}.$$

При этом суммарное $x_n = 1$, если все $x_i = 1$ (логическое перемножение).

Вероятность совместного проявления и совокупности факторов, и наступления эффекта (результата) запишется аналогичным образом:

$$P(Xy) = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n x_{jly} = \frac{x_{ny}}{m}.$$

В этой формуле также $x_{ny} = 1$, если все $x_{iy} = 1$.

Тогда условная вероятность $P(y|X)$ в соответствии с известной формулой теории вероятностей будет равна:

$$P(y|X) = \frac{P(Xy)}{P(X)}. \quad (2)$$

Недостаток использования такого подхода тот же, что и для предложенного ранее варианта, а именно — предположение независимости действия факторов, т.е. отсутствие в формулах эффектов взаимодействия факторов. Следовательно, можно предположить, что рассчитанная таким образом условная вероятность будет меньше, чем гипотетическая «правильная» вероятность, которая должна быть дополнена добавкой, вызванной эффектами взаимодействия.

3. Если изначально предположить взаимозависимость (совместность) эффектов действия различных факторов, то можно воспользоваться формулами произведения событий.

При действии двух факторов формула будет иметь вид:

$$R_{\Sigma y} = p_{1y} + p_{2y} - p_{1y+2y}, \quad (3)$$

где p_{1y+2y} — означает вероятность либо события p_{1y} , либо события p_{2y} (вероятность суммы двух событий или логическое сложение).

Для трех факторов имеем:

$$R_{\Sigma y} = p_{1y} + p_{2y} + p_{3y} - p_{1y+2y} - p_{1y+3y} - p_{2y+3y} + p_{1y+2y+3y}. \quad (4)$$

Недостатком использования такого подхода является значительное усложнение формулы для большего количества факторов, поскольку в нее входят вероятности сумм двух, трех и т.д. событий.

Таким образом, приходим к выводу, что использовать непосредственно формулы (1), (2) или (3) либо некорректно, либо слишком сложно.

В связи с этим нами была предпринята попытка модернизации этих формул. При этом будем основываться на формуле (1), но дополним ее, используя следующие соображения (постулаты).

1. Увеличение вероятности $R_{\Sigma y}$ при увеличении количества сомножителей (факторов) n должно происходить более инерционно, чем в исходной формуле.

2. В формуле должна учитываться возможная связь факторов между собой. При наличии такой связи $R_{\Sigma y}$, рассчитанная по классической формуле, должна уменьшаться, поскольку совместное действие «выбирает на себя» часть от совокупности эффектов независимого действия. При отсутствии совместных взаимодействий факторов формула должна приобретать традиционный вид.

3. Поскольку мы предположили, что исходные данные имеют бинарный вид, мерой такой связности могут быть условные межфакторные вероятности p_{ik} наличия i -го фактора в присутствии k -го фактора (где $i \neq k$).

4. Общая вероятность не должна быть меньше максимальной частной вероятности. То есть, если «наиболее вредный» k -ый фактор вызывает ухудшение здоровья с вероятностью $p_k = p_{max}$ то $P_{\Sigma y} > p_{max}$.

При конструировании итоговой формулы будем использовать подходы формальной логики и теории вероятностей [11, 12].

А именно, если p_{iy} — вероятность действия i -го фактора на показатель здоровья y , то его дополнение $(1 - p_{iy})$ — вероятность «недействия». Произведение вероятностей событий означает вероятность их совместного наступления.

Например, в «классической» формуле

$$P_{\Sigma y} = 1 - \prod_{i=1, \dots, n} (1 - p_{iy})$$

каждый из сомножителей $(1 - p_{iy})$ означает «недействие» i -го фактора, их произведение — *совместное «недействие»*, а дополнение к произведению — *вероятность совместного действия*.

Вероятность совместного наличия любого из факторов i и максимально действующего k запишется как p_{ik} , где $i=1, \dots, k, \dots, n$. Его дополнение $(1 - p_{ik})$ имеет смысл вероятности «отсутствия совместного наличия этих факторов». Произведение этой скобки на p_{iy} , т.е. $p_{iy}(1 - p_{ik})$, означает одновременное *действие* i -го фактора на показатели здоровья (заболеваемость) и *независимость* любого из i -ых факторов от k -го фактора. Соответственно, дополнение к этому произведению $(1 - p_{iy}(1 - p_{ik}))$ имеет смысл «недействия» i -го фактора с учетом его связности с k -ым фактором, т.е. «*совместного зависимого недействия*» факторов i и k .

Далее, перемножая эти скобки для всех факторов, получаем *вероятность «совместного зависимого недействия» для некоторой совокупности воздействующих факторов*. Выделим максимально действующий фактор, и скобку с его «недействием» вынесем за знак произведения.

Тогда *совместное недействие всех факторов*, включая максимально «вредный», запишется в следующем виде:

$$(1 - p_{ky}) \prod_{i=1, i \neq k}^{n-1} (1 - p_{iy}(1 - p_{ik})).$$

Соответственно, дополнение к этому выражению будет иметь смысл вероятности *совместного действия всех факторов с учетом их связи с наиболее действующим, то есть в предположении их зависимого действия*.

Итак, получаем итоговую формулу:

$$P_{\Sigma y} = 1 - (1 - p_{ky}) \prod_{i=1, i \neq k}^{n-1} (1 - p_{iy}(1 - p_{ik})). \quad (5)$$

Очевидно, что при наличии полной независимости всех факторов от ведущего, т.е. при $p_{ik} = 0$, формула приобретает «классический» вид (1). Если хоть одна из вероятностей $p_{ik} \neq 0$, итоговая вероятность по (5) будет меньше вероятности, рассчитанной по (1). При условии полной зависимости всех остальных факторов от «ведущего» ($p_{ik} = 1$), вероятность их «совместного действия» определяется только одним «ведущим» фактором ($P_{\Sigma y} = p_{ky}$).

ПРИМЕР РАСЧЕТА СОВМЕСТНОГО ДЕЙСТВИЯ НЕСКОЛЬКИХ ФАКТОРОВ

Пусть в девяти пунктах наблюдения регистрировалось наличие трех ($n = 3$) вредных факторов среды (x_1, x_2, x_3) и одного показателя здоровья (y). Выраженность переменных кодировалась в бинарном виде: «1» — значимая вредность факторов (например, превышение нормативного значения) или достоверное ухудшение здоровья (например, по сравнению со среднестатистическими значениями); 0 — отсутствие вредности факторов и ухудшения здоровья (табл. 1).

Таблица 1.

Бинарное представление медико-экологической информации

N пункта наблюдения	Факторы среды			Показатель здоровья (y)
	x_1	x_2	x_3	
1	1	1	1	1
2	0	1	1	0
3	1	1	1	1
4	1	1	1	1
5	1	0	1	1
6	1	1	0	1
7	1	0	1	1
8	1	0	1	0
9	1	1	1	0

Событием для расчета вероятностей будем считать совпадения «1» для каждой переменной в сравниваемых столбцах. Вероятность рассчитывается как отношение сумм этих совпадений к количеству замеров (в данном случае, пунктов наблюдения).

Тогда, в соответствии с таблицей число совпадений «1» для факторов x_1 и x_2 равно 5. Соответственно, $p_{12} = 5/9 = 0,56$.

Вероятности связи наличия факторов и ухудшения здоровья будут равны: $p_{1y} = 0,67$, $p_{2y} = 0,44$, $p_{3y} = 0,56$. Максимальной будет вероятность действия первого фактора $p_{1y} = p_{max} = 0,67$, поэтому назовем его «ведущим» фактором (p_{ky}). Вероятности одновременного наличия в пунктах наблюдения этого фактора и двух других равны соответственно: $p_{12} = 0,56$, $p_{13} = 0,78$.

Вероятность совместного действия всех трех факторов на здоровье по предлагаемой формуле (5) будет равна:

$$\begin{aligned}
 P_{\Sigma} &= 1 - (1 - 0,67)[1 - (1 - 0,56) * 0,44] * [1 - (1 - 0,78) * 0,56] = \\
 &= 1 - 0,33 * (1 - 0,44 * 0,44) * (1 - 0,22 * 0,56) = 1 - 0,33 * 0,81 * 0,88 = \\
 &= 1 - 0,14 = 0,76.
 \end{aligned}$$

Для сравнения: по формуле для независимых событий (1) вероятность равна $P_{\Sigma y} = 0,92$, по формуле для условной вероятности (2) — $P_{\Sigma y} = 0,75$, по формуле (3) — $P_{\Sigma y} = 0,33$.

Видно, что вероятность, рассчитанная по формуле (5), находится в интервале между вероятностью изолированного действия наиболее существенного из факторов (нижняя граница) и совместной вероятностью независимого действия (1) всех факторов (верхняя граница), что отвечает сформулированным изначально условиям. Формулы (2) и (3) дают неправильные и явно заниженные результаты, поскольку вероятности, рассчитанные по ним, меньше, чем вероятность изолированного действия наиболее вредного фактора.

Выводы

Показана неадекватность применения методов оценки вероятности совместного действия факторов, используемых в теории вероятностей и математической статистике.

Предложена оригинальная формула расчета вероятности совместного действия факторов, учитывающая их возможную взаимосвязь. При наличии полной независимости всех факторов от ведущего, формула приобретает «классический» вид для расчета совместной вероятности независимых событий. Если имеется хоть одна взаимосвязь, вероятность, рассчитанная по предлагаемой формуле, будет меньше «классической». При условии полной зависимости всех остальных факторов от ведущего, вероятность их совместного действия определяется только максимальной вероятностью эффекта.

1. Дюк В. А., Самойленко А. П. Data Mining. Учебный курс — СПб, 2001. — 368 с.
2. David W. Hosmer, Stanley Lemeshow Applied Logistic Regression, 2nd ed. — New York, Chichester, Wiley, 2002. — 383p.
3. Наследов А. Н. SPSS 19: Профессиональный статистический анализ данных — СПб, 2011. — 400 с.
4. Greenacre M. Correspondence Analysis in Practice, 2nd ed. — London: Chapman & Hall/CRC, — 2007. — 280 p.
5. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. / M.D. Ritchie, L.W. Hahn, N. Roodi & others // Am J Hum Genet. — 2001 Jul;69(1):138-47.
6. Hahn L.W., Ritchie M.D., Moore J.H. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions//Bioinformatics. — 2003 Feb 12;19(3):376-82
7. Орлов А. И. Прикладная статистика — М.: Изд-во «Экзамен» — 2004. — 656 с.
8. Антомонов М.Ю. Математическая обработка и анализ медико-биологических данных — К.: Изд-во „Малый друк“, — 2006. — 558с.
9. Гайдышев И. Анализ и обработка данных — СПб. — 2001. — 750 с.
10. Вентцель Е. С. Теория вероятностей. 10-е изд., стер.. — М.: «Академия», — 2005. — 576 с.
11. Гнеденко Б. В., Хинчин А. Я. Элементарное введение в теорию вероятностей — 1970. — 168с.

THE METHOD OF DETERMINATION OF ENVIRONMENTAL FACTORS JOINT IMPACT IN EPIDEMIOLOGICAL STUDIES FOR BINARY DATA

M.Yu. Antomonov

*State Institution "O.N. Marzeev Institute for Hygiene and Medical Ecology
of NAMS of Ukraine" (Kiev)*

Introduction. Modern approaches for data analysis combine classical methods and focused on their practical application. Sometimes the information is presented in the form of qualitative characteristics that are characterize the contamination of the research object. Such binary variables are easily transformed into a probability (in percent), so the task description of results performed using probability theory.

The purpose of the article is to develop such a common method for calculation joint action of the factors that would allow to operate with qualitative (binary) information and would use techniques and formulas of probability theory

Results. A careful analysis was carried out for the existing approaches in the medical and environmental studies for calculating the effect of the joint action of the factors. It was evaluated disadvantages of these approaches that implemented in the theory of probability and mathematical statistics.

The article proposes an original method of calculating the combined effect of the factors that allows you to work with the information expressed in binary form. The final expression was designed by using approach of formal logic and probability theory.

Conclusions. It is shown that the known methods of probability theory cannot be adequately used to evaluate the combined effect of the factors. The original method of calculating the probability of the joint action of factors that take into account their possible connection is described.

Keywords: qualitative data, binary variables, joint effect of the factors, the probability of independent and interdependent events.

1. Duke V. A. Samoilenko A. P. Data Mining. Training — SPb, 2001. — 368p.
2. David W. Hosmer, Stanley Lemeshow Applied Logistic Regression, 2nd ed. — New York, Chichester, Wiley. 2002. — 383p.
3. Nasledov A. N. SPSS 19: Professional statistical analysis. — SPb, 2011. — 400 p.
4. Greenacre M. *Correspondence Analysis in Practice*, 2nd ed. — London: Chapman & Hall / CRC — 2007. — 280 p.
5. Ritchie M. D., Hahn L. W., Roodi N., Bailey L. R., Dupont W. D., Parl F. F., Moore J. H. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am. J. Hum. Genet.* 2001 Jul; 69 (1): 138-47.
6. Hahn L. W., Ritchie M. D., Moore J. H. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions//*Bioinformatics*. 2003 Feb 12; 19 (3): 376-82
7. Orlov A. I. Applied Statistics — М.: Publisher "Exam", 2004. — 656 p.
8. Antomonov M. Y. Mathematical processing and analysis of medical-biological data. — Kiev: Publishing house "Malii Druk", 2006. — 558p.

9. Gaydyshev I. Analysis and data processing — St. Petersburg, 2001. — 750p.
10. Wentzel E. S. Probability 10th ed., Sr — M.: "Academy", 2005. — 576 p.
11. Gnedenko B. V., Khinchin A. Y. An elementary introduction to the theory of probability, 1970. — 168p.
12. Novikov P. S. Elements of mathematical logic. 2nd ed. — M.: Nauka, 1973 — 400 p.

Получено 02.06.2015

КОМПЛЕКСНАЯ ОЦЕНКА РЕГИОНАЛЬНОГО ДЕМОГРАФИЧЕСКОГО РАЗВИТИЯ

О.А. Кривова, Л.М. Козак

Международный научно-учебный центр информационных технологий и систем НАН Украины и МОН Украины

Проанализированы подходы к формированию композитных индикаторов социально-демографического развития. Рассмотрены особенности информационно-методического обеспечения определения регионального человеческого развития, в частности для расчета индекса воспроизводства населения в регионах Украины. На основе методов кластерного и регрессионного анализа разработана методика формирования комплексной оценки регионального демографического развития. Проанализированы отличия предложенного подхода к формированию композитного индикатора и действующей методики измерения регионального человеческого развития.

Ключевые слова: кластеризация, композитный индикатор, индекс регионального демографического развития, упорядоченная классификация.

Проаналізовано підходи до формування композитних індикаторів соціально-демографічного розвитку. Розглянуто особливості інформаційно-методичного забезпечення визначення індексу регіонального людського розвитку, зокрема для розрахунку індексу відтворення населення в регіонах України. На основі методів кластерного та регресійного аналізу розроблено методику формування комплексної оцінки регіонального демографічного розвитку. Проаналізовано відмінності запропонованого підходу до формування композитного індикатора і діючої методики вимірювання регіонального людського розвитку.

Ключові слова: кластеризація, композитний індикатор, індекс регіонального демографічного розвитку, упорядкована класифікація.

ВВЕДЕНИЕ

Для исследования и анализа аспектов человеческого развития (демографических, социально-экономических, экологических и других) используют набор статистических показателей или индикаторов. При необходимости принятия управляющих решений обращаются к комплексным оценкам или сводным индикаторам. В англоязычной научно-технической литературе для таких сводных оценок используется термин «композитный индикатор», или «индекс явления, процесса». Социально-экономические композитные индикаторы, иначе индексы качества жизни, бедности, экономической безопасности, социального прогресса и другие, применяются для сопоставления этих величин для разных стран. Наиболее известным является "индекс развития человеческого потенциала" (ИРЧП), по которому ООН с 1990 года проводятся международные сравнения возможностей и уровня развития человеческого потенциала. В Украине разработана национальная методика измерения уровня регионального человеческого развития, адаптированная к национальной статистической базе (Методика РЧР) [1]. С 2001 года Государственная служба статистики проводит по этой

методике расчеты, результаты публикуются в ежегодных сборниках [2]. На основе этих результатов можно оценить тенденции и диспропорции в региональных условиях жизни населения: характеристик воспроизводства населения, социальной среды и работы, уровня благосостояния, условий проживания, возможностей образования [3].

В работах Портера М.Е [4], Изарда У.[5]. Гранберга А.Г. [6], Либановой С.М. [3], Герасимчук З.В. [7] разработаны теоретико-методологические основы анализа социально-экономического развития регионов. Прикладным и методическим аспектам оценки регионального развития, в частности совершенствованию методики формирования индексов, посвящены работы ряда отечественных и зарубежных ученых [9–15]. На наш взгляд, целесообразно провести сравнение методологических подходов к оценке региональной дифференциации, основанных на разных типах математических моделей.

Цель — формирование комплексной оценки демографического развития регионов Украины на основе упорядоченной типологической классификации.

ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ ПОСТРОЕНИЯ КОМПОЗИТНЫХ ИНДИКАТОРОВ

Теоретические основы и принципы формирования композитных индикаторов изложены в рекомендациях, разработанных группой прикладной статистики Объединенного исследовательского центра Еврокомиссии (JRC-COIN) [16] и ОБСЕ [17]. Общая схема формирования композитных индикаторов (или композитных индексов) состоит из следующих основных этапов.

Выбор переменных. Отбор наиболее информативных показателей; оценка их качества (полнота, достоверность, отсутствие выбросов), а также определение зависимости между выбранными переменными (проверка мультиколлинеарности). Объединение набора показателей (первичных индикаторов) в отдельные блоки.

Нормализация индикаторов. Для соизмеримости показателей, представленных в разных единицах измерения, их необходимо привести к общему масштабу. Основные методы нормализации переменных:

$$z = \frac{x}{\bar{x}} \quad (1)$$

$$z = \frac{x}{\bar{x} - \sigma} \quad (2)$$

$$z = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

$$z = \frac{x}{x_{\max}} \quad (4)$$

где \bar{x} — среднее значение показателя, σ — среднеквадратическое отклонение исходной переменной, x_{max} и x_{min} — максимальное и минимальное значение x .

Преобразование (2) соответствует методу стандартизации переменных, основанному на нормальном распределении, при котором переменные приводятся в масштаб единичного стандартного отклонения. Минимаксный метод нормализации (3) соответствует масштабированию в диапазон [0, 1]. Используются и другие варианты преобразования переменных. При выборе способа нормализации индикаторов необходимо учитывать их свойства (тип данных, особенности функций распределения). Известно, что разные методы нормализации могут приводить к разным конечным результатам [18].

Методы и модели агрегации первичных данных. Перечислим основные модели формирования композитного индекса, основанные на методах многомерного статистического анализа:

- модель множественной линейной регрессии;
- метод главных компонент (факторный анализ) [1];
- метод оценки внутренней согласованности (коэффициент α Кронбаха);
- кластерный анализ [19];
- многомерное шкалирование;
- нейросетевые модели [20];

Типичный вид композитного индекса при аддитивной свертке индикаторов:

$$I_j = \sum_{i=1}^m w_i x_{ji} \quad (5)$$

$$\sum_{i=1}^m w_i = 1, \quad 0 \leq w_i \leq 1 \quad (6)$$

где x_{ji} — нормализованные индикаторы; m — количество индикаторов, w_{ji} — их весовые коэффициенты. Весовые коэффициенты рассчитываются на основе выбранной модели.

Линейная свертка является традиционным подходом к формированию композитных индикаторов, реже применяют мультипликативный, комбинаторный метод агрегации индикаторов, функциональное шкалирование.

Заключительный этап — интерпретация результатов, проверка обоснованности индекса, определение диапазона для принятия решений, построение рейтинга исследуемых объектов [16, 21].

Каждый из этапов формирования комплексной оценки в виде индекса имеет свои методические проблемы и трудности: выбор системы первичных индикаторов, объединение их в отдельные блоки, согласованность выбора способа нормализации и математической модели, проверка адекватности полученных результатов.

В последние десятилетия на базе теории принятия решений разработаны специальные методы и алгоритмы, где в качестве входных данных рассматриваются альтернативы, а затем оцениваются в соответствии с определенными критериями [22, 23]. Одним из таких подходов к принятию решений, применяемый для формирования интегральных индикаторов, является классический метод анализа иерархий Саати (МАИ) [25], а также его последующие модификации [26]. В основе метода лежат: процедура попарного сравнения показателей на основе шкалы от одного до девяти, формирование матрицы попарных сравнений и вычисление относительных весовых коэффициентов индикаторов как собственного вектора матрицы.

В последние годы активно разрабатываются гибридные алгоритмы «построения упорядоченной классификации» (ordinal classification, multicriteria ordered clustering), которые предварительно разбивают объекты на группы (категории, кластеры), чем упрощается проблема принятия решения о преимуществах по нескольким критериям [25–27]. Покажем использование такого подхода на примере комплексного оценивания и формирования композитного индекса демографического регионального развития Украины.

МЕТОДИЧЕСКИЕ ПОДХОДЫ К РАСЧЕТУ РЕГИОНАЛЬНЫХ ИНДЕКСОВ РАЗВИТИЯ

Методика РЧР. Расчеты индекса регионального человеческого развития в соответствии с первой версией национальной методики проводились на протяжении 10 лет. Выявленные региональные проблемы, а также недостатки методологического подхода были детально проанализированы учеными Института демографии и социальных исследований им. М.В. Птухи НАН Украины. Вносился ряд замечаний, предложений по ее усовершенствованию [12, 14]. В результате произошло кардинальное обновление методики, изменилось все информационно-методическое обеспечение расчетов индексов [28].

В обновленной Методике РЧР изменена система первичных показателей: сократилось число показателей и блоков, соответствующих аспектам человеческого развития. Составляющие регионального развития объединены в шесть блоков: «Воспроизводство населения», «Социальное положение», «Комфортная жизнь», «Благосостояние», «Достойный труд», «Образование». Изменен алгоритм стандартизации показателей, вместо нормировки на размах, была введена калибровка на нормативные (целевые, или стандартные) значения показателей. Выбрана новая модель формирования композитного индекса. В первой версии национальной методики расчеты проводились на основе модели главных компонент, теперь же весовые коэффициенты индексов каждого из блоков определяются методом попарных сравнений с применением экспертных оценок [23, 28, с. 36]. Интегральный индекс представляет собой сумму отдельных индексов. На основе индексов блоков рассчитываются рейтинги областей по отдельным блокам. Заключительный этап — зонирование (визуализация результатов для

принятия решений) представлен в виде карт группировки регионов Украины по составляющим регионального развития и по интегральному показателю. Зоны соответствуют группам, сформированным по градациям индексов (высокий уровень, выше среднего, средний, ниже среднего, низкий).

В научной литературе был высказан ряд аргументов против традиционных моделей композитных индексов [18]. Основное замечание, относящееся к линейной агрегации переменных, состоит в том, что весовые коэффициенты индикаторов ошибочно трактуют как меру важности (или преимущества) одного индикатора по сравнению с другим. Линейная модель (5) и (6) предполагает взаимную компенсацию составляющих, а весовые коэффициенты являются мерой взаимозаменяемости индикаторов (составляющих процесса).

Алгоритм расчета коэффициентов модели композитного индекса устанавливает предпочтения объектов. По нашему мнению, расчет весовых коэффициентов с использованием МАИ, основанный на согласовании суждений экспертов, объективизирует оценки и рейтинги, однако при этом теряется исходная информация о реальных соотношениях составляющих региональных процессов. Кластерный подход к оценке дифференциации регионального развития дает информацию о взаимосвязанных изменениях различных составляющих. Этот подход, основанный на выделении классов регионов с однородными показателями, применялся при разработке типологии субъектов Российской Федерации [19], а также для оценки влияния комплексных факторов регионального развития на рождаемость населения Украины [29].

МЕТОДИКА КОМПЛЕКСНОЙ ОЦЕНКИ РЕГИОНАЛЬНОГО СОЦИАЛЬНО-ДЕМОГРАФИЧЕСКОГО РАЗВИТИЯ

Задачу оценивания регионального социально-демографического развития и формирования композитного индикатора можно рассматривать как задачу построения упорядоченной классификации [19, 25]. Основное предположение: в исходных данных существует структура (кластеры), обусловленная взаимосвязями показателей (индикаторов процесса). Особенности типологических кластеров определяются соотношением составляющих (индикаторов). После определения типологических кластеров их необходимо упорядочить, построив функцию предпочтения на основе выявленных альтернатив соотношений индикаторов в кластерах. При этом не отдается предпочтение каким-либо индикаторам (назначается вес), как это происходит при экспертном оценивании, а выявляются реально существующие альтернативы соотношений индикаторов в группах, и таким способом определяются «весовые коэффициенты». На примере расчета индекса воспроизводства населения по данным, представленным в действующей Национальной методике РЧР [28], рассмотрим основные этапы предложенной методики.

Этап I. Выбор и нормализация индикаторов. Нормализация показателей проводится по минимаксному методу:

$$x_{ji} = (X_{ji} - X_{\min, i}) / (X_{\max, i} - X_{\min, i}) \quad (7)$$

$$x_{ji} = (X_{\max, i} - X_{ji}) / (X_{\max, i} - X_{\min, i}) \quad (8)$$

где $j = 1, \dots, N$, N — число объектов; $i = 1, \dots, m$, m — количество показателей X_{ji} . Для показателей-стимуляторов применяется выражение (7), для дестимуляторов — (8). Показатель-дестимулятор — это показатель такой составляющей процесса, увеличение которой приводит к уменьшению результирующего процесса. После нормировки по минимаксному методу индикаторы изменяются в диапазоне от 0 до 1, а максимальное значение показателя-дестимулятора становится равным 0, а минимальное 1, для показателей-стимуляторов — наоборот. Таким образом, все индикаторы одинаково влияют на результирующий процесс.

Этап II. Выделение категорий, или определение оптимальной кластерной структуры (устойчивой группировки регионов).

Имеем N объектов — регионов $\{X_j\}_{j=1}^N$. В результате процедуры кластерного анализа, описанной в работе [29], получим оптимальное число (NC) кластеров $\{Y_k\}_{k=1}^{NC}$, каждый из которых имеет характерное соотношение средних значений показателей в кластере (координат кластерных центров). Эти соотношения выявляют имеющиеся диспропорции между индикаторами (составляющими процесса в регионах).

Этап III. Сортировка категорий, или же построение функции предпочтения:

$$f_{NC}(Y_k^{(NC)}) = k, \quad k = 1, \dots, NC \quad (9)$$

Функция предпочтения является монотонной функцией от номера объекта. Чем больше ее значение, тем лучше альтернатива (предпочтительнее соотношение индикаторов в комплексной оценке). Упорядочим кластеры $\{Y_k\}_{k=1}^{NC}$ по альтернативам соотношений координат кластерных центров так, чтобы функция предпочтения принимала значения упорядоченных номеров кластеров. Комбинаций (вариантов упорядочения) в общем случае может быть несколько. Предположим, что существует линейная функция предпочтения.

Этап IV. Аппроксимация функции предпочтения линейной регрессионной моделью:

$$\tilde{f}(X^{(k)}) = \sum_{i=1}^m \alpha_i x_i^{(k)} \quad (10)$$

Лучший по статистическим оценкам вариант регрессии (результат моделирования функции предпочтения) выбираем в качестве индекса I_j —

функции упорядоченных по номерам объектов. Полученную комплексную оценку — индекс можно нормировать в диапазон $[0, 1]$.

Этап V. Ранжирование. Расчет для каждого объекта значения композитного индекса I_j (индекса аспекта регионального развития) и ранжирование этой функции I_j — определение рейтинга регионов.

РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ МЕТОДИКИ КОМПЛЕКСНОЙ ОЦЕНКИ

Предложенная методика использовалась для расчета индекса регионального демографического развития. Для сравнения результатов разных подходов использовались исходные статистические данные за 2010 год из блока «Воспроизводство населения» Методики РЧР [28, с.30]. Расчеты проводились с применением статистического пакета STATISTICA 10.

В блок «Воспроизводство населения» включены демографические показатели, представляющие собой «критерии-индикаторы результативности развития — как собственно социально-экономического, так и человеческого» [28, с. 4]:

1) суммарный коэффициент рождаемости (TFR), характеризующий среднее число детей, которое рождается у одной женщины за детородный период, считается наиболее корректным индикатором уровня детородной активности и замещения поколений;

2) детская смертность (смертность детей в возрасте до пяти лет), ‰ (ChM) является общепризнанным показателем воздействий медико-демографического и социального благополучия в регионе и стране;

3) средняя ожидаемая продолжительность жизни при рождении, число лет (L); индикатор состояния здоровья, условий жизни и труда населения показывает число лет, которые проживет новорожденный при условии, если в течение жизни интенсивность смертности в каждой возрастной группе будет оставаться такой же, как в год рождения;

4) вероятность мужчин дожить от 20 до 65 лет — (Pm);

5) вероятность женщин дожить от 20 до 65 лет — (Pw).

Выбор последних двух показателей обусловлен тем, что в Украине наблюдается значительная диспропорция преждевременной смертности по полу.

Среди пяти индикаторов процесса воспроизводства населения дестимулятором является показатель детской смертности.

Нормирование индикаторов выполнялось по формулам (7), (8). Отметим, что показатель ожидаемой продолжительности жизни (L) коррелирует с показателями вероятностями дожития (Pm , Pw), а также детской смертностью (ChM). Соответствующие коэффициенты корреляции Спирмена равны: $R_{S1} = 0,898$, $R_{S2} = 0,927$, $R_{S3} = -0,445$ при $p < 0,05$.

Проводился кластерный анализ 25-ти областей Украины по пяти нормированным переменным. Применялись два метода: метод Уорда и k -средних (расстояние Евклида).

На рис. 1 представлена дендрограмма, полученная методом Уорда, где обозначены метки 4-х основных кластеров. Сравнение состава двух

разбиений, полученных разными методами, показало, что незначительное различие (индекс Ренда $R = 0,92$) касается принадлежности Киевской и Херсонской области к смежным кластерам. Приходим к выводу о подобии кластерной структуры, выделенной разными методами, что свидетельствует об обоснованности полученной классификации.

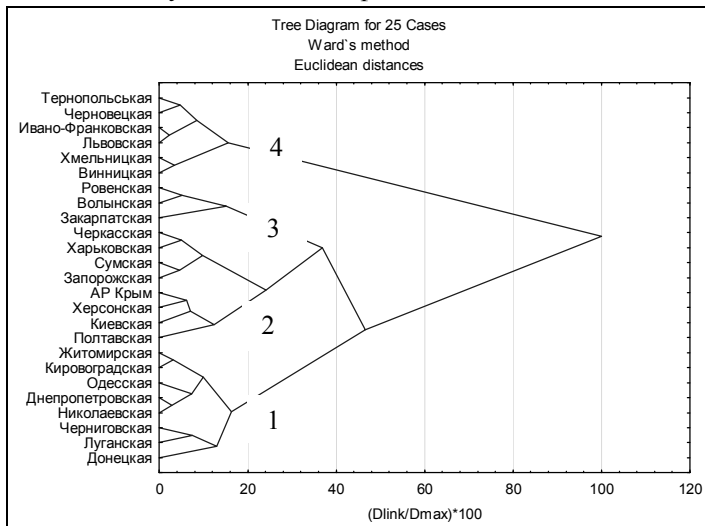


Рис. 1. Дендрограмма объединения регионов в кластеры по индикаторам воспроизводства населения (по методу Уорда)

В таблице 1 приведены средние значения показателей в 4-х кластерах (абсолютные, до нормировки), соответствующие координатам центров кластеров, номера кластеров как на рис. 1. Средние значения индикаторов в кластерах существенно отличаются. Охарактеризуем особенности каждого из кластеров, зависящие от соотношения демографических индикаторов.

В кластер 1 вошли восемь областей с наибольшей диспропорциональностью между индикаторами, а именно низкие уровни рождаемости, продолжительности жизни сочетаются с наибольшей детской смертностью и наибольшее различие по полу в вероятности дожития.

Кластер 2 (восемь областей) отличается от кластера 1 меньшей детской смертностью, большей продолжительностью жизни и вероятностью дожития для обоих полов, но значимо не отличается уровнем рождаемости.

Таблица 1

Средние значения показателей воспроизводства населения в кластерах

Кластеры	Число объектов	TFR	ChM	L	Pm	Pw
1	8	1,4	2,46	69,12	0,549	0,812
2	8	1,38	2,19	70,32	0,584	0,830
3	3	1,86	2,43	70,51	0,594	0,834
4	6	1,50	2,12	72,23	0,631	0,864

Особенность кластера 3 (Волынская, Ровенская, Закарпатская области) — наивысший уровень рождаемости в Украине и пропорциональная ей детская смертность.

Кластер 4 объединяет шесть областей Западного региона с такими особенностями индикаторов для Украины: наибольшая продолжительность жизни, наименьшая детская смертность, наименьшее различие между дожитием мужчин и женщин.

Таким образом, определены альтернативные соотношения демографических индикаторов, характерные для регионов Украины.

Упорядочим выделенные кластеры, построим функцию предпочтения кластеров по альтернативным соотношениям индикаторов. В нашем случае, когда нормировка переменных проводилась по минимаксному методу, функция суммарного вклада показывает преимущества кластеров.

На рис. 2 приведена диаграмма с накоплением, которая отображает вклад каждой переменной по категориям. Вертикальная ось — сумма вкладов нормированных индикаторов (средних значений в кластере). На горизонтальной оси категорий указаны номера кластеров, упорядоченные по возрастанию суммы вкладов индикаторов. Так наглядно можно представить функцию предпочтения кластеров, где кластер 1, кластер 2, кластер 3, кластер 4 упорядочены от «худшего» (кластер 1) к «лучшему» кластеру (кластер 4).

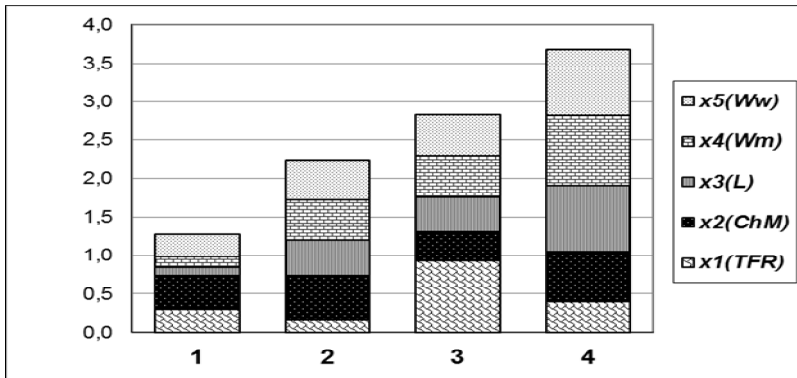


Рис. 2 Диаграмма суммарного вклада средних значений индикаторов в кластерах

Функцию предпочтения, соответствующую представленному на рис. 2 упорядочению кластеров, можно аппроксимировать линейной регрессией.

Для получения устойчивого результата применялась гребневая регрессия (ridge regression), так как некоторые независимые переменные являются высоко коррелированными, что уже отмечалось. При $\lambda = 0,1$ получен лучший статистический результат моделирования с высоким коэффициентом детерминации $R^2 = 0,96$, ($F(5,20) = 96,89$, $p = 0,00000$), другие результаты указаны в таблице 2. Таким образом, рассчитанный нами композитный индикатор имеет вид:

$$I = 1,12 x_1 + 0,61 x_2 + 1,16 x_3 + 1,17 x_4 + 0,92 x_5, \quad (11)$$

где переменные x_i — нормированные индикаторы в диапазоне $[0,1]$.

В таблице 2 представлены коэффициенты регрессионной модели (стандартизированные и обычные), а для сравнения — весовые коэффициенты интегрального индекса воспроизводства населения, полученные по методике РЧР [28].

Таблица 2.

Основные результаты линейной регрессии

Переменные	Весовые коэффициенты (Методика РЧР)	β	Ошибка β	b	Ошибка b	$t(20)$	p
$x_1 (TFR)$	0,1426	0,200	0,06	1,12	0,3	3,28	0,003*
$x_2 (ChM)$	0,1956	0,134	0,07	0,61	0,3	1,74	0,096
$x_3 (L)$	0,2834	0,234	0,10	1,16	0,5	2,17	0,041*
$x_4 (Wm)$	0,216	0,260	0,10	1,17	0,4	2,58	0,017*
$x_5 (Ww)$	0,1624	0,205	0,101	0,92	0,4	1,96	0,063

Как видно из таблицы 2, коэффициенты регрессионной модели β отличаются от весовых коэффициентов, полученных методом парных сравнений экспертных оценок. Согласно экспертным оценкам, наибольший «вес» получил индикатор «ожидаемая продолжительность жизни» [28, с.26]. Статистически значимыми переменными кластерной модели композитного индикатора являются переменные: Wm , L , TFR , а их относительные вклады в модель сопоставимы между собой.

Таким образом, получена комплексная оценка дифференциации демографического развития в виде упорядоченной классификации (кластерной модели), на основе которой определяются весовые коэффициенты композитного индикатора воспроизводства населения Украины.

Индекс воспроизводства населения для каждой области I_j рассчитывается по формуле (11). Рейтинг региона определяется на основе ранга функции I_j . Область, у которой наибольшее значение индекса, получает первое место в рейтинге, область с наименьшим индексом — последнее.

В таблице 3. представлены значения индекса (I) и рейтинга областей Украины, рассчитанные по предложенной модели (11), а также индекс и ранг областей для блока «Воспроизводство населения», полученные согласно методике РЧР.

Как можно заметить, совпадают первые и последние позиции рейтингов областей, рассчитанных по кластерной модели и по методике РЧР [28]. Различия в промежуточных позициях рейтингов некоторых областей, что вполне ожидаемо для разных подходов, можно объяснить другим методом нормирования показателей, выбором в качестве индикаторов коррелированных показателей, а также использованием для получения типологической группировки данных только за один год. Рейтинги, полученные по различным методикам формирования композитного индекса, показали относительную устойчивость.

Преимущество использования предлагаемой модели для определения индекса регионального демографического развития состоит в том, что индекс и рейтинг несут информацию о месте области в пределах типологического кластера, а также дают представление о тенденции перехода в другой кластер.

Таблица 3.

Индексы демографического развития и рейтинги регионов Украины

Области	Модель (11)			Методика РЧР [28]	
	Кластеры	<i>I</i>	рейтинг	индекс	ранг
Тернопольская	4	4,11	1	0,841	1
Черновицкая		4,02	2	0,829	3
Ивано-Франковская		3,76	3	0,806	9
Львовская		3,73	4	0,808	7
Хмельницкая		3,38	5	0,834	2
Винницкая		3,30	7	0,813	6
Ровенская	3	3,21	6	0,821	5
Закарпатская		2,74	8	0,794	11
Волынская		2,73	9	0,806	10
Черкасская	2	2,52	10	0,778	14
Харьковская		2,46	11	0,781	13
АР Крым		2,25	12	0,791	12
Полтавская		2,20	13	0,822	4
Сумская		1,84	14	0,763	18
Запорожская		1,82	15	0,757	21
Киевская		1,81	16	0,807	8
Херсонская		1,62	17	0,774	15
Житомирская		1,37	18	0,757	20
Одесская		1,24	19	0,764	17
Черниговская	1	1,16	20	0,768	16
Кировоградская		1,10	21	0,747	23
Луганская		1,09	22	0,744	24
Днепропетровская		1,02	23	0,752	22
Николаевская		0,94	24	0,758	19
Донецкая		0,38	25	0,711	25

Для получения обоснованных результатов существенно требование, чтобы расчеты кластерной типологической структуры проводились по статистическим данным за длительный период наблюдения.

Особенность предлагаемого подхода состоит в применении объективных, математически обоснованных критериев, по которым определяются типологические кластеры и ранжируются регионы.

Выводы

Разработанная методика комплексной оценки регионального социально-демографического развития позволяет рассчитывать композитный индекс воспроизводства населения по региональным статистическим данным Украины на основе многомерных статистических методов (кластерный, регрессионный анализ). Сравнение результатов, полученных с помощью разных подходов, показало, что преимущество применения кластерной модели состоит в формировании композитного индекса на основе математических критериев выявления типологической кластерной структуры, отражающей региональные диспропорции составляющих демографического процесса.

Предложенный подход к построению упорядоченной классификации региональных объектов целесообразно применять для создания

інформаційної технології, орієнтованої на підтримку процесу прийняття рішень при розробці програм регіонального розвитку.

1. Методика вимірювання людського розвитку регіонів України; затверджена Постановою Колегії Держкомстату України та Президії НАН України від 14.03.2001 р. № 76 [Електронний ресурс]. — Режим доступу: http://www.ukrstat.gov.ua/od_polog/metod_doc/sp/sp_04.pdf.
2. Регіональна статистика. Статистичний збірник «Регіональний людський розвиток» [Електронний ресурс]. — Режим доступу: https://ukrstat.org/uk/druk/publicat/Arhiv_u/15/Arch_rir_zb.htm
3. Лібанова Е.М. Людський розвиток в Україні: трансформація рівня життя та регіональні диспропорції / Інститут демографії та соціальних досліджень НАН України . — У 2-х томах. — К.: ІДСД ім. М. В. Птухи НАН України. — 2012. — 436С.
4. Porter M. E. Location, competition, and economic development: Local clusters in a global economy // *Economic Development Quarterly*. — 2000. — Vol. 14. — №1. — pp. 15–34.
5. Изард У. Методы регионального анализа: Введение в науку о регионах — М.: Прогресс, 1960. — 660 С.
6. Гранберг А.Г. Основы региональной экономики. 4-е изд. — М.: Изд. дом ГУ ВШЭ, 2004. — 495 С.
7. Людський розвиток регіонів України: методика оцінки і сучасний стан / Лібанова Е.М., Власенко Н. С., Власюк О. С. та ін. — К.: 2002. — 110 С.
8. Герасимчук З.В., Регіональна політика соціально-орієнтованого розвитку економіки: теорія, методологія, практика — Луцьк: РВВ ЛНТУ, 2011. — 260С.
9. Айвазян С.А. Интегральные индикаторы качества жизни населения: их построение и использование в социально-экономическом управлении и межрегиональных сопоставлениях — Рос. акад. наук, Центр. экон.-мат. ин-т. — М.: ЦЭМИ РАН, 2000. — 117 с.
10. Hagerty M. R., Land K. C. Constructing Summary Indices of Quality of Life A Model for the Effect of Heterogeneous Importance Weights // *Sociological Methods Research* — 2007. vol. 35. — № 4. — pp.455–496.
11. Шишкін В.С. Концептуальні та методологічні підходи до вивчення регіональних особливостей людського розвитку // *Статистика України*. — 2001. — № 1. — С. 30–38.
12. Макарова О.В., Гладун О.М. Регіональний індекс людського розвитку: причини та напрями удосконалення методики розрахунку // *Статистика України*. — 2012. — № 1. — С.10–15.
13. Григорук П.М., Ткаченко І.С. Методи побудови інтегрального показника // *Бізнес Інформ*. — 2012. — № 4. — С. 34–38.
14. Макоцьоба М.В. Інтегральна оцінка, рейтинг та кластеризація регіонів України за рівнем прогресу в людському розвитку // *Економіка и управление*. — 2012. — №4. — С. 144–151.
15. Огліх В.В, Єфанова Т.І. Регіональний простір: методичний підхід до оцінки внутрішньорегіональної диференціації // *Регіональна економіка*. — 2014. — № 4. — С. 40–47.
16. Composite Indicators Research Group — Режим доступу: <https://composite-indicators.jrc.ec.europa.eu/?q=content/overview>
17. Handbook on constructing composite indicators. Methodology and user guide OECD/JRC (2008) — [Електронний ресурс]. — Режим доступу: — <http://www.oecd.org/std/42495745.pdf>
18. Freudenberg M. Composite indicators of country performance: a critical assessment, OECD — Paris. STI WORKING PAPER. — 2003. — №16. — 34 P. — [Електронний ресурс]. — Режим доступу: https://www.itu.int/osg/spu/ni/wsisbridges/linked_docs/Background_papers/otherdocs/OECD_WP_2003_16.pdf
19. Типология российских регионов / Б. Бутс, С. Дробышевский, О. Кочеткова и др. — К. — М.: Институт экономики переходного периода, CEPRA. — 2002. — 159С.

20. Игнатъева Е.Д., Мариев О.С. Методический подход к анализу устойчивости регионального развития с использованием самоорганизующихся карт // Экономика региона. — 2008. — №2. — С.116–129.
21. Saisana M., Saltelli A., Tarantola S. Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators // J Royal Statistical Society — 2005. — 168(2) — pp.307–323.
22. Ishizaka A, Nemery P. Multi-criteria decision analysis: methods and software. 1st ed. Chichester: Wiley, 2013. — 310 P.
23. Саати Т. Принятие решений. Метод анализа иерархий — М.: Радио и Связь. — 1993. — 278С.
24. Миронова Н.А. Интеграция модификаций метода анализа иерархии для систем поддержки принятия групповых решений // Радіоелектроніка, інформатика, управління. — 2011. — №2. — С. 47–54.
25. Smet, Y.D., Nemery, P., Selvaraj, R.: An exact algorithm for the multicriteria ordered clustering problem // Omega. — 2012. — Vol. 40. — № 6. — pp.861–896.
26. Миркин Б.Г. Методы кластер анализа для поддержки принятия решений: обзор: препринт WP7/2011/03 Б.Г. Миркин; Национальный исследовательский университет «Высшая школа экономики». — М.: Изд. дом Национального исследовательского университета «Высшая школа экономики», 2011. — 88 с.
27. Fernandez. E., Navarro J., Bernal S. Handling multicriteria preferences in cluster analysis // European Journal of Operational Research. — 2010. — Vol. 202. — № 3. — pp. 819–827.
28. Методика вимірювання регіонального людського розвитку. — К.: Ін-т демографії та соціальних досліджень НАН України, Державна служба статистики України, 2012. — 41 С. [Електрон. ресурс] — Режим доступу: <http://www.idss.org.ua/>
29. Кривова О.А., Курило И.А. Региональная типология естественного прироста населения Украины: кластерный подход // Кибернетика и вычислительная техника. — 2011. — Вып. 164. — С. 89–102.

UDC 004.9:314.303

COMPLEX ESTIMATION OF REGIONAL DEMOGRAPHIC DEVELOPMENT

O.A. Krivova, L.M. Kozak

International Research and Training Center for Information Technologies and Systems of the National Academy of Sciences of Ukraine and Ministry of Education and Science of Ukraine (Kiev)

Introduction. Several studies are being conducted in the world to measure developmental disparities between countries, regions and territorial units. Composite indicators (or indexes) are used whenever a lot of variables are needed for evaluating developmental disparities between territories. Demographic variables are considered as important indicators of socio-economic development of regions. We show how cluster analysis can be combined with elements of multi-criteria decision analysis (MCDA) to construct composite index regional demographic development of Ukraine.

The purpose of this article is the development of regional socioeconomic systems analysis methodology and construction of composite indicators of regional demographic development.

Results. We have used 5 territorial social-demographic indicators: 1) total fertility rate; 2) death rate of children under age of five; 3) life expectation at birth;

4) survivorship probabilities for men from 20 to 65 years; 5) survivorship probabilities for women from 20 to 65 years. The following strategy can be pursued in order to construct composite index. First, a cluster analysis (algorithms Ward and K-means) for defining clusters of regions based on the value of the individual indicators is used. The result of the cluster analysis is typological clusters of the selected regions. Second, such as each cluster can be characterized with a centroid, these centroids must be ordered from best to worst. Weights of composite index are calculated as coefficients of the best linear regression model of preference function.

Conclusion. The composite index of regional demographic development allows to assess the degree of variance in regional demographic development and ranking of regions.

Keywords: clustering, a composite indicator, the index of regional demographic development, ordered classification.

1. Methods of measuring human development regions of Ukraine; Resolution adopted by the College of the State Statistics Committee of Ukraine and Presidium of NAS of Ukraine 14.03.2001 p. № 76 Available at: http://www.ukrstat.gov.ua/metodod_polog/metod_doc/sp/sp_04.pdf. (in Ukrainian).
2. Regional statistics. Statistical Yearbook Regional Human Development Available at: https://ukrstat.org/uk/druk/publicat/Arhiv_u/15/Arch_rir_zb.htm. (in Ukrainian).
3. Libanova E.M. Human development in Ukraine: transformation of the standard of living and regional disproportion — Kyiv: IDSS NANU — 2012. — 436 p. (in Ukrainian).
4. Porter M. E. Location, competition, and economic development: Local clusters in a global economy // *Economic Development Quarterly* — 2000. — vol. 14. — №1 — pp. 15–34.
5. Izard W. Methods of Regional Analysis: An Introduction to Regional Science. — M.: Progress — 1960. — 660 p. (in Russian).
6. Granberg A.G. Fundamentals of regional economy. 4th ed. — M.: GU HSE — 2004. — 495 p. (in Russian).
7. Libanova E.M. Human development of the regions of Ukraine: assessment methodology and current state — Kyiv. — 2002. — 110 p. (in Ukrainian).
8. Gerasymchuk Z.V. Regional policy of sustainable development: theory, methodology, practice — Lutsk: LNTU — 2011. — 260 p. (in Ukrainian).
9. Ayvazian S. A. Integrated indicators of the quality of life of the population: their construction and use in socio-economic governance interregional comparisons — M.: TsEMI RAN — 2000. — 117 p. (in Russian).
10. Hagerty. M.R., Land K.C. Constructing Summary Indices of Quality of Life A Model for the Effect of Heterogeneous Importance Weights // *Sociological Methods Research* — 2007. — vol. 35. — № 4. — pp.455–496.
11. Shishkin V.S. Conceptual and methodological approaches to regional disparities in human development // *Statistics of Ukraine* — 2001. — № 1. — pp. 30–38. (in Ukrainian).
12. Makarova, O.V., Hladun O.M. Regional Human Development Index: Causes and areas of improvement methods of calculation // *Statistics of Ukraine* — 2012. — № 1. — pp. 10–15. (in Ukrainian).
13. Grygoruk P.M., Tkachenko I.S. Methods of Integral Index Construction // *Business Inform* — 2012 — №4. — pp. 34–38. (in Ukrainian).
14. Makotsoba M.B. Integrated estimation, ranking and clustering regions of Ukraine in terms of progress in human development // *Economy and management* — 2012. — № 4. — pp. 144–151. (in Ukrainian).
15. Ohlikh V.V., Yefanova T.I. Regional space: methodical approach to the assessment of intraregional differentiation // *Regional economy* — 2014. — № 4. — pp. 40–47. (in Ukrainian).

16. Composite Indicators Research Group Available at: <https://composite-indicators.jrc.ec.europa.eu/?q=content/overview>
17. Handbook on constructing composite indicators. Methodology and user guide OECD/JRC. 2008 Available at: <http://www.oecd.org/std/42495745.pdf>
18. Freudenberg M. Composite indicators of country performance: a critical assessment — Paris: OECD, STI WORKING PAPER — 2003. — № 16. — 34 p.
https://www.itu.int/osg/spu/ni/wsisbridges/linked_docs/Background_papers/otherdocs/OECD_WP_2003_16.pdf
19. The typology of Russian regions/B. Boots, S. Drobyshevskiy, O. Kochetkova — M.: Gaidar institute for Economic Policy, CEPRA — 2002. — 159 p. (in Russian).
20. Ignatieva A., Maryev O. Methodical attitude to analyses of regional development stability on the basis of Kohonen's self-organizing maps//*Economy of region*, Institute of Economic — Yekaterinburg — 2008. — № 2. — pp. 116–129. (in Russian).
21. Saisana M., Saltelli A., Tarantola S. Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators//*Journal of the Royal Statistical Society* — Blackwell Publishing — 2005. — vol. 168. — № 2. — pp. 307–323.
22. Ishizaka A, Nemery P. Multi-criteria decision analysis: methods and software. 1st ed. — Chichester: Wiley — 2013. — 310 p.
23. Saaty T. Decision Making. The Analytic Hierarchy Process — M.: Radio Sviaz. — 1993. — 278 p. (in Russian).
24. Mironova N. O. Integration modifications of the analytical hierarchy process for group decision making support systems//*Radio Electronics, Computer Science, Control* — Zaporizhzhya, National Technical University — 2011. — № 2. — pp. 47–54. (in Russian).
25. Smet Y.D., Nemery P., Selvara R. An exact algorithm for the multicriteria ordered clustering problem//*Omega*. — 2012. — vol. 40. — № 6. — pp. 861–896.
26. Mirkin B. G. Methods of cluster analysis for support decision-making. Review. Preprint WP7/2011/ 03 Available at: http://www.hse.ru/data/2013/03/23/1303511006/WP7_2011_03f.pdf. (in Russian).
27. Fernandez. E., Navarro J., Bernal S. Handling multicriteria preferences in cluster analysis//*European Journal of Operational Research*. — 2010, — vol. 202, — №3, — pp. 819–827.
28. Methods of measuring human development regions of Ukraine — Kyiv: IDSS NANU, State Statistics Service of Ukraine. — 2012. — 41 p. Available at: <http://www.idss.org.ua/>
29. Krivova O.A., Kurilo I.A. The regional typology the natural growth of the population of Ukraine: Cluster approach//*Cybernetics and Computer Engineering* — 2011. — №. 164. — pp. 89–102. (in Russian).

Получено 01.06.2015

СИНТЕЗ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ НА ПРИНЦИПАХ САМООРГАНИЗАЦИИ МОДЕЛЕЙ

Е.А. Настенко, А.Л. Бойко, Е.К. Носовец, К.И. Тепляков,
В.А. Павлов

Национальный технический университет Украины «КПИ»

В работе решается задача структурно-параметрического синтеза модели логистической регрессии. Предложенный алгоритм осуществляет автоматическую оптимизацию параметров шагового алгоритма многомерной логистической регрессии на принципах самоорганизации моделей. Оптимизация параметров осуществляется с помощью предложенного внешнего критерия баланса, отражающего точность классификации на обучающей и проверочных выборках, с одной стороны, и требование к балансу качества распознавания в каждом классе, с другой. Рассмотрен пример моделирования классификатора функциональных состояний сердечнососудистой системы человека. Сравнение результатов моделирования стандартным и предложенным алгоритмами показало преимущество последнего на экзаменационной выборке данных.

Ключевые слова: логистическая регрессия, шаговый алгоритм, оптимизация параметров, принципы самоорганизации моделей, внешний критерий, функциональное состояние, сердечнососудистая система.

У роботі вирішується завдання структурно-параметричного синтезу моделі логістичної регресії. Запропонований алгоритм здійснює автоматичну оптимізацію параметрів крокового алгоритму багатовимірної логістичної регресії на принципах самоорганізації моделей. Оптимізація параметрів здійснюється за допомогою запропонованого зовнішнього критерію балансу, який відображає точність класифікації на навчальній та перевірочних вибірках, з одного боку, та вимогу до балансу якості розпізнавання в кожному класі, з іншого. Розглянуто приклад моделювання класифікатора функціональних станів серцево-судинної системи людини. Порівняння результатів моделювання за стандартним та запропонованим алгоритмами показало перевагу останнього на екзаменаційній вибірці даних.

Ключові слова: логістична регресія, кроковий алгоритм, оптимізація параметрів, принципи самоорганізації, зовнішній критерій, функціональний стан, серцево-судинна система.

ВВЕДЕНИЕ

Классические алгоритмы шаговой многомерной регрессии в задачах аппроксимации и классификации [1] реализуют процедуру включения и исключения аргументов в/из модели, исходя из заданных параметров: пороговых значений вероятности ошибок 1-го и 2-го родов или соответствующих пороговых значений критерия Фишера — $F_{\text{вкл}}$ и $F_{\text{искл}}$. При этом задание параметров остается за пользователем, что приводит к субъективному характеру получаемой модели: в зависимости от уровня, типа распределения и точек приложения шумов в данных, при одних и тех же значениях уровней значимости можем получить как переобученную модель, так и модель недостаточной сложности. Принципы решения поставленной проблемы были разработаны в теории самоорганизации [2, 3] на основе введения в алгоритм структурно-параметрического синтеза внешних

критериев — критериев регулярности и несмещенности [2], в методах индуктивного моделирования — критерии Акаике [4], Шварца [5], Маллоуза [6], на основе которых с успехом решается проблема единственности структуры в задачах аппроксимации и прогноза. В настоящей работе рассматривается возможность автоматической оптимизации на принципах самоорганизации параметров шагового алгоритма многомерной регрессии на примере синтеза логистической модели.

ЦЕЛЬ РАБОТЫ

Повышение качества классификации шаговых логистических регрессионных моделей за счет автоматической оптимизации параметров алгоритма многомерной бинарной логистической регрессии.

ПОСТАНОВКА ЗАДАЧИ

Задана матрица входных наблюдений $x \in R^M$ и вектор зависимой переменной $Y \in \{0, 1\}$

$$\begin{pmatrix} x_{11} & \dots & x_{1M} & y_1 \\ x_{21} & \dots & x_{2M} & y_2 \\ \dots & \dots & \dots & \dots \\ x_{n1} & \dots & x_{nM} & y_n \end{pmatrix},$$

где n — число наблюдений, M — число переменных, из которых необходимо выбрать наилучших m объясняющих аргументов в модель.

Необходимо построить алгоритм структурно — параметрического синтеза модели оптимальной структуры:

$$Y = \frac{1}{1 + \exp^{-z}}, \quad z = b_0 + \sum_{j=1}^m b_j x_{ij}, \quad \text{где } m < M. \quad (1)$$

Оптимальность структуры модели понимается в смысле достижения наилучшего значения рационально выбранного внешнего критерия на тестовой выборке данных.

МОДИФИКАЦИЯ КЛАССИЧЕСКОГО АЛГОРИТМА ШАГОВОЙ РЕГРЕССИИ НА ПРИНЦИПАХ САМООРГАНИЗАЦИИ МОДЕЛЕЙ

В работе предлагается модифицировать на принципах самоорганизации моделей классический алгоритм шаговой регрессии типа Stepwise [6]. Данный алгоритм выбран по следующим причинам.

1. Stepwise реализует наиболее распространенный метод структурно-параметрического синтеза, поэтому его эффективная модификация будет легко воспринята пользователями программного продукта.

2. Метод является одним из первых средств индуктивного моделирования, где используется явный штраф за сложность модели (выражение $d = n - m - 2$ в статистике Фишера при оценке эффективности вводимого/выводимого аргумента). Недостаток такого механизма — крайне низкая чувствительность штрафа при $n \gg m$ (n — количество точек, m — сложность модели), что, как правило, и имеет место в практических задачах. В работе предлагается увеличить точность модели на "свежих" (проверочной, экзаменационной) выборках за счет использования принципов самоорганизации моделей [2]: выделение проверочной выборки, используемой для расчета внешнего критерия, осуществление селекции моделей (аргументов) в соответствии с механизмом неявного штрафа, введение внешнего критерия оценки структуры модели.

3. Алгоритм Stepwise не является "жадным", как подавляющее число алгоритмов метода группового учета аргументов [3], реализующих принципы самоорганизации моделей. Это позволяет предполагать высокую эффективность его модификации.

Как отмечено выше, основной проблемой классической процедуры Stepwise есть "ручное" задание оператором моделирования пороговых значений критерия F (или уровней значимости α), используемых при включении или исключении аргумента модели. В работе предлагается задаться сеткой параметров $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$ и, включив в шаговую процедуру Stepwise расчет внешнего критерия, выбирать модели при тех $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$, при которых будут получены наилучшие значения внешнего критерия. Внешний критерий должен включать результаты классификации на проверочной выборке данных. Окончательный вариант модели может быть получен ее пересчетом уже на полной выборке наблюдений, при найденных выше оптимальных значениях вероятностей ошибок первого и второго родов $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$.

Этапы предлагаемого в работе алгоритма выглядят следующим образом.

1. Вводится расширенная (нелинейными аргументами) матрица переменных x .

2. Выравнивается количество наблюдений в классах. Этап введен ввиду несбалансированности результатов классификации логистической моделью, полученной стандартной процедурой итерационного взвешенного метода наименьших квадратов (МНК), при неравенстве наблюдений в классах.

3. Полная выборка наблюдений W делится на обучающую (А), проверочную (В) и экзамен (С) в заданном соотношении случайным образом.

4. Задается текущее значение $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$ сетки оптимизируемых параметров алгоритма.

5. Для текущего значения параметров алгоритма проводится модифицированная шаговая процедура с определением структуры, коэффициентов логистической модели и расчетом значения внешнего критерия $I_{\text{вн}}$. Решением шаговой процедуры считается модель для которой получен максимум внешнего критерия. Структура, коэффициенты и значение внешнего критерия $I_{\text{вн}}$ запоминаются.

6. Если все значения сетки перебраны, переходим к п.7, если нет —

задается следующее значение сетки оптимизируемых параметров алгоритма, переход к п. 4.

7. Из полученных моделей выбирается та, для которой получено наилучшее значение внешнего критерия.

8. Окончательной оценкой качества полученной модели вида (1) считаем качество классификации на экзаменационной выборке.

9. Осуществляем пересчет модели на полной выборке W для найденных оптимальных параметров $\alpha_{\text{искл}}$ и $\alpha_{\text{вкл}}$.

Ниже более подробно описаны отдельные этапы алгоритма:

Этап 3: выборка наблюдений W делится на обучающую A , проверочную B и экзамен C в заданном соотношении случайным образом. Расчет коэффициентов логистических моделей внутри шаговой процедуры будет проводиться на обучающей выборке, значение внешнего критерия для выбора структуры модели будет оцениваться как сбалансированная на классах точность классификации на обучающей и проверочной выборках, результирующая оценка — точность классификации на экзаменационной выборке.

Этап 5: модифицированная шаговая процедура для фиксированных значений $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$ состоит из нескольких шагов.

1. Включение предиктора в модель:

1.1. Проводим F -тест, сравнивая модель, полученную на предыдущей итерации, с моделью, включающую предиктор (переменную) x_i для каждого предиктора, еще не включенного в модель:

$$F_i = \frac{SSR_{\text{prev}+x_i} - SSR_{\text{prev}}}{MSR_{\text{prev}+x_i}}, \quad (2)$$

$$\text{где } MSR_{\text{prev}+x_i} = \frac{SSR_{\text{prev}+x_i}}{d}, \quad d = n - m - 2, \quad (3)$$

$$SSR = \sum_{i=1}^n \bar{Y}_i - Y, \quad (4)$$

$i = 1, \dots, k_I, k_I$ — количество предикторов претендентов ранее не включенных в модель, Y — табличное значение выходной переменной, Y_i — значение регрессионной модели, n — количество наблюдений в выборке, m — количество переменных в модели, индекс prev — означает модель, полученную на предыдущей итерации.

1.2. Выбираем предиктор с наибольшим значением F . Если уровень значимости α , соответствующий полученному значению критерия F , меньше фиксированного $\alpha_{\text{вкл}}$ ($\alpha < \alpha_{\text{вкл}}$), то принимается гипотеза о включении предиктора в модель.

2. Исключение предикторов:

2.1. Проводим F -тест, сравнивая текущую модель с моделью, не включающую предиктор x_i для каждого предиктора из модели:

$$F_i = \frac{SSR_{\text{curr}} - SSR_{\text{curr}-x_i}}{MSR_{\text{curr}}} \quad (5)$$

$$MSR_{\text{curr}} = \frac{SSR_{\text{curr}}}{d} \quad (6)$$

где $i = 1, \dots, k_2$, k_2 — количество предикторов ранее включенных в модель, индекс curr — означает модель, полученную на текущей итерации.

2.2. Если уровень значимости α , соответствующий полученному значению критерия F , превышает фиксированный $\alpha_{\text{искл}}$ ($\alpha > \alpha_{\text{искл}}$), то принимается гипотеза об исключении предиктора из модели. Среди переменных прошедших F -тест для исключения выбираем предиктор с наименьшим значением F .

Для каждой из сравниваемых выше моделей, их коэффициенты считаются по итерационному взвешенному методу наименьших квадратов [8] на обучающей выборке данных.

Рассчитываем и запоминаем значение внешнего критерия $I_{\text{вн}}$ для модели полученной на данной итерации. Если предикторы для включения не исчерпаны, переходим к п. 2.1., если исчерпаны — к п. 2.3.

2.3. Для фиксированного значения параметров сетки $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$ запоминается модель с наилучшим значением внешнего критерия $I_{\text{вн}}$.

Определим далее целесообразную форму внешнего критерия. Отразим в нем требования к точности классификации на обучающей и проверочных выборках, с одной стороны, и требование к балансу качества распознавания в каждом классе, с другой. Тогда внешний критерий (назовем его критерий баланса качества классификации — БКК) имеет вид:

$$I_{\text{бн}} = \alpha \Delta^A + (1 - \alpha) \Delta^B, \quad (7)$$

где α — коэффициент баланса качества классификации на обучающей и проверочной выборках, Δ^A — критерий качества классификации на обучающей выборке A , Δ^B — критерий качества классификации на проверочной выборке B .

Критерий Δ^A рассчитывается на обучающих точках A для фиксированной структуры и соответствующей подматрицы X :

$$\Delta^A = (1 - \gamma) * (\Delta^{A1} + \Delta^{A2}) + \gamma * \frac{1}{1 + |\Delta^{A1} - \Delta^{A2}|}. \quad (8)$$

Критерий Δ^B рассчитывается для проверочных точек B и моделей с коэффициентами, найденными на обучающей выборках A :

$$\Delta^B = (1 - \gamma) * (\Delta^{B1} + \Delta^{B2}) + \gamma * \frac{1}{1 + |\Delta^{B1} - \Delta^{B2}|}. \quad (9)$$

где r — коэффициент баланса между суммарным качеством распознавания по классам и симметричностью (относительно каждого класса) полученного результата;

$$\Delta^{A_1} = \frac{n^{A_1^*}}{n^{A/2}} = \frac{2n^{A_1^*}}{n^A}, \quad \Delta^{A_2} = \frac{n^{A_2^*}}{n^{A/2}} = \frac{2n^{A_2^*}}{n^A}, \quad (10)$$

$$\Delta^{B_1} = \frac{n^{B_1^*}}{n^{B/2}} = \frac{2n^{B_1^*}}{n^B}, \quad \Delta^{B_2} = \frac{n^{B_2^*}}{n^{B/2}} = \frac{2n^{B_2^*}}{n^B}, \quad (11)$$

где $n^{A_1^*}$, $n^{B_1^*}$ — количество правильно распознанных объектов 1-го класса;
 $n^{A_2^*}$, $n^{B_2^*}$ — количество правильно распознанных объектов 2 класса;
 n^A , n^B — количество объектов в выборке.

СРАВНИТЕЛЬНАЯ ОЦЕНКА РАБОТЫ АЛГОРИТМОВ ПРИ РЕШЕНИИ ЗАДАЧИ ПОСТРОЕНИЯ КЛАССИФИКАТОРА ФУНКЦИОНАЛЬНЫХ СОСТОЯНИЙ СЕРДЕЧНОСОСУДИСТОЙ СИСТЕМЫ

Для сравнительной оценки базового и предложенного алгоритмов решалась следующая задача: получить классифицирующую функцию, выделяющую группу испытуемых с некоторым определенным функциональным состоянием сердечно-сосудистой системы (23 наблюдения, значение классифицирующей переменной "1") от остальной выборки испытуемых (138 наблюдений, значение классифицирующей переменной "0"). Для расчета были взяты данные, полученные в лаборатории функциональной диагностики кафедры Физического воспитания НТУУ «КПИ». Выборка содержит 161 наблюдение и 22 переменных, характеризующих антропометрические параметры и психо-физиологические тесты, проведенные для студентов 2-го курса университета.

Рассмотренный выше алгоритм реализован в программной среде R. Ниже приведены результаты расчета логистической модели и сравнения качества классификации стандартным алгоритмом логистической регрессии glm в программной среде R с предложенной в работе версией алгоритма шаговой регрессии на принципах самоорганизации моделей с оптимизацией параметров $\alpha_{\text{вкл}}$ и $\alpha_{\text{искл}}$ (АШРО).

Выборка была разбита на обучающую, проверочную и экзаменационную в пропорции — 55/30/15. Модели логистической регрессии сравниваемыми алгоритмами считались на рабочей выборке (обучающая + проверочная), а окончательная оценка алгоритмов осуществлялась на экзаменационной выборке.

На рис. 1 приведен график чувствительности (Sensitivity) и специфичности (Specificity), а также ROC кривая результатов работы стандартного алгоритма glm логистической регрессии в R на экзаменационной выборке.

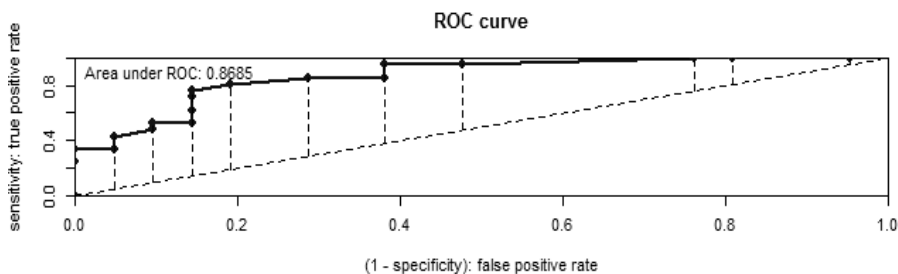
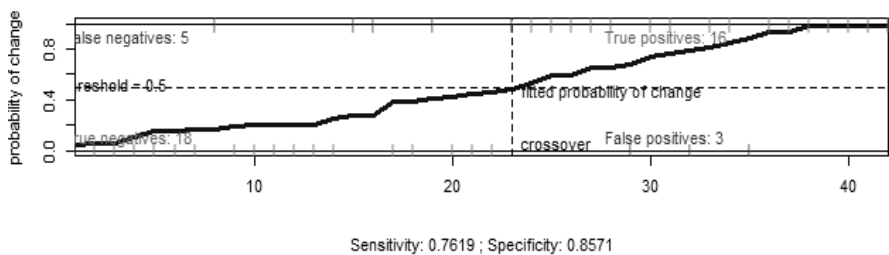


Рис. 1. Результаты работы стандартного алгоритма glm

Стандартный алгоритм показал на экзаменационной выборке качество классификации — 81 %, область под ROC-кривой — 0,87.

На рис. 2 приведены графики чувствительности (Sensitivity) и специфичности (Specificity) и ROC-кривая для АШПО рассчитанные на экзаменационной выборке данных. Качество классификации данного алгоритма — 90,5 %, область под ROC — кривой — 0,97.

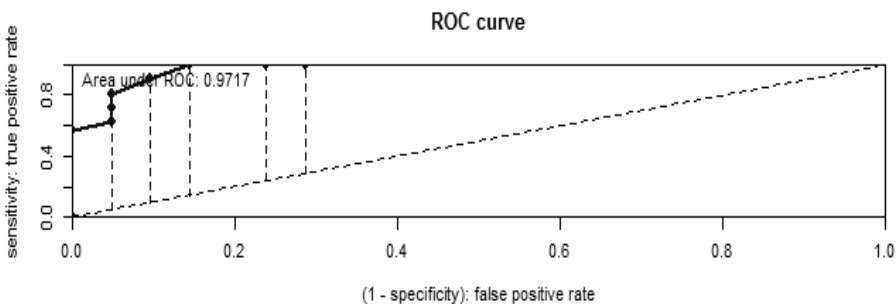
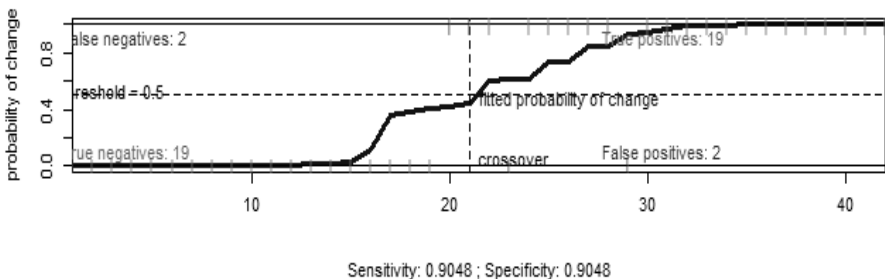


Рис. 2. Результаты работы АШПО алгоритма

Выводы

В решении задачи классификации функциональных состояний сердечнососудистой системы, предложен алгоритм синтеза логистической регрессии на принципах самоорганизации моделей, который при сравнении со стандартным шаговым алгоритмом логистической регрессии, показал улучшение результатов классификации экзаменационной выборки на 10 %,. Эффект получен за счет оптимизации параметров шагового алгоритма в соответствии с предложенным в работе внешним критерием, Критерий отражает требования к балансу точности классификации на обучающей и проверочных выборках, с одной стороны, и балансу качества распознавания в каждом классе с другой.

1. Стрижов В.В., Крымова Е.А. Методы выбора регрессионных моделей — М.: ВЦ РАН. — 2010. — С. 20–23.
2. Ивахненко А.Г., Степашко В. С. Помехоустойчивость моделирования — Киев.: Наук.думка. — 1985. —216 С.
3. Ивахненко А.Г., Мюллер Й.А. Самоорганизация прогнозирующих моделей — К.: Техніка ; Берлин : Фоб Ферлаг техник. — 1984. — 223 С.
4. Akaike H. A new look at the statistical model identification // IEEE Transactions on Automatic Control. — 1974. — Vol. 19. — P. 716–723.
5. Schwarz E. Estimating the dimension of a model // Annals of Statistics. — 1978. — Vol. 6, № 2. — P. 461–464.
6. Mallows C.L. Some Comments on CP / C.L. Mallows // Technometrics — 1973. — Vol. 15, № 4 — P. 661–675.
7. Efroymson M.A. Multiple regression analysis //Mathematical Methods for Digital Computers, — 1960.
8. Green P.G. Iteratively Reweighted Least Squares for Maximum Likelihood Estimation, and some Robust and Resistant Alternatives (with discussions) // Journal of the Royal Statistical Society, Series — 1984. — B 46, — P. 149–192.

UDC 0.04:519.584:57.041

SYNTHESIS OF LOGISITIC REGRESSION, BASED ON SELF-ORGANISATION PRINCIPLES OF MODELS

E.A. Nastenکو, A.L. Boyko, O.K. Nosovets, K.I. Teplyakov, V.A. Pavlov

National Technical University of Ukraine “Kiev Polytechnical Institute” (Kiev)

Introduction. Requirements for modeling algorithms and their implementations varies depending upon the desired properties of the models, which has to be received in restrictions on the available computational resources. Examples of desired properties — accuracy, efficiency ratings, the lowest sensitivity to a change in the data of the model error, variance estimation of parameters, p values etc. Depending on the specific use of models, those or other criteria are taken as a basis for designing specific algorithm simulation. However, choice of the solution of resulting model is usually left to the user. This article considers the possibility of stepwise regression algorithm’s automatic optimization of parameters that is based on principles of self-organization on an example of the synthesis of the logistic model.

The purpose of this article is the improvement the quality of logistic regression

© E.A. Настенко, А.Л. Бойко, Е.К. Носовец, К.И.Тепляков, В.А. Павлов, 2015

classification models due to automatic optimization multivariate binary logistic regression algorithm parameters.

Results. The essence of the modification of stepwise logistic regression standard algorithm: defines p_{enter} , p_{leave} grid for each combination of the thresholds calculates stepwise logistic algorithm and the corresponding value of the external criteria. Proposed external criteria reflects the classification accuracy on the training and test datasets, on the one hand, and the requirement to balance the quality of recognition in each class on the other. The stated procedure is repeated for the next value of the grid parameters of the algorithm. Final evaluation of the model is given in the exam sample data. For logistic model calculation and quality's comparison of classification between standard logistic regression (glm function in R software) and proposed version of modified stepwise algorithm were taken data obtained in the laboratory of functional diagnostics at Department of Physical Education NTUU "KPI". The purpose of the example is to get a classifying function, of group of subjects with certain states of the cardiovascular system from the rest of the test sample. Standard algorithm demonstrated on examination sample classification quality — 81%, the area under the ROC — curve — 0.8685. Graphs of sensitivity and specificity, and ROC curve for modified algorithm showed the results: quality of the classification algorithm — 90.5 %, area under the ROC — curve — 0.9717.

Conclusions. Article proposes stepwise logistic regression based on the principles of self-organization synthesis algorithm. In order to optimize the parameters of the algorithm proposed by external criterion, which reflects the classification accuracy on the training and test samples and requirement to balance the quality of recognition in each class the effect was received. For the aboved example the classification of functional states of the cardiovascular system in comparison of the standard stepwise algorithm with the proposed algorithm has shown classification quality improvement on 10 % on examination sample.

Keywords: logistic regression, stepwise regression, self-organization's principles.

1. Strighov V., Krimova E., Selection methods of regression models — Moscow: CC RAS — 2010. — 45 p. (in Russian).
2. Ivakhnenko A., Stepushko V. Noisestability modelling — Kiev: «Nauk.dumka». — 1985, — 216 p. (in Russian).
3. Ivakhnenko A. Muller J. Self-organization of predictive models — Kiev: Technic. — 1984, — 223 p. (in Russian).
4. Akaike H.A new look at the statistical model identification // IEEE Transactions on Automatic Control — 1974. — Vol. 19. — P.716–723.
5. Schwarz E. Estimating the dimension of a model // Annals of Statistics — 1978. — Vol. 6. — № 2. — P. 461–464.
6. Mallows C.L. Some Comments on CP//Technometrics — 1973. — Vol. 15. — № 4. — P. 661–675.
7. Efrogmson M.A. Multiple regression analysis // Mathematical Methods for Digital Computers — 1960.
8. Green P.G. Iteratively Reweighted Least Squares for Maximum Likelihood Estimation, and some Robust and Resistant Alternatives (with discussions) // Journal of the Royal Statistical Society, Series — 1984. — B 46. — P. 149–192.

Получено 15.06.2015

АВТОРЫ НОМЕРА

Антомонов Михаил Юрьевич, доктор биол. наук, профессор, заведующий отделом медицинской информатики ГУ «Институт гигиены и медицинской экологии имени А.Н. Марзеева НАМН Украины», ул. Попудренко, 50, г. Киев-94, 02660, email: antomonov@gmail.com

Аралова Наталья Игоревна, кандидат техн. наук, старший научный сотрудник Института кибернетики им.И.В. Глушкова НАН Украины, проспект Глушкова, 40, г. Киев, 03680 ГСП, e-mail: aralova@ukr.net

Баловсяк Сергей Васильевич, кандидат физ.-мат наук, доцент, доцент кафедры компьютерных систем и сетей Черновицкого национального университета имени Юрия Фельдовича, ул. Коцюбинского, 2, 58012, г. Черновцы, Украина, e-mail: s.balovsyak@chnu.edu.ua.

Бойко Ганна Леонидовна, кандидат пед. наук, доцент, заведующий кафедрой физического воспитания факультета биомедицинской инженерии, Национального Технического Университета Украины «КПИ», Киев, 03056, пр. Победы, 37, e-mail: anna_anna2612@mail.ru

Козак Людмила Михайловна, доктор биол. наук, старший научный сотрудник, ведущий научный сотрудник отдела медицинских информационных систем Международного научно-учебного центра информационных технологий и систем НАН Украины и МОН Украины, пр. Академика Глушкова, 40, г. Киев, 03680 ГСП, e-mail: lmkozak52@gmail.com

Копец Мирослав Михайлович, кандидат физ.-мат. наук, доцент, доцент кафедры математической физики Национального технического университета Украины «Киевский политехнический институт», пр. Победы, 37, г. Киев, Украина, 03056, e-mail: optimal201214@yandex.ua

Кривова Ольга Анатольевна, научный сотрудник Международного научно-учебного центра информационных технологий и систем НАН Украины и МОН Украины, проспект Глушкова, 40, г. Киев, 03680 ГСП, e-mail: ol_krilova@yahoo.com

Луцык Ярослав Васильевич, аспирант кафедры физики твердого тела Черновицкого национального университета имени Юрия Фельдовича, ул. Коцюбинского, 2, 58012, г. Черновцы, Украина, e-mail: clavikfire@gmail.com

Настенко Евгений Арнольдович, кандидат техн. наук, доктор биол. наук, заведующий кафедрой биомедицинской кибернетики факультета биомедицинской инженерии Национального Технического Университета Украины «КПИ», Киев, 03056, пр. Победы, 37, e-mail: nastenko@inbox.ru

Носовец Елена Константиновна, кандидат техн. наук, ассистент кафедры биомедицинской кибернетики факультета биомедицинской инженерии Национального Технического Университета Украины «КПИ», Киев, 03056, пр. Победы, 37, e-mail: e.nosovets@yandex.ru

Павлов Андрей Владимирович, кандидат техн. наук, научный сотрудник отдела информационных технологий индуктивного моделирования Международного научно-учебного центра информационных технологий и систем НАН Украины и МОН Украины, пр. Академика Глушкова, 40, г. Киев, 03680 ГСП, e-mail: andriypavlove@gmail.com

Павлов Владимир Анатольевич, кандидат техн. наук, доцент, доцент кафедры биомедицинской кибернетики факультета биомедицинской инженерии Национального технического университета Украины «КПИ», Киев, 03056, пр. Победы, 37, e-mail: vpavlo@bk.ru

Соловей Юрий Николаевич, кандидат медицинских наук, докторант кафедры ухода за больными и высшего медсестринского образования Буковинского государственного медицинского университета, Театральная пл. 2, 58000, г. Черновцы, Украина, e-mail: aspirant-surgeon@ukr.net.

Суровцев Игорь Викторович, кандидат технических наук, старший научный сотрудник Международного научно-учебного центра информационных технологий и систем НАН Украины и МОН Украины, проспект Глушкова, 40, г. Киев, 03680 ГСП, e-mail: igorsur52@gmail.com

Тепляков Кирилл Игоревич, студент кафедры биомедицинской кибернетики факультета биомедицинской инженерии Национального Технического Университета Украины «КПИ», Киев, 03056, пр. Победы, 37, e-mail: kiriltepl@gmail.com

Фодчук Игорь Михайлович, доктор физ.-мат наук, профессор кафедры физики твердого тела Черновицкого национального университета имени Юрия Федьковича, ул. Коцюбинского, 2, 58012, г. Черновцы, Украина, e-mail: ifodchuk@ukr.net

Системный указатель статей, опубликованных в 2015 году

Информатика и информационные технологии

<i>Вовк М.И., Галян Е.Б.</i> Персоналифицированная биотехническая система для восстановления речи	179
<i>Доценко С.И.</i> Теоретико-игровой подход к задачам координации действий с обменом информации	179
<i>Суровцев И.В.</i> Преобразование структуры данных при определении концентрации методами инверсионной хронопотенциометрии	180
<i>Доценко С.И.</i> Решение задачи оптимального выбора с групповым просмотром с помощью теоретико-игрового подхода	180
<i>Савченко Е.А.</i> Технология решения задачи моделирования и прогнозирования на основе индуктивного подхода	180
<i>Файнзильберг Л.С., Ориховская К.Б.</i> Информационная технология оценки адаптационных резервов организма в полевых условиях	181
<i>Даник Ю.Г., Писарчук А.А., Тимчук С.В.</i> Ситуативный синтез автоматизированной системы сбора и обработки информации	181
<i>Суровцев И.В.</i> Метод цифровой фильтрации электрохимических сигналов в хронопотенциометрии	182
<i>Баловсяк С.В., Луцик Я.В., Соловей Ю.Н., Фодчук И.М.</i> Многоуровневый метод повышения локального контраста и удаления неоднородного фона изображений	182
<i>Павлов А.В.</i> Подход к организации рекуррентно-параллельных вычислений в системе автоматизированной структурно-параметрической идентификации	182

Интеллектуальное управление и системы

<i>Norkin V.V.</i> On the Approximation of Vector Optimization Problems	179
<i>Романенко В.Д., Милявский Ю.Л.</i> Стабилизация импульсных процессов в когнитивных картах сложных систем на основе модальных регуляторов состояния	179
<i>Городецкий В.Г., Осадчук Н.П.</i> Алгоритм реконструкции динамических систем по одной наблюдаемой переменной	179
<i>Мащенко С.О., Шушарин Ю.В.</i> Критерий минимаксного сожаления в задаче принятия решений с нечетким множеством состояний окружающей среды	180
<i>Павлов В.В., Волков А.Е., Волошенюк Д.А.</i> Инвариантная сетевая система управления конфликтными ситуациями воздушных кораблей на этапе захода на посадку	180
<i>Коцур М.П., Наконечный А.Г.</i> Оптимальное управление нестационарным режимом каскадного термоэлектрического охладителя	180
<i>Гриценко Ю.В., Скрипец А.Н., Тронько В.Д.</i> Анализ возможности управления траекторией движения самолета по корреляционной функции	181
<i>Zhiteckii L.S., Nikolaienko S.A., Solovchuk K.Yu.</i> Adaptation and Learning in Some Classes of Identification and Control Systems	181
<i>Мельничук С.В.</i> Метод идентификации многосвязных систем по частотным характеристикам	181
<i>Конец М.М.</i> Оптимальное управление колебаниями балки переменного поперечного сечения	182
<i>Аралова Н.А.</i> Математическая модель краткосрочной и среднесрочной адаптации системы дыхания лиц, работающих в экстремальных условиях высокогорья	182

Медицинская и биологическая кибернетика

<i>Наджафян Тумаджани М., Будник Н.Н., Коваленко А.С.</i> Оценка степени неомогенности электрических процессов в желудочках сердца по данным магнитокардиографии	179
<i>Пашиная С.Л., Антомонов М.Ю.</i> Обратная задача интегрального оценивания: выявление критических составляющих медико-экологической ситуации	179
<i>Чернышенко С.В., Рузич Р.В.</i> Дискретные эффекты в непрерывных моделях сукцессионных процессов	180
<i>Майоров О.Ю., Фенченко В.Н.</i> Мультифрактальный анализ в исследовании биоэлектрической активности мозга	181
<i>Антомонов М.Ю.</i> Метод определения эффектов совместного действия факторов среды в эпидемиологических исследованиях для бинарных переменных	182
<i>Кривова О.А., Козак Л.М.</i> Комплексная оценка регионального демографического развития	182
<i>Настенко Е.А., Бойко А.Л., Павлов В.А., Тепляков К.И.</i> Синтез логистической регрессии на принципах самоорганизации моделей	182

ВНИМАНИЮ АВТОРОВ!

Требования к рукописям статей

1. Рукопись предоставляется на белой бумаге в двух экземплярах (язык — русский, 9–10 с.) и электронная версия. К рукописи прилагаются:

— аннотации — на русском и украинском языках (УДК, фамилия, инициалы автора/ов, название статьи, 5–6 строк текста, ключевые слова), на английском языке (фамилия, инициалы автора/ов, место работы, город, страна, название статьи, от 250 слов, с выделением рубрик: введение, цель, результаты, выводы, ключевые слова);

— согласие автора/ов на публикацию;

— сведения об авторе должны включать: ФИО, ученую степень, научное звание, должность, отдел, место работы, почтовый адрес организации, телефон (служ., моб., дом.), E-mail.

2. Текст статьи подается с обязательными рубриками: введение, постановка задачи, цель, результаты, четко сформулированные выводы.

3. Поданные статьи проходят рецензирование ведущими специалистами в данной области.

Структура статьи в КВТ

УДК

НАЗВАНИЕ

И.О.Фамилия / и

Организация/и

Аннотация (русс., 5-10 строк)

Ключевые слова: (5-8 слов)

Анотація (укр., 5-10 рядків)

Ключові слова: (5-8 слів)

Введение

Цель

Постановка задачи

Результаты (1-3 раздела с названиями)

Выводы

Список литературы на языке оригинала (в порядке упоминания в тексте, по стандарту ДСТУ ГОСТ 7.1 : 2006, ДСТУ ГОСТ 7.80 : 2007).

Abstract 0,5-0,7 p.

ФИО, Название статьи на английском языке

Название организации – на английском языке.

С обязательными рубриками:

Introduction.

Purpose.

Results.

Conclusion.

Keywords:

Список литературы — перевод источников на английский язык, фамилии и инициалы авторов — транслитерация

(Author A.A. Article. *Journal*, 2000, vol. 1, no. 2, pp. 111–112.
Author A.A., Author B.B. *Book*. City: Publisher, 2000. 111 p. (in Russian).

Требования к текстовому файлу

Формат файла *.doc, *.rtf. Файл должен быть подготовлен с помощью текстового редактора Microsoft Word 2003.

Используемые стили: шрифт Times New Roman, высота 12 пт, межстрочное расстояние — полуторное. Формат бумаги А4, поля (слева, справа, сверху, внизу) 2 см.

Формулы набираются в редакторе формул Microsoft Equation Editor 3.0.

Опции редактора формул — (10,5; 8,5; 7,5; 14; 10). Ширина формул не более 12 см.

Рисунки должны быть достаточно качественными, созданы встроенным редактором рисунков Word Picture либо иными Windows-приложениями (в этом случае рисунки должны быть представлены отдельными файлами соответствующих форматов). Ширина рисунков не более 12 см.

Таблицы выполняются стандартным встроенным в Word инструментарием «Таблица».

В розничную продажу журнал не поступает

Оформление подписки для желающих опубликовать статьи в нашем журнале обязательно.

Подписка осуществляется:

- „Каталог видань України”, подписной индекс 86598
- в агентстве «Укринформнаука» НАН Украины ukrinformnauka@gmail.com

Підп. до друку 14.12.2015. Формат 70×108/16. Папір офсетний.
Ум. друк. арк. 10,33. Обл. вид. арк. 9,84. Тираж 100. Зам. №3211

Видавничий дім “Академперіодика” НАН України
01004, Київ 4, вул. Терещенківська, 4.
Свідоцтво про внесення до Державного реєстру суб’єкта видавничої справи
Серії ДК №544 від 27.07.2001 р.