

ISSN 2074-050X



ВІСНИК

ПРИКЛАДНОЇ

МАТЕМАТИКИ

МАТЕМАТИЧНОГО

МОДЕЛЮВАННЯ

2011

ISSN 2074-5897

Міністерство освіти і науки, молоді та спорту України

Дніпропетровський національний університет імені Олеся Гончара

ПІТАННЯ ПРИКЛАДНОЇ МАТЕМАТИКИ І МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ

Збірник наукових праць

Дніпропетровський
Висхідний Дніпропетровський
національний університет

2011

УДК 001, 004, 007, 378, 514, 519, 532, 534, 539, 621, 631, 637, 669, 681
ББК 22.12я431+22.311я431+22.176я431+22.18я431
П 32

*Друкується за ухвалою вченої ради
Дніпропетровського національного університету імені Олеся Гончара*

Уміщені результати фундаментальних досліджень та практичних розробок з проблем математичного та програмного забезпечення інтелектуальних систем. Статті присвячені питанням математичного моделювання складних прикладних систем і розробки ефективних обчислювальних методів та алгоритмів їхнього розв'язання, а також оптимізації, чисельних методів, функціонального аналізу, математичної фізики. Для науковців, викладачів вузів, може бути корисним аспірантам і студентам

П 32 Питання прикладної математики і математичного моделювання:
зб. наук. пр. / ред. кол. ... О. М. Кісельова (голов. ред.) та ін. – Д.: Вид-во Дніпропетр. нац. ун-ту, 2011. – 332 с.

Содержатся результаты фундаментальных исследований и практических разработок по проблемам математического и программного обеспечения интеллектуальных систем. Статьи посвящены вопросам математического моделирования сложных прикладных систем и разработке эффективных вычислительных методов и алгоритмов их решения, а также оптимизации, численных методов, функционального анализа, математической физики. Для ученых, преподавателей вузов, может быть полезен аспирантам и студентам

УДК 001, 004, 007, 378, 514, 519, 532, 534, 539, 621, 631,
637, 669, 681
ББК 2.12я431+22.311я431+22.176я431+22.18я431

Рецензенти:

д-р фіз.-мат. наук, проф. **В. В. Лобода**
д-р фіз.-мат. наук, проф. **С. Б. Вакарчук**

Редакційна колегія:

д-р фіз.-мат. наук, проф. **О.М. Кісельова** (головний редактор), д-р фіз.-мат. наук, проф. **О.О. Кочубей** (заст. відп. редактора), д-р фіз.-мат. наук, проф. **В. Є. Белозьоров**, д-р фіз.-мат. наук, проф. **В. О. Капустян**, д-р фіз.-мат. наук, проф. **П. І. Когут**, д-р фіз.-мат. наук, проф. **В. І. Кузьменко**, д-р фіз.-мат. наук, проф. **А.М. Пасічник**, д-р техн. наук, проф. **А. А. Босов**, д-р техн. наук, проф. **О. І. Михальов**, д-р техн. наук, проф. **Н. І. Ободан**, д-р техн. наук, проф. **О. П. Приставка**, д-р техн. наук, проф. **О. І. Федякін**, канд. фіз.-мат. наук, проф. **В. Д. Ламзюк**, канд. фіз.-мат. наук, доц. **Л. С. Коряшкіна** (відп. секретар)

© Дніпропетровський національний університет
імені Олеся Гончара, 2011
© Видавництво ДНУ, оформлення, 2011

УДК 577.4:517.9

¹А.Г. Балакирева, С.Н. Герасин, С.В. Яковлев
*Харьковский национальный университет радиоэлектроники
Харьковский национальный университет внутренних дел*

ИССЛЕДОВАНИЯ АСИМПТОТИЧЕСКИХ СВОЙСТВ НЕОДНОРОДНОЙ МОДЕЛИ ЛЕСЛИ

Розглядається неоднорідний добуток невід'ємних матриць та доводиться, що за певних умов цей добуток має властивість слабкої ергодичності. Уводиться поняття неоднорідної популяційної моделі Леслі та досліджуються асимптотичні властивості моделі. Установлюється, що неоднорідна модель Леслі має властивість слабкої ергодичності.

Рассматривается неоднородное произведение неотрицательных матриц и доказывается, что при определенных условиях данное произведение обладает свойством слабой эргодичности. Вводится понятие неоднородной популяционной модели Лесли и исследуются асимптотические свойства данной модели. Устанавливается, что неоднородная модель Лесли обладает свойством слабой эргодичности.

In this work the inhomogeneous product of non-negative matrices is considered and it is proved that under certain conditions the given product is weakly ergodic. The concept of inhomogeneous population Leslie model is entered and asymptotic properties of this model are investigated. It is ascertained that inhomogeneous population Leslie model has weak ergodicity property.

Ключевые слова: модель Лесли, слабая эргодичность, матрица, популяция.

Введение. Матричный аппарат для описания эволюции сложных многовидовых популяций был предложен П. Лесли. Формализм Лесли опирается на допущение, что популяция разбита на конечное число последовательных n возрастных классов одинаковой длительности, а численность всех классов изменяется в дискретном времени с равномерным шагом. Данная модель называется однородной моделью Лесли и ее свойства достаточно хорошо изучены [1– 5]. Однако, классическая однородная модель

Лесли имеет весьма жесткие ограничения для использования ее в прогнозировании динамики реальных популяций. В связи с этим мы распространяем данную модель на неоднородный случай, предполагая, что статистические показатели популяции изменяются на каждом шаге.

Неоднородная модель Лесли. Перед определением неоднородной модели Лесли введем понятие классической однородной модели Лесли. Однородная модель Лесли

$$X(t_{j+1}) = LX(t_j), \quad j = 0, 1, 2, \dots,$$

где $X(t_j) = [x_1(t_j) \quad x_2(t_j) \quad \dots \quad x_n(t_j)]^T$ вектор-столбец описывает структуру популяции в момент времени t_j . Элементы $x_i(t_j)$ – численность i -ой возрастной группы ($i = 1, \dots, n$) в момент времени t_j , если не учитывается разделение по полу, и численность самок i -ой группы в момент времени t_j , если половое разделение существенно для рассматриваемой популяции.

Матрица L называется переходной матрицей Лесли и имеет вид:

$$L = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{n-1} & \alpha_n \\ \beta_1 & 0 & \dots & 0 & 0 \\ 0 & \beta_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \beta_{n-1} & 0 \end{pmatrix},$$

где α_i ($\alpha_i \geq 0, i = 1, 2, \dots, n$) – возрастные коэффициенты рождаемости, (то есть средняя плодовитость особей i -й возрастной группы), и β_i ($0 < \beta_i \leq 1, i = 1, 2, \dots, n$) – коэффициенты выживания, равные вероятности перехода из возрастной группы i в $i+1$ группу к следующему моменту времени (причем $\sum_{i=1}^{n-1} \beta_i$ может быть больше 1).

Таким образом, зная структуру матрицы L и начальное состояние популяции (вектор_столбец $X(t_0)$), можно прогнозировать состояние популяции в любой наперед заданный момент времени.

$$\begin{aligned} X(t_1) &= LX(t_0), \\ X(t_2) &= LX(t_1) = LLX(t_0) = L^2 X(t_0), \\ &\dots\dots\dots \\ X(t_n) &= LX(t_{n-1}) = L^n X(t_0), \end{aligned}$$

где L^n – n -ая степень матрицы L .

В однородной модели Лесли переходная матрица L с течением времени остается неизменной, что не соответствует действительности. Предположим, что коэффициенты выживаемости и рождаемости популяции изменяются на каждом шаге, что соответствует изменению переходной матрицы Лесли. Тогда однородная модель Лесли модифицируется на неоднородный случай.

Неоднородную модель Лесли для прогнозирования развития популяции с течением времени имеет вид:

$$X(t_n) = L_{0,n} X(t_0), \quad L_{0,n} = L_1 \cdot L_2 \cdot \dots \cdot L_n, \quad n = 1, 2, \dots, \quad (1)$$

где L_i – матрица Лесли на i -ом шаге ($1 \leq i \leq n$).

Свойство эргодичности для неоднородного произведения неотрицательных матриц. Введем некоторые обозначения: x – вектор_столбец, тогда x^T – вектор_строка.

Определение. Функция $\rho(x, y)$, определенная на множестве $(1 \times n)$ положительных векторов, имеет свойства метрики или расстояния, если для любых двух векторов $x^T = (x_1, x_2, \dots, x_n) > 0$ и $y^T = (y_1, y_2, \dots, y_n) > 0$ она определена как

$$\rho(x^T, y^T) = \ln \left[\frac{\max_i (x_i / y_i)}{\min_i (x_i / y_i)} \right] = \max_{i,j} \ln \left(\frac{x_i y_j}{x_j y_i} \right)$$

и $\rho(x^T, y^T) = 0$ тогда и только тогда, когда $x = \lambda y$ ($\lambda > 0$).

Свойства функции $\rho(x, y)$:

1. $\rho(x^T, y^T) \geq 0$;
2. $\rho(x^T, y^T) = \rho(y^T, x^T)$;
3. $\rho(x^T, y^T) \leq \rho(x^T, z^T) + \rho(z^T, y^T)$;
4. $\rho(x^T, y^T) = \rho(\alpha x^T, \beta y^T)$,

где x^T, y^T, z^T любые положительные вектора $(1 \times n)$, а α, β – два положительных скаляра.

Данная псевдо-метрика показывает «проекционное расстояние» между векторами $x^T > 0$ и $y^T > 0$. Будем предполагать, что все векторы имеют длину n , а все матрицы являются не отрицательными и имеют размерность $(n \times n)$, если не сказано иное.

Введем некоторые понятия. Будем говорить, что матрица A не имеет нулевых столбцов, если в каждом столбце данной матрицы есть, по крайней мере, один элемент отличный от нуля. Аналогично, матрица A не имеет нулевых строк (без нулевых строк), если в каждой строке данной матрицы есть, по крайней мере, один элемент отличный от нуля.

Пусть A матрица без нулевых столбцов, тогда если $x^T, y^T > 0$, то $x^T A, y^T A > 0$.

Неравенство (2) дает свойство сжатия функции $\rho(\cdot, \cdot)$:

$$\rho(x^T A, y^T A) \leq \rho(x^T, y^T). \quad (2)$$

Докажем данное свойство.

Лемма 1. Если $x, y > 0$ и $\tilde{x} = Ax$, $\tilde{y} = Ay$, где $A = \{a_{ij}\} \geq 0$ и не имеет нулевых строк, то

$$\max_i \left(\frac{\tilde{x}_i}{\tilde{y}_i} \right) \leq \max_i \left(\frac{x_i}{y_i} \right), \quad \min_i \left(\frac{\tilde{x}_i}{\tilde{y}_i} \right) \geq \min_i \left(\frac{x_i}{y_i} \right). \quad (3)$$

Доказательство.

$$\frac{\tilde{x}_i}{\tilde{y}_i} = \frac{\sum_j a_{ij} x_j}{\sum_k a_{ik} y_k} = \sum_j p_{ij} \frac{x_j}{y_j},$$

где $p_{ij} = a_{ij}y_j / \sum_k a_{ik}y_k$ – (i, j) элемент стохастической матрицы P .

В частности, $\sum_j p_{ij} = 1$, тогда неравенства (3) доказано. Учитывая определение для функции $\rho(\cdot; \cdot)$ и результаты леммы 1 нетрудно показать, что неравенство (2) выполняется для матриц, которые не имеют нулевых столбцов.

Определение. Матрица $A = \{a_{ij}\} \geq 0$ без нулевых строк называется кодирующей, если любые две строки данной матрицы имеют, по крайней мере, один положительный элемент с одинаковым индексом j_0 .

Лемма 2. Если A^T кодирующая матрица и $x, y > 0$, то выполняется следующее неравенство

$$\rho(x^T A, y^T A) < \rho(x^T, y^T). \quad (4)$$

Следствие. Неравенство (4) будет также выполняться, если матрица A имеет хотя бы одну строку, каждый элемент которой отличен от нуля.

Учитывая неравенство (2) и свойство функции $\rho(\cdot; \cdot)$ ($\rho(x^T, y^T) = 0$ тогда и только тогда, если $x = \lambda y$ ($\lambda > 0$)), определим величину $\tau_B(A)$ следующим образом:

$$\tau_B(A) = \sup_{\substack{x, y > 0 \\ x \neq \lambda y}} \frac{\rho(x^T A, y^T A)}{\rho(x^T, y^T)},$$

$$0 \leq \tau_B(A) \leq 1. \quad (5)$$

Ясно, если A_1 и A_2 матрицы без нулевых столбцов, то для произведения данных матриц $A_1 A_2$ и для $x, y > 0$ имеем, что

$$\rho(x^T A_1 A_2, y^T A_1 A_2) \leq \tau_B(A_2) \rho(x^T A_1, y^T A_1) \leq \tau_B(A_2) \tau_B(A_1) \rho(x^T, y^T)$$

Тогда

$$\tau_B(A_1 A_2) \leq \tau_B(A_1) \tau_B(A_2). \quad (6)$$

Определение. Величины $\tau_B(\cdot)$ называются коэффициентами сжатия Биркгоффа (или коэффициентами эргодичности), если они удовлетворяют свойствам (5) и (6). [6]

Если из последовательности матриц $\{H_k\}$, где каждая матрица не имеет нулевых столбцов, выбрать матрицы $H_{p+1}, \dots, H_{p+r}$ и составить из них произведение $H_{p,r}$ (то есть $H_{p,r} = \prod_{k=p+1}^{p+r} H_k$) в произвольном порядке, то в силу неравенства (6) имеем

$$\tau_B(H_{p,r}) \leq \prod_{k=p+1}^{p+r} \tau_B(H_k). \quad (7)$$

Матрица A будет сжимающейся, если $\tau_B(A) < 1$.

Определим вид матрицы A , для которой $\tau_B(A) = 0$. Заметим, что если A имеет ранг 1, то есть A представимо в виде $A = wv^T = \{w_i v_j\}$, где $v > 0$; $w \geq 0, \neq 0$, а также A не имеет нулевых столбцов, тогда

$$\tau_B(A) = \sup_{\substack{x, y > 0 \\ x \neq \lambda y}} \frac{\rho(x^T w v^T, y^T w v^T)}{\rho(x^T, y^T)} = 0,$$

поскольку $\rho((x^T w)v^T, (y^T w)v^T) = \rho(v^T, v^T) = 0$. Таким образом, мы получаем, что для матрицы A , которая не имеет нулевых строк и столбцов, $\tau_B(A) = 0$ тогда и только тогда, когда A имеет ранг 1, то есть $A = wv^T$, $w, v > 0$.

Далее определим явную форму коэффициентов $\tau_B(A)$ для матрицы A , не содержащую нулевых строк и нулевых столбцов [6].

$$\tau_B(A) = \{1 - [\phi(A)]^{1/2}\} / \{1 + [\phi(A)]^{1/2}\},$$

где

$$\phi(A) = \begin{cases} \min_{i,j,k,l} \frac{a_{ik}a_{jl}}{a_{jk}a_{il}}, & \text{якщо } A > 0 \\ 0, & \text{якщо } A \leq 0 \end{cases},$$

Определение. Произведение $H_{p,r} = \{h_{ij}^{(p,r)}\}$, сформированное из матриц $H_{p+1}, H_{p+2}, \dots, H_{p+r}$, каждая из которых без нулевых строк и столбцов, умноженные в некотором особом порядке для каждого $p \geq 0, r \geq 1$, обладает свойством слабой эргодичностью, если существуют положительные матрицы $S_{p,r} = \{s_{ij}^{(p,r)}\}$ ($p \geq 0, r \geq 1$) ранга 1, такие, что для любого фиксированного p при $r \rightarrow \infty$ имеем:

$$h_{ij}^{(p,r)} / s_{ij}^{(p,r)} \rightarrow 1 \text{ для всех } i, j \quad (8)$$

Лемма 3. Произведение матриц $H_{p,r}$ является слабой эргодичностью тогда и только тогда, когда для всех $p \geq 0$ при $r \rightarrow \infty$ выполняется

$$\tau_B(H_{p,r}) \rightarrow 0. \quad (9)$$

Доказательство.

(достаточность) Из явной формы $\tau_B(A)$, которая связана с $A > 0$, ясно, что (8) очевидно включает в себя (9).

(необходимость) Пусть условие (9) выполняется и матрица $H_{p,r}$ без нулевых строк и столбцов. Определим матрицы ранга 1 в следующем виде:

$$H_{p,r} E E^T H_{p,r} / E^T H_{p,r} E.$$

Тогда из явной формы $\tau_B(\cdot)$ следует, что $H_{p,r} > 0$ для достаточного большого r . Отношение

$$\frac{h_{ij}^{(p,r)}}{s_{ij}^{(p,r)}} = h_{ij}^{(p,r)} / \left\{ \sum_{k,s} \frac{h_{ik}^{(p,r)} h_{sj}^{(p,r)}}{h_{sk}^{(p,r)} h_{ij}^{(p,r)}} \cdot \frac{h_{sk}^{(p,r)} h_{ij}^{(p,r)}}{\sum_{k,s} h_{ks}^{(p,r)}} \right\} \rightarrow 1,$$

поскольку выполняется условие (9), следовательно $\phi(H_{p,r}) \rightarrow 1$.

Далее обратим внимание на два вида произведений $T_{p,r} = H_{p+1}H_{p+2}...H_{p+r}$ и $U_{p,r} = H_{p+r} \cdots H_{p+2}H_{p+1}$, $r \rightarrow \infty$.

Лемма 4. Если $H_{p,r} = H_{p+1}H_{p+2}...H_{p+r}$, то есть $H_{p,r} = T_{p,r}$, и все H_k не имеют нулевых строк и столбцов, то $\tau_B(T_{p,r}) \rightarrow 0$ при $r \rightarrow \infty$ для всякого $p \geq 0$ тогда и только тогда, когда выполняются два условия:

а) $T_{p,r} > 0, r \geq r_o(p)$;

$$\text{б) } t_{ik}^{(p,r)} / t_{jk}^{(p,r)} \rightarrow W_{ij}^{(p)} > 0 \quad (10)$$

для всех i, j, p, k , где предел не зависит от k (то есть строки $T_{p,r}$ стремятся к пропорциональности при $r \rightarrow \infty$).

Доказательство. (достаточность) Очевидно, что из (10) следует (9), поскольку в силу (10) ясно, что $\Phi(T_{p,r}) \rightarrow 1$.

(необходимость) Пусть (9) выполняется, тогда ясно $T_{p,r} > 0$ для достаточно больших r ($r \geq r_o(p)$). Зафиксируем i и j , тогда

$$\frac{t_{ik}^{(p,r+1)}}{t_{jk}^{(p,r+1)}} = \sum_s \frac{d_{ks}^{(p,r)} t_{is}^{(p,r)}}{t_{js}^{(p,r)}},$$

где $d_{ks}^{(p,r)} = t_{js}^{(p,r)} h_{sk}(p+r+1) / t_{jk}^{(p,r+1)} - (k, s)$ элемент стохастической матрицы со строго положительными элементами, и, следовательно, данная матрица является кодирующей. Тогда из леммы 1 имеем

$$\max_k \left(\frac{t_{ik}^{(p,r)}}{t_{jk}^{(p,r)}} \right) \text{ не возрастает с ростом } r;$$

$$\min_k \left(\frac{t_{ik}^{(p,r)}}{t_{jk}^{(p,r)}} \right) \text{ не убывает с ростом } r.$$

Поскольку $\tau_B(T_{p,r}) \rightarrow 0$, $\Phi(T_{p,r}) \rightarrow 1$, то при $r \rightarrow \infty$ имеем

$$\frac{t_{ik}^{(p,r)}}{t_{jk}^{(p,r)}} \cdot \frac{t_{js}^{(p,r)}}{t_{is}^{(p,r)}} \rightarrow 1 \text{ для всех } i, j, k, s,$$

значит две монотонные величины при $r \rightarrow \infty$ имеют один и тот же положительный предел, который не зависит от k и может быть записан как $W_{ij}^{(p)}$.

Замечание 1. В лемме 4, $W_{i,j}^{(p)} = \lim_{r \rightarrow \infty} t_{ik}^{(p,r)} / t_{jk}^{(p,r)}$ может быть записана как $w_i^{(p)} / w_j^{(p)}$ для некоторого $w^{(p)} = \{w_i^{(p)}\}$, где $w^{(p)} > 0$, $(w^{(p)})^T E = 1$.

Теорема 1. Для последовательности $\{H_k\}$, $k = 1, 2, \dots$ неотрицательных и не имеющих нулевых строк и столбцов матриц ($H_{p,r} = T_{p,r}$ и $H_{p,r} = U_{p,r}$), слабая эргодичность существует тогда и только тогда, когда существует строго возрастающая последовательность положительных целых чисел $\{k_s\}$, $s = 0, 1, 2, \dots$ такая, что

$$\sum_{s=0}^{\infty} [\phi(H_{k_s, k_{s+1}-k_s})]^{1/2} = \infty. \quad (11)$$

Доказательство. Предположим, что $H_{p,r} = T_{p,r}$, $p \geq 0$, $r \geq 1$. Возьмем для простоты $p = 0$ и достаточно большое r .

$$T_{0,r} = T_{0,k_0} T_{k_0, k_1 - k_0} T_{k_1, k_2 - k_1} \cdots T_{k_{t-1}, k_t - k_{t-1}} T^*,$$

где $T^* \geq 0$ некоторая матрица без нулевых строк и столбцов, k_t – ближайший член последовательности $\{k_s\}$ не больший чем r . Ясно, что $t \rightarrow \infty$ при $r \rightarrow \infty$. Тогда из неравенства (5) следует, что

$$\tau_B(T_{0,r}) \leq \prod_{s=0}^{t-1} \tau_B(T_{k_s, k_{s+1}-k_s}),$$

и при $r \rightarrow \infty$ правая часть вышеприведенного неравенства $\rightarrow 0$ тогда и только тогда, когда

$$\sum_{s=0}^{\infty} \{1 - \tau_B(T_{k_s, k_{s+1}-k_s})\} = \infty.$$

Из определения для функции $\phi(\cdot)$, $0 \leq \phi(\cdot) \leq 1$, и явной формы $\tau_B(\cdot)$, следует, что дивергенция суммы вытекает из (11). Таким образом, условия (11) достаточно для слабой эргодичности для $T_{p,r}$.

Теперь пусть выполняется слабая эргодичность, тогда учитывая лемму 3 имеем, что $\tau_B(T_{p,r}) \rightarrow 0$ при $r \rightarrow \infty$, $p \geq 0$.

Зафиксируем $\delta: 1 > \delta > 0$. Тогда, по определению, последовательность $\{k_s\}$ рекуррентна по выбору произвольного k_0 , и k_{s+1} были определены как

$$\tau_B(T_{k_s, k_{s+1} - k_s}) \leq \delta.$$

То

$$[\phi(T_{k_s, k_{s+1} - k_s})]^{1/2} = \left\{ \frac{1 - \tau_B(T_{k_s, k_{s+1} - k_s})}{1 + \tau_B(T_{k_s, k_{s+1} - k_s})} \right\} \geq \frac{1 - \delta}{1 + \delta} > 0,$$

что доказывает выполнения условия (11) для данной последовательности

$$H_{p,r} = T_{p,r}.$$

Доказательство для произведения $U_{p,r}$ аналогично.

Следствие. Если для последовательности $\{k_s\}$, $s \geq 0$ положительных целых чисел таких, что $k_{s+1} - k_s = g(const)$, выполняется условие:

$$\phi(T_{k_s, k_{s+1} - k_s}) \geq \varepsilon^2,$$

тогда $\tau_B(T_{p,r}) \rightarrow 0$ при $r \rightarrow \infty$, $p \geq 0$, по геометрической норме. Например, в случае $p = 0$ имеем:

$$\tau_B(T_{0,p}) \leq \{(1 - \varepsilon)/(1 + \varepsilon)\}^{-(k_0/g)^{-1}} \{(1 - \varepsilon)/(1 + \varepsilon)\}^{r/g}$$

для достаточного большого r . Аналогичный результат можно получить и для произведения $U_{p,r}$.

Теорема 2. Если для последовательности неотрицательных и без нулевых строк и столбцов матриц $H_k = \{h_{ij}(k)\}$, $k \geq 1$, выполняются условия

i) $H_{p,r_0} > 0$ для $p \geq 0$, где r_0 (≥ 1) некоторое фиксированное целое число независимое от p

$$i) \min_{i,j}^+ h_{ij}(k) / \max_{i,j} h_{ij}(k) \geq v > 0 \quad (12)$$

(где $\min^+$ – минимум по всем положительным элементам и v не зависит от k), тогда если $H_{p,r} = T_{p,r}$ (или $U_{p,r}$), $p \geq 0$, $r \geq 1$, слабая эргодичность (по геометрической норме) существует.

Доказательство. Из структуры $\tau_B(\cdot)$ и его зависимости от $\phi(\cdot)$ очевидно, что значение $\tau_B(H_{p,r})$ не изменится, если каждую H_k умножить на некоторый положительный скаляр (каждый скаляр зависит от k). Поскольку из леммы 3 ясно, что слабая эргодичность зависит только от таких значений (величин), мы можем заменить условие (12) без потери общности на следующие неравенства:

$$0 < v \leq \min_{i,j}^+ h_{ij}(k), \quad \max_{i,j} h_{ij}(k) \leq 1. \quad (13)$$

Из того, что $H_{p,r_0} > 0$, $p \geq 0$, следует, что

$$v^{r_0} \Pi^T \leq H_{p,r_0} \leq n^{r_0-1} \Pi^T. \quad (14)$$

Значит $\phi(H_{p,r_0}) \geq (v^{r_0} / n^{r_0-1})^2 = \varepsilon^2$, $p \geq 0$. Далее применяем следствие теоремы 1 с $g = r_0$ и $k_0 = 1$.

Утверждение 1. Если все матрицы H_k , $k > 1$ имеют одну и ту же матрицу событий и являются неразложимыми, а также хотя бы один элемент главной диагонали каждой матрицы отличен от нуля, то $H_{p,r} > 0$, $p \geq 0$, $r \geq 2(n-1)$.

Если произведение $H_{p,r} = \{h_{ij}^{(p,r)}\}$ обладает свойством слабой эргодичности, то выполняется условие

$$\lim_{r \rightarrow \infty} \frac{h_{ik}^{(p,r)}}{h_{jk}^{(p,r)}} = W_{ij}^{(p)} = \frac{w_i^{(p)}}{w_j^{(p)}} \text{ для всех } i, j, p, k,$$

где $w^{(p)} = \{w_i^{(p)}\}$ некоторый стохастический вектор.

Свойство эргодичности для неоднородной модели Лесли. Рассмотрим неоднородную модель Лесли вида (1). Основываясь на утверждении 1 теоремы 2 и на том, что все элементы матрицы Лесли находятся в диапа-

зоне от нуля до бесконечности, можно утверждать, что условия теоремы 2 выполняются для $L_{o,n}$.

Тогда, определим два различных начальных распределения как $X(t_0^1) = \{x_i^1\}$ и $X(t_0^2) = \{x_i^2\}$. Рассмотрим отношение $L_{o,n} X(t_0^1)$ к $L_{o,n} X(t_0^2)$, при этом числитель и знаменатель данной дроби умножим на $l_{qk}^{(0)}$:

$$\frac{L_{o,n} X(t_0^1)}{L_{o,n} X(t_0^2)} = \frac{\sum_{i=1}^n x_i^{(1)} l_{ik}^{(0,n)}}{\sum_{j=1}^n x_j^{(2)} l_{jk}^{(0,n)}} = \frac{\sum_{i=1}^n x_i^{(1)} l_{ik}^{(0,n)} \cdot l_{qk}^{(0,n)}}{\sum_{j=1}^n x_j^{(2)} l_{jk}^{(0,n)} \cdot l_{qk}^{(0,n)}}.$$

Применяя лемму 4 к последнему выражению и устремляя $n \rightarrow \infty$ имеем:

$$\frac{\sum_{i=1}^n x_i^{(1)} l_{ik}^{(0,n)}}{\sum_{j=1}^n x_j^{(2)} l_{jk}^{(0,n)}} \rightarrow \frac{\sum_{i=1}^n x_i^{(1)} w_{iq}^{(0)}}{\sum_{j=1}^n x_j^{(2)} w_{jq}^{(0)}} = \frac{\sum_{i=1}^n x_i^{(1)} w_i^{(0)}}{\sum_{j=1}^n x_j^{(2)} w_j^{(0)}},$$

то есть значение предела не зависит от k . Данная независимость называется свойством слабой эргодичности неоднородной модели Лесли и для популяции интерпретируется следующим образом: доля особей каждого возрастного класса по отношению к общей численности популяции остается неизменной, хотя общая численность может изменяться с течением времени. То есть, при $n \rightarrow \infty$ имеем

$$\begin{aligned} & \frac{\sum_{i=1}^n x_i(t_o) l_{ik}^{(0,n)}}{\sum_{i=1}^n \sum_{s=1}^n x_i(t_o) l_{is}^{(0,n)}} \sim \\ & \sim \left\{ l_{qk}^{(0,p)} \sum_{i=1}^n \frac{x_i(t_o) w_i^{(0)}}{w_q^{(0)}} \right\} / \left\{ \left(\sum_{s=1}^n l_{qs}^{(0,n)} \right) \left(\sum_{i=1}^n \frac{x_i(t_o) w_i^{(0)}}{w_q^{(0)}} \right) \right\} \sim \\ & \sim l_{qk}^{(0,n)} / \sum_{s=1}^n l_{qs}^{(0,n)} \end{aligned}$$

для любого фиксированного $q = 1, \dots, n$ и всех $k = 1, \dots, n$. Видим, что последний предел не зависит от $X(t_0)$.

Выводы. Введение неоднородности в модель Лесли обусловлена достоверностью описания динамики реальных популяций. В связи с этим встает вопрос об исследовании свойств данной модели. В данной статье установлено, что неоднородная модель Лесли обладает свойством слабой эргодичности.

Бibliографические ссылки

1. **Leslie P.H.** On the use of matrices in certain population mathematics // Biometrika. – 1945. – V.33, N3. – P.183-212.
2. **Свирижев Ю.М.** Устойчивость биологических сообществ / Ю.М. Свирижев, Д.О. Логофет. – М., 1978.
3. **Ризниченко Г.Ю.** Математические модели биологических продукционных процессов / Г.Ю. Ризниченко, А.Б. Рубин. – М., 1993.
4. **Логофет Д.О.** Неотрицательные матрицы как инструмент моделирования динамики популяций: классические модели и современные обобщения / Д.О. Логофет, И.Н. Белова // Фундаментальная и прикладная математика, Т. 13. – вып. 4. – 2007. – С.145-164.
5. **Horn R.** Matrix Analysis / R. A. Horn, C. R. Johnson. — London: Cambridge Univ. Press, 1990
6. **Seneta E.** Non-negative matrices and Markov chains. - Springer-Verlag New-York-1981. - 284p.

Надійшла до редколегії 21.02.11

В.Н. Богомаз

Днепропетровский национальный университет имени Олеся Гончара

О РАЗРЕШИМОСТИ ОДНОЙ ЗАДАЧИ ВЕКТОРНОЙ ОПТИМИЗАЦИИ ДИНАМИКИ ВИБРОСИСТЕМЫ

Досліджується питання розв'язності задачі векторної оптимізації динаміки вібросистеми. Одержані достатні умови існування розв'язків.

Исследуется вопрос разрешимости задачи векторной оптимизации динамики вибросистемы. Получены достаточные условия существования решений.

The question of solvability of task to vector optimization of dynamics of the vibrosystem is in-process investigated. The sufficient terms of existence of decisions are got.

Ключевые слова: полунепрерывность снизу, эффективные решения, частично упорядоченные пространства, динамическая система.

Введение. В предыдущей работе автора [3] была рассмотрена задача оптимального управления вибросистемой, встроенной в уплотняющую машину каткового типа. Конструкция вибросистемы подробно была описана в работе [1]. Особенности рассмотренной задачи оптимального управления являются:

- наличие ограничений на управления и фазовые координаты;
- система обыкновенных дифференциальных уравнений существенно нелинейна относительно фазовых координат и линейна относительно управлений;
- функционал качества линейно зависит от управлений.

Задача имеет следующий физический смысл: среди допустимых законов изменения моментов, приложенных к валам дебалансов и водила, найти такие, которые обеспечивают увеличение воздействия машины на уплотняемую среду в течение всего промежутка времени работы вибросистемы, сохраняя угловые скорости вращения дебалансов и водила в преде-

лах допустимых. Другими словами, необходимо обеспечить отрицательность суммарной вертикальной возмущающей силы в течение всего промежутка времени работы вибросистемы.

Функционалом качества выбран интеграл от функции изменения суммарной возмущающей силы. Таким образом, минимизировалось среднее значение возмущающей силы на промежутке времени от 0 до 40 секунд.

Сглаженная функция изменения суммарной возмущающей силы для вибросистемы с двумя дебалансами, соответствующая оптимальному управлению в рассмотренной задаче приведена на рисунке 1.

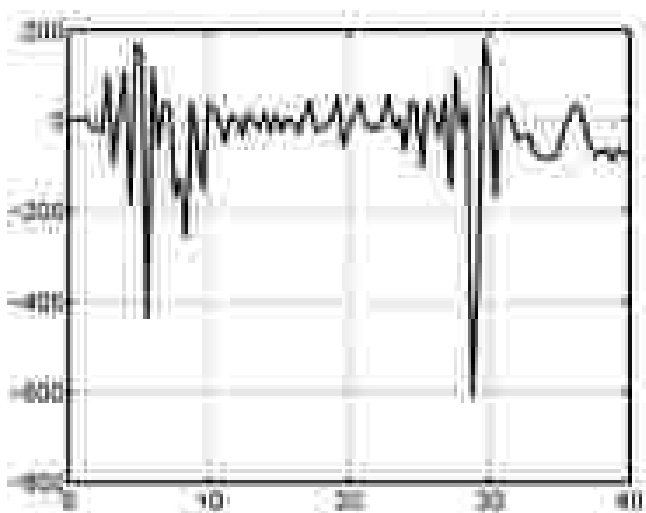


Рис. 1 Суммарная возмущающая сила $L(t, x(t))$

Рассматривая кривую изменения суммарной возмущающей силы, приходим к выводу, что среднее ее значение отрицательно, но есть отрезки времени, где она строго положительна, что является нежелательным. Таким образом, рассматривать подобную задачу с точки зрения скалярной оптимизации является неэффективным.

Основной целью данной работы является постановка задачи векторной оптимизации динамики вибросистемы, а также приведение достаточных условий ее разрешимости.

Основные понятия и вспомогательные результаты. Приведем некоторые известные понятия, необходимые для дальнейших рассуждений.

Определение 1.1. Непустое множество Λ банахового пространства Y с нулевым элементом θ называется конусом, если выполняется условие $\alpha\Lambda \subseteq \Lambda \forall \alpha \geq 0$. Если к тому же выполняется:

- 1) условие $\Lambda + \Lambda \subseteq \Lambda$, то конус $\Lambda \subseteq Y$ называется выпуклым;
- 2) условие $\Lambda \cap (-\Lambda) = \{\theta\}$, то конус $\Lambda \subseteq Y$ называется острым.

Конус $\Lambda \subseteq Y$ называется телесным, если его внутренность $Int\Lambda$ непуста.

Всюду далее будем считать, что пространство Y частично упорядочено выпуклым замкнутым острым конусом $\Lambda \subset Y$. Обозначим частичный порядок, порожденный конусом Λ как $\leq_\Lambda$, что означает следующее:

$$y_1 \leq_\Lambda y_2 \Leftrightarrow y_2 - y_1 \in \Lambda, \quad y_1 <_\Lambda y_2 \Leftrightarrow y_2 - y_1 \in \Lambda \setminus \{\theta\}. \quad (1.1)$$

Пусть Y_0 – некоторое непустое подмножество пространства Y . Обозначим, через $IntY_0$, clY_0 , $\partial_\tau Y_0$ соответственно внутренность, замыкание и границу относительно топологии τ множества $Y_0 \subset Y$. Всюду далее, если τ ассоциируется с сильной топологией и это не приводит к непониманию, этот индекс будем опускать. Когда τ ассоциируется со слабой топологией, то в качестве индекса будем использовать букву w .

Приведем некоторые известные факты из теории векторной оптимизации.

Элемент $y^* \in Y_0$ будем называть Λ -минимальным, если выполняется условие $(y^* - \Lambda) \cap Y_0 = \{y^*\}$. Совокупность всех Λ -минимальных элементов множества $Y_0 \subset Y$ обозначим через $\Lambda - Min(Y_0)$.

Элемент $y^* \in Y_0$ будем называть Λ -минорантой множества $Y_0 \subset Y$, если выполняется условие $y^* \leq_\Lambda y$ для всех $y \in Y_0$. Множество Y_0 , для которого множество его Λ -минорант непусто, будем называть Λ -ограниченным снизу.

Последовательность $\{y_n \in Y_0\}$ будем называть невозрастающей, если для всех $n \in \mathbb{N}$ выполняется условие $y_{n+1} \leq_\Lambda y_n$. Выпуклый конус Λ

будем называть правильным, если каждая невозрастающая и Λ_- -ограниченная снизу последовательность $\{y_n \in Y\}$ y^* сильно в Y .

Определение 1.2. Множество $Y_0 \subset Y$ называется Λ_- -полуограниченным снизу, если любая невозрастающая последовательность $\{y_n\} \subset Y_0$ является Λ_- -ограниченной снизу.

Обозначим через $\bar{Y}$ расширенное пространство $Y \cup \{\pm\infty\}$, а через $Y^\bullet$ - полурасширенное пространство $Y \cup \{\infty\}$. При этом допустим, что $\|\pm\infty\|_Y = \pm\infty$.

Приведем определение эффективного Λ_τ -инфимума множества $Y_0 \subset Y$.

Определение 1.3. Эффективным Λ_τ -инфимумом множества $Y_0 \subset Y$ будем называть множество Λ_- -минимальных элементов τ -замыкания множества Y_0 в пространстве Y в случае, когда это множество непусто, и множество $\{-\infty\}$ в противоположном случае, т.е.:

$$\Lambda_\tau \text{ Inf } (Y_0) = \begin{cases} \Lambda_- \text{ Min}(Y_{\tau 0}), \text{ Min}(Y) \neq \emptyset \\ \{-\infty\}, \text{ Min}(Y_{\tau 0}) = \emptyset \end{cases} \quad (1.2)$$

Пусть X и Y – банаховы пространства, и Y частично упорядочено с помощью некоторого выпуклого замкнутого острого конуса Λ . Для любой последовательности $\{y_n\} \subset Y$ обозначим через $L_Y^\tau\{y_n\}$ множество всех ее предельных точек относительно топологии τ . Будем считать, что:

- 1) $-\infty \in L_Y^\tau\{y_n\}$, если последовательность $\{y_n\} \subset Y$ является убывающей и неограниченной снизу;
- 2) $+\infty \in L_Y^\tau\{y_n\}$, если последовательность $\{y_n\} \subset Y$ является возрастающей и неограниченной сверху.

Пусть задано некоторое отображение $f: X \rightarrow Y^\bullet$. Множество, являющееся эффективным Λ_τ -инфимумом образа $f(X)$ в $Y^\bullet$, будем называть

эффективным Λ_{τ} -инфимумом отображения f и обозначать $\Lambda_{\tau} \text{ Inf}_{x \in X} f(x)$.

Для произвольного отображения $f : X \rightarrow Y^{\bullet}$ введем для рассмотрения следующие множества:

$$L^{s, \tau \tau}(f(x_0)) = \bigcap_{\{x_{ks}\}_{k=1}^{\infty} \in S(x_0)} L(f(x_k)), \quad (1.3)$$

$$L^{w, \tau \tau}(f(x_0)) = \bigcap_{\{x_{kw}\}_{k=1}^{\infty} \in S(x_0)} L(f(x_k)), \quad (1.4)$$

где $S_x(x_0)$ и $S_w(x_0)$ – совокупность всех последовательностей $\{x_k\}_{k=1}^{\infty} \subset X$ таких, что $x_k \rightarrow_0$ сильно в X и слабо в X соответственно.

Для приведения понятия Λ_{τ} -нижней грани отображения введем следующие множества:

$$L_{\min}^{s, \tau \tau}(f(x_0)) = L(f(x)) \cap \Lambda_{\tau} \text{ Inf}_{x \in X} f(x), \quad (1.5)$$

$$L_{\min}^{w, \tau \tau}(f(x_0)) = L(f(x)) \cap \Lambda_{\tau} \text{ Inf}_{x \in X} f(x). \quad (1.6)$$

Определение 1.4. Λ_{τ} -нижней гранью отображения f в точке $x_0 \in X$ будем называть множество:

$$\Lambda_{\tau} \liminf_{x \rightarrow_0} f(x) = \begin{cases} L_{\min}^{s, \tau \tau}(f(x_0)), & L_{\min}(f(x)) \neq \emptyset \\ \Lambda_{\tau} \text{ Inf}_{x \in X} L^{s, \tau \tau}(f(x_0)), & L_{\min}(f(x)) = \emptyset \end{cases} \quad (1.7)$$

и слабой Λ_{τ} -нижней гранью отображения f в точке $x_0 \in X$ – множество:

$$\Lambda_{\tau} \liminf_w f(x) = \begin{cases} L_{\min}^{w, \tau \tau}(f(x_0)), & L_{\min}(f(x)) \neq \emptyset \\ \Lambda_{\tau} \text{ Inf}_{x \in X} L^{w, \tau \tau}(f(x_0)), & L_{\min}(f(x)) = \emptyset \end{cases} \quad (1.8)$$

где $x \xrightarrow{w} 0$ означает, что $x \rightarrow_0$ слабо в X .

Пусть $Domf = \{x^{**} \in X : f(x) \neq +\infty\}$ – эффективное множество отображения $f : X \rightarrow Y^*$.

Определение 1.5. Отображение $f : X \rightarrow Y^*$ будем называть Λ_τ -полунепрерывным снизу в точке $x_0 \in Domf$, если $f(x_0) \in \Lambda_\tau \liminf_{x \rightarrow x_0} f(x)$ и слабо Λ_τ -полунепрерывным снизу в точке $x_0 \in Domf$, если $f(x_0) \in \Lambda_\tau \liminf_{x \xrightarrow{w} x_0} f(x)$.

Пусть X_∂ некоторое непустое подмножество пространства X .

Как известно, для задач векторной оптимизации основным понятием является понятие эффективного решения.

Определение 1.6. Элемент $x^* \in X_\partial$ назовем Λ -эффективным решением задачи векторной оптимизации, если $f(x^*)$ является Λ -минимальным элементом множества $f(X_\partial)$, т.е. выполняется условие:

$$(f(x^{**}) - \Lambda) \cap f(X_\partial) = \{f(x)\}.$$

Задачу векторной оптимизации, при которой необходимо найти множество всех Λ -эффективных решений, будем сокращенно записывать:

$$f(x) \rightarrow \Lambda - Min, \quad x \in X_\partial. \quad (1.9)$$

Множество всех Λ -эффективных решений задачи векторной оптимизации (1.9) обозначим через $Eff(X_\partial; f, \Lambda)$, т.е.:

$$Eff(X_\partial; f, \Lambda) = \left\{ x^{**} \in X : f(x) \in \Lambda - \inf_{x \in X_\partial} f(x) \right\}.$$

Определение 1.7. Элемент $x^* \in X_\partial$ назовем Λ_τ -эффективным решением задачи векторной оптимизации, если $f(x^*)$ является Λ -минимальным элементом множества $cl_\tau f(X_\partial)$. Множество всех Λ_τ -

эффективных решений задачи векторной оптимизации обозначим через $Eff_{\tau}(X_{\partial}; f; \Lambda)$, т.е.:

$$Eff_{\tau}(X_{\partial}; f; \Lambda) = \left\{ x^{**} \in X : f(x) \in \Lambda - \inf_{x \in X_{\partial}} f(x) \right\}.$$

Задачу векторной оптимизации, при которой необходимо найти множество всех Λ_{τ} -эффективных решений, будем сокращенно записывать:

$$f(X_{\partial}) \rightarrow \Lambda_{\tau} \text{ inf}, \quad x \in X_{\partial}. \quad (1.10)$$

Приведем один известный результат из [5].

Теорема 1. Пусть X и Y – банаховы пространства, пространство Y полуупорядочено некоторым конусом Λ , множество $X \subset Y$ является

μ -компактным, а отображение $f : X_{\partial} \rightarrow Y^*$ является Λ_{τ} -полунепрерывным снизу относительно топологии μ пространства X . Тогда отображение f является Λ -полуограниченным снизу. Если к тому же, конус Λ – правильный и для любого $x \in X_{\partial}$ выполнено условие $\{f(x)\} = \Lambda_{\tau} \liminf_{u \rightarrow x} f(u)$, то $Eff_{\tau}(X_{\partial}; f; \Lambda) \neq \emptyset$.

Замечание 1. Внутренность конуса Λ может быть пустой.

Необходимо заметить, что для векторнозначных отображений существует несколько понятий полунепрерывности. Для формулировки следствий из теоремы 1, приведем некоторые из них.

Определение 1.8. Отображение $f : X_{\partial} \rightarrow Y^*$ называют полунепрерывным снизу в точке $x_0 \in X_{\partial}$, если для любой окрестности V точки $f(x_0)$ в Y существует окрестность точки x_0 в X такая, что выполняется условие $f(U) \subset V + \Lambda \cup \{\emptyset\}$.

Определение 1.9. Отображение $f : X_{\partial} \rightarrow Y^*$ называют квазиполунепрерывным снизу в точке $x_0 \in X_{\partial}$, если для любого $b \notin f(x_0)$ существует

окрестность U точки $x_0 \in X$ такая, что выполняется условие $b \neq f(x_0)$ $\forall x \in U$.

Определение 1.10. Отображение $f : X_\partial \rightarrow Y^*$ называют o -полу непрерывным снизу в точке $x_0 \in X$, если для любой последовательности $\{x_n\} \subset X$, которая удовлетворяет следующим условиям:

- 1) $x_n \rightarrow x_0$ в X ;
- 2) существует последовательность $\{\varepsilon_n\} \subset Y$, $\{\varepsilon_n\} \rightarrow 0$ такая, что последовательность $\{f(x_n) + \varepsilon_n\}$ является невозрастающей;
- 3) существует последовательность $\{g_n\} \subset Y$, $\{g_n\} \rightarrow 0$ такая, что $f(x_0) \leq f(x_n) + g_n$ для всех $n \in \mathbb{N}$.

Отображение $f : X_\partial \rightarrow Y^*$ будем называть полу непрерывным снизу (соответственно квазиполу непрерывным снизу, o -полу непрерывным снизу), если оно является полу непрерывным снизу (соответственно квазиполу непрерывным снизу, o -полу непрерывным снизу) в каждой точке $x_0 \in X_\partial$.

Приведем следствия из теоремы 1.

Следствие 1.1. Пусть X и Y - банаховы пространства, пространство Y полуупорядочено с помощью правильного конуса Λ , множество $X_\partial \subset X$ является компактным (относительно топологии, порожденной нормой), а отображение $f : X_\partial \rightarrow Y^*$ квазиполу непрерывное снизу (o -полу непрерывное снизу). Тогда $Eff_\tau(X_\partial; f, \Lambda) \neq \emptyset$.

Следствие 1.2. Пусть X - рефлексивное пространство, Y - банахово пространство, полуупорядоченное с помощью правильного конуса Λ , множество X_∂ ограничено и слабо замкнуто в X , отображение $f : X_\partial \rightarrow Y^*$ слабо квазиполу непрерывно снизу (o -полу непрерывное снизу). Тогда $Eff_\tau(X_\partial; f, \Lambda) \neq \emptyset$.

Постановка задачи и исследование ее разрешимости. Пусть $L^{pn}(t_0, T)$ - пространство управлений, а $W^{1,pn}(t_0, T)$ - простран-

ство фазовых траекторий, где $1 < p < \infty$. Пусть U_{∂} – некоторое подмножество допустимых управлений в пространстве $L^{pn}(t_0, T)$, а K – некоторое фиксированное множество в пространстве $W^{1,pn}(t_0, T)$.

Рассмотрим следующую задачу векторной оптимизации:

$$L(x, u) \rightarrow \inf, \quad (2.1)$$

$$\dot{x} = f(x, u), \quad (2.2)$$

$$u \in U, \quad (2.3)$$

$$x \in K, \quad (2.4)$$

$$x_{ii}(t_0) = x_i, \quad \forall i \in [1, n], \quad (2.5)$$

где $L: L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t_0, T; \mathbb{R}^n) \rightarrow \mathbb{R}^l$ – целевое отображение.

Всюду далее будем говорить, что задача (2.1) – (2.5) является регулярной, если существует хотя бы одна пара $(x, u) \in L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t_0, T; \mathbb{R}^n)$, где $x = x(t)$ соответствующее решение системы (2.2) такое, что пара (x, u) удовлетворяет условиям (2.3)–(2.5). В этом случае пару (x, u) будем называть допустимой в задаче (2.1)–(2.5).

Обозначим через Ξ множество всех допустимых пар в задаче (2.1) – (2.5):

$$\Xi = \left\{ (x, u) \in L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t_0, T; \mathbb{R}^n) \mid \begin{aligned} &\dot{x} = f(x, u), x(t_0) = x_0, u \in U, \\ &x \in K \end{aligned} \right\}. \quad (2.6)$$

Всюду далее задачу векторной оптимизации (2.1) – (2.5) будем ассоциировать с тройкой (Ξ, L) .

Введем для рассмотрения расширение отображения $L: \Xi \rightarrow L^p(t_0, T; \mathbb{R}^l)$ на все пространство $L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t_0, T; \mathbb{R}^n)$:

$$L(u, \cdot) = \begin{cases} L(u, \cdot), & (u, \cdot) \in \Xi, \\ +\infty, & (u, \cdot) \notin \Xi. \end{cases}$$

Будем говорить, что отображение $L^p(t_0, T; \cdot) \times W^{1,p}(t_0, T; \cdot) \rightarrow L^p(t_0, T)$ ограничено снизу, если существует хотя бы один элемент $y \in \Lambda$ такой, что $y \leq L(u, \cdot)$ для всех $(u, \cdot) \in L^p(t_0, T; \cdot) \times W^{1,p}(t_0, T; \cdot)$.

Определение 2.1. Будем говорить, что последовательность пар $\{(u_k, x)\}_{k=1}^\infty \subset L^p(t_0, T; \cdot) \times W^{1,p}(t_0, T; \cdot)$ является слабо сходящейся к паре $(u_0, x_0) \in L^p(t_0, T; \cdot) \times W^{1,p}(t_0, T; \cdot)$ (сокращенно $(u_k, x) \xrightarrow{w} (u_0, x)$), если $u_k \rightarrow_0$ слабо в $L^{pn}(t_0, T)$ и $x_k \rightarrow_0$ слабо в $W^{1,pn}(t_0, T)$ при $k \rightarrow \infty$.

Пару $(u^*, x^*) \in \Xi$ будем называть Λ_w -эффективным решением задачи (Ξ, L) , если $(L(u^*, x^*) - \Lambda) \cap cl_w L(\Xi) = \emptyset$ где $cl_w L(\Xi)$ - замыкание множества $L(\Xi)$ в слабой топологии пространства $L^p(t_0, T)$.

Таким образом, имеем:

$$L(u^*, x^*) \in \Lambda_w^- \left(\inf_{(u, x) \in \Xi} L(u, x) \right).$$

Для определения условий, которые гарантируют существование эффективных решений сделаем некоторые предположения. Пусть

- 1) U_0 – ограниченное слабо замкнутое множество в $L^{pn}(t_0, T)$;
- 2) система $\dot{x} = f(u, x)$ – система уравнений Каратеодори, правые части которых непрерывны по Липшицу относительно переменных x ;
- 3) K – секвенциально слабо замкнутое множество в $W^{1,pn}(t_0, T)$;

4) целевое отображение $L: \Xi \rightarrow L^p(t_0, T)$ является секвенциально Λ_τ -полу непрерывным снизу относительно слабой топологии в пространстве $L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t, T; \mathbb{R}^n)$;

5) задача (2.1)-(2.5) регулярна.

При таких предположениях для задачи (2.1) -(2.5) получен следующий результат.

Теорема 2. Пусть U_0 – ограниченное множество пространства $L^{pn}(t_0, T)$, K – секвенциально слабо замкнутое подмножество пространства $W^{1,pn}(t_0, T)$ для всех $1 < p < +\infty$, система $\dot{x} = f(u, x)$ относится к типу уравнений Каратеодори, правые части которых непрерывны по Липшицу относительно переменных x . Пусть целевое отображение $L: \Xi \rightarrow L^p(t_0, T)$ является секвенциально Λ_τ -полу непрерывным снизу относительно слабой топологии в пространстве $L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t, T; \mathbb{R}^n)$. Пусть пространство $L^p(t_0, T)$ упорядочено некоторым правильным конусом Λ . Пусть для любых пар $(u, x) \in \Xi$ выполняется условие $\{L(u, x)\} = \Lambda_\tau \liminf_{(u_{kk}, x) \xrightarrow{w} (u, x)} L(u_{kk}, x)$.

Тогда $Eff_\tau(\Xi; L, \Lambda) \neq \emptyset$, если и только если задача регулярна.

Доказательство. Очевидно, что если $Eff_\tau(\Xi; L, \Lambda) \neq \emptyset$, то задача (2.1)-(2.5) является регулярной.

Рассмотрим множество допустимых пар в задаче (2.1) – (2.5). Из теории дифференциальных уравнений известно, что решения уравнений Каратеодори при любом $u \in U_0$ $L^{pn}(t_0, T)$ образуют ограниченное множество в пространстве $W^{1,pn}(t_0, T)$ для всех $1 < p < +\infty$. Согласно теореме Банаха о слабой компактности ограниченного множества в рефлексивном банаховом пространстве и вследствие секвенциальной слабой замкнутости множества Ξ , имеем: Ξ – секвенциально слабо компактное

множество в пространстве $L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t, T; \mathbb{R}^n)$. Таким образом, из любой последовательности $\{(u_k, x_k)\}_{k=1}^\infty \subset \Xi$ можно извлечь слабо сходящуюся подпоследовательность, что означает: существует такая пара $(u^{**}, x^{**}) \in L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t, T; \mathbb{R}^n)$, что при $k \rightarrow \infty$ выполняется $(u_{k_k}, x_{k_k}) \xrightarrow{w} (u^{**}, x^{**})$ (с точностью до подпоследовательности). Т.к. при каждом k выполняется $(u_k, x_k) \in \Xi$, а множество $U_K \times X_K$ является секвенциально слабо замкнутым, то, очевидно, что $(u^{**}, x^{**}) \in U_\partial \times K$. По условию пара (u_k, x_k) удовлетворяет системе уравнений (2.2) при любом $k \in \mathbb{N}$. Используя теорему вложения Соболева, получим что пространства $W^{1,p}(t_0, T; \mathbb{R}^m)$ компактно вложены в $C([t_0, T; \mathbb{R}^m])$ для всех $1 < p < +\infty$. Переходя к пределу в правой и левой частях уравнений и используя свойство слабой сходимости последовательности $\{u_k\}_{k=1}^\infty$ в $L^p(t_0, T; \mathbb{R}^m)$, получим, что пара (u^{**}, x^{**}) удовлетворяет системе уравнений (2.2). Таким образом, $(u^{**}, x^{**}) \in \Xi$.

Учитывая, что целевое отображение $L: \Xi \rightarrow L^p(t_0, T)$ является секвенциально Λ_τ -полунепрерывным снизу относительно слабой топологии в пространстве $L^p(t_0, T; \mathbb{R}^m) \times W^{1,p}(t, T; \mathbb{R}^n)$ и применяя теорему 1, получим, что при условии $\{L(u_k, x_k)\} = \Lambda_\tau \liminf_{(u_{k_k}, x_{k_k}) \xrightarrow{w} (u, x)} L(u_{k_k}, x_{k_k}), \forall (u, x) \in \Xi$, выполняется $Eff_\tau(\Xi; L) \neq \emptyset$.

Теорема доказана.

Приведем следствия из теоремы 2.

Следствие 2.1. Пусть выполняются требования теоремы 2 относительно системы уравнений $\dot{x} = f(u, x)$. Пусть множества U_∂ и K являются

компактными относительно сильных топологий пространств соответственно $L^{pn}(t_0, T)$ и $W^{1,pn}(t_0, T)$, целевое отображение $L; \Xi \rightarrow L^p(t_0, T)$ квазиполунепрерывно снизу (o -полунепрерывно снизу). Тогда $Eff_\tau(\Xi; L) \neq \emptyset$.

Следствие 2.2. Пусть выполняются предположения теоремы 2 относительно системы уравнений $\dot{x} = f(u, x)$ и множеств U_∂, K . Пусть целевое отображение $L; \Xi \rightarrow L^p(t_0, T)$ слабо квазиполунепрерывно снизу (o -полунепрерывно снизу). Тогда $Eff_\tau(\Xi; L) \neq \emptyset$.

Применение полученных результатов. Покажем применение полученных выше результатов к оптимизации динамики вибросистемы. Объект управления представляет собой вибросистему некоторой конструкции, встроенную в уплотняющую машину каткового типа. Конструкция вибросистемы была подробно описана в [1].

Пусть m, λ, R, i, E_i - заданные массовые и геометрические характеристики вибросистемы, $k_i, C_{\dot{y}i}, I_{i3\theta} \forall i \in [1, 2n+2]$ - константы при заданных массовых и геометрических характеристиках вибросистемы; $M(x)$ - функция, равномерно непрерывная относительно x .

Математическая модель объекта управления имеет следующий вид:

$$\left\{ \begin{array}{l} \dot{x}_{2n+1} = \dots; \\ \dot{x}_{2n+1} = \dots; \\ \dot{x}_{2n+2} = \frac{1}{I_{10}} \left[y_1 + k_1 \cdot \sin x_1 - C_1 \sin(x_1 - \beta\beta) \cdot x_{2n+2}^2 - C_{f \ 2 \ 2}(u_x) \cos(x_1 - \dots) \right]; \\ \dot{x}_{2n+3} = \frac{1}{I_{20}} \left[y_2 + k_2 \cdot \sin x_2 - C_2 \sin(x_2 - \beta\beta) \cdot x_{2n+2}^2 - C_{f \ 2 \ 2}(u_x) \cos(x_2 - \dots) \right]; \\ \dots \\ \dot{x}_{2n+1} = \frac{1}{I_{n0}} \left[y_n + k_n \cdot \sin x_n - C_n \sin(x_n - \beta\beta) \cdot x_{2n+2}^2 - C_{f \ 2 \ 2}(u_x) \cos(x_n - \dots) \right]; \\ \dot{x}_{2n+1} = \frac{1}{M(x)} \left(\sum_{i=1}^n C_{x \ 22} \sin(x_i - \beta_i) - \sum_{i=1}^n \frac{C_{f \ 22}}{2I_{f \ 22}} x_{2n+2}^2 \sin(x_i - 2\beta_i) - \sum_{i=1}^n u_i \cos(x_i - \beta_i) + \right. \\ \left. + \sum_{i=1}^n \frac{C_i}{I_{i0}} k_i \sin(x_i) \cos(x_i - \beta_i) + u_{n+1} + \sum_{i=1}^n N_i \cdot \sin(\beta_i) + \sum_{i=1}^n P_i \cdot [k_i \sin(\beta_i)] \right); \end{array} \right.$$

(3.1)

$$u \in U_D = \left\{ u \in L^2(t_0, T; \mathbb{R}^{n+1}) \mid m_{\min} \leq u_i(t) \leq m_{\max}, \forall i \in [1, n+1] \right\}$$

(3.2)

$$x \in K = \left\{ x \in H^1(0, T; \mathbb{R}^{2n+2}) \mid \alpha_{\min} \leq x_i(t) \leq \alpha_{\max}, \alpha_{\min} = -\alpha_{\max}, \forall i \in [n+2, 2n+2] \right\}$$

(3.3)

$$x_i(t_0) = x, \quad \forall i \in [1, 2n+2]. \quad (3.4)$$

Очевидно, что объект управления (3.1) – (3.4) удовлетворяет предположениям 1)–4) предыдущего пункта.

Пусть Λ – конус положительных элементов в $L^2(t_0, T)$, т.е.:

$$\Lambda = \left\{ f \in L^2(t_0, T) : f(t) \geq 0 \text{ на } [t_0, T] \right\}.$$

Как известно из [3], конус Λ является правильным.

Для каждой пары $(u_x) \in L^2(t_0, T; \mathbb{R}^{n+1}) \times H^1(0, T; \mathbb{R}^{2n+2})$ определим целевое отображение $L: L^2(t_0, T; \mathbb{R}^{n+1}) \times H^1(0, T; \mathbb{R}^{2n+2}) \rightarrow L^2(t_0, T)$ по правилу:

$$L(x(t), x(t)) = \sum_{i=1}^n \left(m_i \left[x_{2n+2}^2(t) R_i \cos x_{n+1}(t) + x_{2n+1}^2(t) R_i \sin x_{n+1}(t) + x_{n+1}^2(t) r_i \cos x_i(t) + x_{n+1}^2(t) r_i \sin x_i(t) \right] + m_{3vi} \frac{R_i}{2} \left[x_{2n+2}^2(t) \sin(x_{n+1}(t) + \lambda(i-1)) + x_{2n+1}^2(t) \cos(x_{n+1}(t) + \lambda(i-1)) \right] \right) \quad (3.5)$$

Функция $L(x(t), x(t))$ с физической точки зрения означает вертикальную проекцию суммарной возмущающей силы вибросистемы в момент времени t .

Множество всех допустимых пар в задаче (3.1)–(3.5) имеет вид:

$$\Xi = \left\{ (u_x) \in L^2(t_0, T; \mathbb{R}^{n+1}) \times H^1(0, T; \mathbb{R}^{2n+2}) \mid \dot{x} = f(u_x), x(0) = x_0, u \in U, \right. \quad (3.6)$$

Тогда задачу векторной оптимизации для объекта (3.1)–(3.5) сформулируем следующим образом:

$$L(x, \cdot) \rightarrow \Lambda_{\bar{\tau}} \quad \inf_{(u_x) \in \Xi} \quad (3.7)$$

С физической точки зрения это означает попытку минимизировать возмущающую силу на всем промежутке времени $[t_0, T]$, т.е. увеличить давление катка на уплотняемую среду в течение всего промежутка времени работы вибросистемы.

Поскольку объектом управления выступает реальная механическая вибросистема, то всюду далее будем считать, что задача (3.1)–(3.5) регулярна.

Для доказательства разрешимости задачи (3.1)–(3.5) рассмотрим свойства целевого отображения (3.5). Учитывая теорему о вложении Соболева, получим, что функция $x(t)$, как элемент пространства $H^1(0, T; \mathbb{R}^{2n+2})$, является непрерывной. Согласно определению пространства $H^1(0, T; \mathbb{R}^{2n+2})$ имеем: $\dot{x} \in L^2(0, T; \mathbb{R}^{2n+2})$. Взяв во внимание вид целевого отображения (3.5) видим, что $L(x, \cdot) \in L^2(t_0, T)$ для любых пар $(u_x) \in \Xi$.

Более того, используя вид уравнений системы (3.1) и учитывая выполнение условий Каратеодори можно показать, что существует суммируемая на $[t_0, T]$ функция $M(t)$ такая, что выполняется неравенство $|L(t, x(t))| \leq M(t)$ почти для всех $t \in [t_0, T]$. Отсюда следует, что $L(x, \cdot) \in L^\infty(t_0, T)$.

Исходя из результатов, полученных в работе [3], можно сделать вывод, что отображение $L(x, \cdot)$ является слабо непрерывным относительно слабой топологии в пространстве $L^2(t_0, T; \mathbb{R}^{n \times 1}) \times H^1(\cdot, T; \mathbb{R}^{2 \times 2})$, что означает: из того, что $(u_{kk}, x) \xrightarrow{w} (u^{**}, x)$ следует $L(u_{kk}, x) \rightarrow L(u^{**}, x)$ слабо в $L^2(t_0, T)$ при $k \rightarrow \infty$.

Из того, что $L(u_{kk}, x) \rightarrow L(u^{**}, x)$ слабо в $L^2(t_0, T)$ при $k \rightarrow \infty$ следует, что выполняется условие $\{L(x, \cdot)\} = \Lambda_w^- \liminf_{(u_{kk}, x) \xrightarrow{w} (u, x)} L(u_{kk}, x)$ для всех пар $(u, x) \in \Xi$. Применяя к задаче (3.1)-(3.5) теорему 2, делаем вывод, что множество Λ_τ^- -эффективных решений $Eff_\tau(\Xi_{\Lambda_\tau^-})$ в задаче (3.1)-(3.5) непусто.

Выводы. Для задач векторной оптимизации объекта управления, который описывается системой нелинейных обыкновенных дифференциальных уравнений с начальными условиями и ограничениями типа неравенств на управления и фазовые траектории получены достаточные условия их разрешимости в классе Λ_τ^- -эффективных решений. Приведено применение полученных результатов на примере векторной оптимизации динамики вибросистемы.

Бibliографические ссылки

1. Богомаз В.Н. Об одной задаче оптимизации механической вибросистемы // В.Н. Богомаз // Питання прикладної математики та математичного моделювання: зб. наук. праць. – Д., 2010. – С. 23–40.

2. **Богомаз В.Н.** Необходимые условия экстремума в задаче оптимизации механической вибросистемы / В.Н. Богомаз // Вісник Дніпропетровського ун-ту, Моделювання – 2010. Т.18, №2 – С. 90–102.
3. **Богомаз В.Н.** Численный анализ задачи оптимального управления механической вибросистемой / В.Н. Богомаз, И.В. Шаповал // Вісник Дніпропетровського ун-ту, Моделювання – 2011. Т.18, №2 – С. 30–43.
4. **Красносельский М.А.** Положительные решения операторных уравнений. / М.А. Красносельский – М., 1962. – 396с.
5. **Kogut P.I.** On the existence of efficient solutions of vector optimization problems in Banach spaces / P.I. Kogut, R. Manzo, I.V. Nechay // Диференціальні рівняння та їх застосування. – 2008 – № 5 – С. 105–121.
6. **Luc D.T.** Theory of vector optimization / Luc D.T. – Lnems 319: Springer-Verlag, 1989. – 343p.
7. **Ehrgott M.** Multicriteria Optimization / Ehrgott M. – Berlin: Springer, 2005. – 323 p.

Надійшла до редколегії 28.01.11.

Я.С. Бондаренко, Д.С. Бабічева, В.М. Турчин

Дніпропетровський національний університет імені Олеся Гончара

КРИТЕРІЙ КЛАСИФІКАЦІЇ ТОЧКИ БІФУРКАЦІЇ ЯК ТОЧКИ РОСТУ ПІДДЕРЕВА ДЕНДРИТА НЕЙРОНА

Запропонована методика оцінювання ймовірності класифікації точки біфуркації як точки росту піддерева дендрита нейрона. Визначені характеристики дендрита, за якими можна класифікувати точки біфуркації як точки росту піддерева. Побудований статистичний критерій дозволяє класифікувати «сумнівні» точки біфуркації дендрита нейрона, а також точки біфуркації інших дендритів даного класу нейронів.

Предложена методика оценивания вероятности классификации точки бифуркации как точки роста поддерева дендрита нейрона. Определены характеристики дендрита, по которым можно классифицировать точку бифуркации как точку роста поддерева. Построенный статистический критерий позволяет классифицировать «сомнительные» точки бифуркации дендрита нейрона, а также точки бифуркации других дендритов данного класса нейронов.

The estimation's technique of probability of classification of bifurcation's point as point of subtree's dendritic growth is offered. The dendritic characteristics, on which it is possible to classify a bifurcation's point as a point of growth subtree, are determined. The constructed statistical criterion allows to classify «doubtful» bifurcation's points and bifurcation's points other dendrites of the given class of neurons.

Ключові слова: нейрон, дендрит, точка біфуркації, ймовірність, логістична регресія, функція ризику, дециль ризику.

Вступ. Сучасний етап розвитку фізіології нервової системи характеризується побудовою чисельних біофізичних моделей нервових клітин, які все повніше описують реальні біологічні нейрони. Проте, при розробці таких моделей недостатньо уваги приділяється структурі нейронів, зокре-

ма геометрії дендритів (рис. 1, 2), хоча спостереження природи розгалужених відростків вказують на їх виключну роль.

Активні математичні дослідження дендритів біологічних нейронів проводилися в двох основних напрямках: по _перше, аналізувалися і моделювалися метричні та топологічні характеристики, ймовірності розгалуження сегментів, по _друге, проводилося моделювання росту дендрита в цілому.

Метричні характеристики дендрита досліджувалися в [2–6]. Описувалися і вивчалися діаметр сегмента, число сегментів дендрита, загальна довжина дерева, довжина проміжних та кінцевих сегментів, відстань кінців дендрита й точок розгалуження від соми і т. ін. Більшість робіт була сконцентрована на вивченні двох метричних характеристик: довжини сегмента та кутів розгалуження.

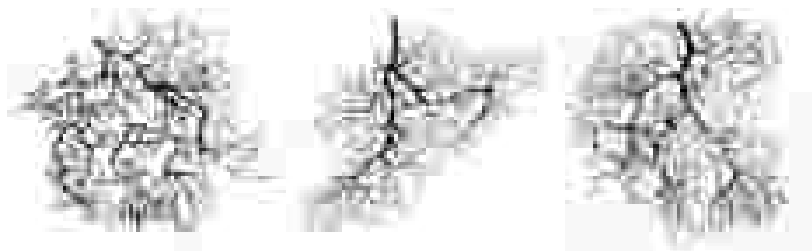


Рис. 1. Дендрити Пуркін'є клітин кори мозочка морської свинки [1]



Рис. 2. Дендрити мотонейронів спинного мозку кішки [1]

Топологічні характеристики дендрита вивчалися в [7 ; 8]. Досліджувалися порядок сегмента, ступінь сегмента, структура розгалуження, число кінцевих сегментів і т. ін. Топологічна модель росту дендрита була запропонована в [9]. За допомогою цієї моделі в [10] отримано опис структур дендритних дерев Пуркін'є клітин і пірамідальних нейронів.

У [11; 12] запропоновано альтернативний підхід до вивчення топології дендритного дерева, а саме, описання дендрита як реалізації стохастичного процесу, побудова моделі цього процесу та оцінка параметрів моделі. Цей підхід важливий як з точки зору вивчення механізму росту окремих дендритів нейрона, так і з точки зору порівняння структур різних типів нейронів.

Невід'ємною складовою дослідження біологічних нервових клітин стало математичне моделювання дендритів. Адекватними методами математичного моделювання дендритних дерев виявилися ймовірнісні методи. Розробці алгоритмів моделювання дендритів та порівнянню моделей з реальними дендритами присвячені [13–22], зокрема в [18–22] авторами цієї статті запропоновано принципово нову загальну ймовірнісну модель побудови дендритів нейронів.

Постановка задачі. Дендрит нейрона є складним стохастичним об'єктом. Ймовірнісна природа дендрита проявляється у стохастичності його характеристик: довжині сегментів, кутах між сегментами, діаметрі сегментів, відстані між точками росту піддерев та інших.

Структура, розміри та форма нейронів суттєво різняться, при цьому спостерігаються відмінності як між дендритами нейронів одного типу, так і різних типів (див. рис. 1, 2). Відмінності проявляються у формі, розмірі, структурі біфуркації, що визначають функціональні властивості дендритних дерев.

Структура біфуркації дендрита визначається точками біфуркації. Точкою біфуркації сегмента може виступати як точка розгалуження сегмента (рис. 3, праворуч), так і точка росту піддерева дендрита нейрона (див. рис. 3, ліворуч). Точці розгалуження сегмента притаманні, як правило, симетричність структури біфуркації, несуттєві відмінності діаметрів дочірніх сегментів, чого не можна сказати про точку росту піддерева дендрита нейрона.



Рис. 3. Структура біфуркації сегмента: ліворуч – точка росту піддерева, у центрі – «сумнівні» точки, праворуч – точка розгалуження

При вивченні дендрита низку точок біфуркації можна класифікувати як точки розгалуження сегмента та точки росту піддерева дендрита нейрона, а для деяких точок біфуркації класифікація викликає труднощі навіть у досвідчених дослідників (рис. 3, по центру). Тому виникає задача знаходження характеристик дендрита, за значеннями яких можна було б диференціювати точку біфуркації як точку росту піддерева або як точку розгалуження сегмента. Такими характеристиками можуть виступати діаметри дочірніх сегментів, діаметр материнського сегмента, кути між материнським та дочірнім сегментами тощо.

У роботі запропоновано статистичний критерій класифікації точки біфуркації як точки росту піддерева дендрита нейрона. Побудований критерій дозволяє класифікувати «сумнівні» точки біфуркації дендрита нейрона, а також точки біфуркації інших дендритів даного класу нейронів. Розроблена методика оцінювання ймовірності класифікації точки біфуркації як точки росту піддерева дозволить глибше пізнати особливості структури різних класів біологічних нейронів.

Термінологія. Для опису структури дендрита використано таку термінологію. *Точка розгалуження* – точка біфуркації (маркер 1 на рис. 4). *Дендритний сегмент* – частина дендрита між двома точками розгалуження (маркер 2 на рис. 4) або між сомою та найближчою точкою розгалуження, або між точкою розгалуження та кінцем дендрита (маркер 3 на рис. 4). *Материнський сегмент даного сегмента* – сегмент, що його породжує (на рис. 4 сегмент 2 є материнським сегментом для сегментів 5). *Дочірній сегмент даного сегмента* – це породжений ним сегмент (на рис. 4 сегменти 5

є дочірніми для сегмента 2); кожен сегмент має 0 чи 2 дочірніх сегменти. *Кінцевий сегмент* – сегмент, що має 0 дочірніх сегментів (маркер 4 на рис. 4). *Ланка сегмента* – частина сегмента між двома сусідніми «оцифрованими» точками сегмента (маркер 6 на рис. 4). *Кут між материнським та дочірнім сегментами* – це кут, утворений прямою, проведеною у напрямку першої ланки дочірнього сегмента й прямою, проведеною у напрямку останньої ланки материнського сегмента (маркер 7 на рис. 4). *Проміжний кут* – це кут, утворений прямою, проведеною у напрямку даної ланки і прямою, проведеною у напрямку попередньої ланки того самого дендритного сегмента (маркер 8 на рис. 4). *Кореневий сегмент* – дочірній сегмент, який утворює з материнським сегментом кут, близький до прямого (маркер 9 на рис. 4). Цей кут називатимемо *кутом між кореневим та материнським сегментами* (маркер 10 на рис. 4). *Піддерево (піддендрит)* – упорядкована множина сегментів, утворена кореневим сегментом та іншими некореновими сегментами, кожен із яких закінчується точкою розгалуження чи кінцем дендрита. *Точка росту піддерева* – точка росту кореневого сегмента. *Рівень піддерева* визначається рівнем його кореневого сегмента. Сегмент, який росте з соми, будемо називати кореневим сегментом 1-го рівня (маркер 11 на рис. 4). Кореневий сегмент, який росте на піддереві 1-го рівня, називатимемо кореневим сегментом 2-го рівня і т. д. *Сегмент з піддеревами* – сегмент, що має принаймні одне піддерево. *Сегмент без піддерев* – сегмент, що не має жодного піддерева.

Отже, дендрит нейрона розглядається як упорядкована множина сегментів, кожен із яких закінчується точкою розгалуження чи кінцем дендрита. Будемо вважати, що піддерево має *напрямок росту ліворуч*, якщо його кореневий сегмент розташований ліворуч відносно материнського сегмента (маркер 9 на рис. 4), у супротивному разі піддерево має *напрямок росту праворуч*.

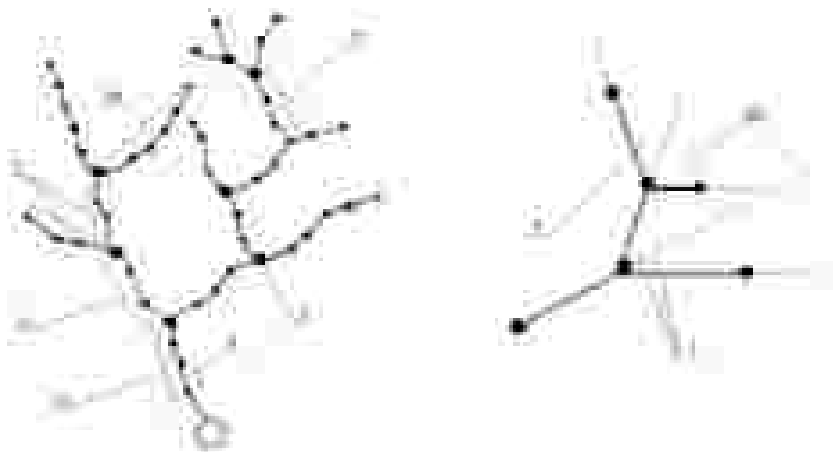


Рис. 4. До визначення структури дендрита: 1 – точка розгалуження; 2 – материнський сегмент для сегментів 5; 3 – дендритний сегмент; 4 – кінцевий сегмент; 5 – дочірні сегменти для сегмента 2; 6 – ланка сегмента; 7 – кут між материнським та дочірнім сегментами; 8 – проміжний кут; 9 – кореневий сегмент 2-го рівня; 10 – кут між кореневим і материнським сегментами; 11 – кореневий сегмент 1-го рівня

Модель логістичної регресії. Позначимо ймовірність диференціації точки біфуркації як точки росту піддерева через p , тоді ймовірність диференціації точки біфуркації як точки розгалуження дорівнює $1 - p$.

Нехай щільність розподілу ймовірностей вектора характеристик дендрита $\mathbf{x} = (x_1, \dots, x_k)$ для точок росту піддерева є $f_1(x_1, \dots, x_k)$, а для точок розгалуження – $f_0(x_1, \dots, x_k)$. Тоді за формулою Байєса ймовірність $P(x_1, \dots, x_k)$ того, що точка біфуркації з вектором характеристик $\mathbf{x} = (x_1, \dots, x_k)$ є точкою росту піддерева дорівнює

$$P(x_1, \dots, x_k) = \frac{pf_1(x_1, \dots, x_k)}{pf_1(x_1, \dots, x_k) + (1-p)f_0(x_1, \dots, x_k)} = \frac{1}{1 + \frac{1-p}{p} \frac{f_0(x_1, \dots, x_k)}{f_1(x_1, \dots, x_k)}}. \quad (1)$$

Будемо виходити з припущення, що вектор характеристик дендрита $\mathbf{x} = (x_1, \dots, x_k)$ має нормальний розподіл зі щільністю

$$f_0(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det \mathbf{C}}} \exp \left\{ -\frac{(\mathbf{C}^{-1}(\mathbf{x} - \mathbf{a}_0), \mathbf{x} - \mathbf{a}_0)}{2} \right\}, \quad (2)$$

для точки розгалуження і

$$f_1(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det \mathbf{C}}} \exp \left\{ -\frac{(\mathbf{C}^{-1}(\mathbf{x} - \mathbf{a}_1), \mathbf{x} - \mathbf{a}_1)}{2} \right\}, \quad (3)$$

для точки росту піддерева дендрита нейрона. Матриця коваріацій $\mathbf{C}$ одна й та сама. Тоді маємо

$$\frac{1-p}{p} \frac{f_0(x_1, \dots, x_k)}{f_1(x_1, \dots, x_k)} = \exp \left\{ -\left(\alpha_0 + \sum_{i=1}^k \alpha_i x_i \right) \right\}, \quad (4)$$

де α_0 та α_i подані через $p, \mathbf{a}_0, \mathbf{a}_1, \mathbf{C}^{-1}$.

Таким чином, отримали ймовірність $P(x_1, \dots, x_k)$ класифікації точки біфуркації з вектором характеристик $\mathbf{x} = (x_1, \dots, x_k)$ як точки росту піддерева

$$P(x_1, \dots, x_k) = \frac{1}{1 + \exp \left(-\alpha_0 - \sum_{i=1}^k \alpha_i x_i \right)}. \quad (5)$$

Імовірність $P(x_1, \dots, x_k)$ виявилася функцією від лінійної комбінації характеристик $\mathbf{x} = (x_1, \dots, x_k)$

$$y = \alpha_0 + \sum_{i=1}^k \alpha_i x_i. \quad (6)$$

Функцію $y = \alpha_0 + \sum_{i=1}^n \alpha_i x_i$ називатимемо *функцією ризику за значеннями характеристик* $x_1, \dots, x_k$. Оцінки невідомих коефіцієнтів $\alpha_0, \alpha_1, \dots, \alpha_n$ знаходяться за методом максимальної правдоподібності.

Визначення характеристик дендрита, що описують точку біфуркації. Вихідним матеріалом для дослідження дендрита був оцифрований дендрит Пуркін'є клітини кори мозочка морської свинки (клітини Пуркін'є є основним структурним елементом кори мозочка, див. рис. 1).

Побудуємо статистичний критерій класифікації точки біфуркації як точки росту піддерева дендрита нейрона (необхідно встановити зв'язок між імовірністю класифікації точки біфуркації як точки росту піддерева дендрита нейрона та значеннями цих характеристик).

Спочатку розглянемо випадок, коли піддерево має напрямок росту ліворуч (рис. 5, ліворуч). Вихідні дані складаються з 367 точок біфуркації, з них 286 точок виявилися напевно точками розгалуження, 24 точки виявилися напевно точками росту піддерев, 57 точки біфуркації виявилися «сумнівними».

286 напевно точок розгалуження та 24 напевно точок росту піддерева використаємо для побудови критерію – інструменту для класифікації точок біфуркації, який далі застосуємо для класифікації 57 «сумнівних» точок біфуркації.

Точку біфуркації будемо характеризувати *величинами лівого (x_l) та правого (x_n) кутів між материнським та дочірнім сегментами, діаметрами лівого (d_l) та правого (d_n) дочірніх сегментів, діаметром материнського сегмента (d_m)* (див. рис. 5).

Зазначені характеристики є нормально розподіленими випадковими величинами. Гіпотези про нормальний розподіл характеристик перевірено за допомогою критерію χ^2 на рівні значущості 0,05.

Виявлено суттєву відмінність середніх значень таких характеристик, які описують точку розгалуження та точку росту піддерева: *величини лівого (x_l) кута між материнським та дочірнім сегментами, величини правого (x_n) кута між материнським та дочірнім сегментами, діаметра*

правого (d_n) дочірнього сегмента, діаметра материнського сегмента (d_m). Гіпотези про рівність середніх значень перелічених характеристик перевірено за допомогою критерію Стюдента на рівні значущості 0,05 (табл. 1).

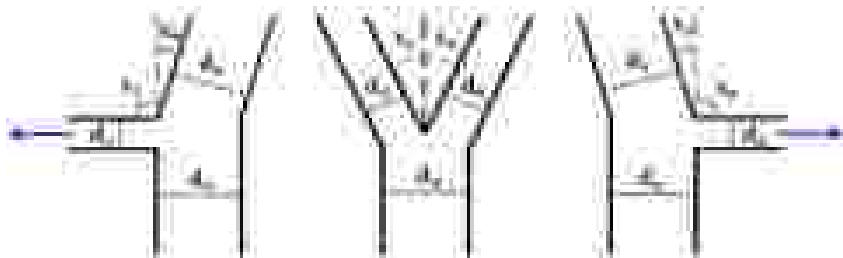


Рис. 5. Структура біфуркації сегмента: ліворуч – напрямок росту піддерва ліворуч, у центрі – розгалуження сегмента, праворуч – напрямок росту піддерва праворуч

Таблиця 1

Результати перевірки гіпотези про рівність середніх значень характеристик

Характеристика	Середнє значення		Значення статистики критерію	$t_{0,05;308}$	Гіпотеза про рівність середніх значень
	Точка росту піддерва	Точка розгалуження			
x_l	84,62	36,31	13,23	1,65	Відхиляється
x_n	8,07	37,12	8,13	1,65	Відхиляється
d_l	0,74	0,71	0,35	1,65	Не відхиляється
d_n	1,15	0,71	4,56	1,65	Відхиляється
d_m	1,24	0,93	2,47	1,65	Відхиляється

Величина лівого (x_l) кута між материнським та дочірнім сегментами, величина правого (x_n) кута між материнським та дочірнім сегментами, діаметр правого (d_n) дочірнього сегмента – незалежні. Гіпотези про незалежність характеристик перевірено за допомогою критерію χ^2 незалежності ознак на рівні значущості 0,05.

Тепер розглянемо випадок, коли піддерево має напрямок росту праворуч (див. рис. 5, праворуч). Вихідні дані складаються з 387 точок біфуркації, з них 286 точок виявилися напевно точками розгалуження, 21 точки виявилися напевно точками росту піддерев, 79 точки біфуркації виявилися „сумнівними”. Точку біфуркації будемо характеризувати величинами лівого (x_l) та правого (x_n) кутів між материнським та дочірнім сегментами, діаметрами лівого (d_l) та правого (d_n) дочірніх сегментів, діаметром материнського сегмента (d_m) (див. рис. 5).

Зазначені характеристики є нормально розподіленими. Гіпотези про нормальний розподіл характеристик перевірено за допомогою критерію χ^2 на рівні значущості 0,05.

Таблиця 2

Результати перевірки гіпотези про рівність середніх значень характеристик

Характеристика	Середнє значення		Значення статистики критерію	$t_{0,05;307}$	Гіпотеза про рівність середніх значень
	Точка росту піддерев	Точка розгалуження			
x_l	7,56	36,31	7,45	1,65	Відхиляється
x_n	85,08	37,12	12,5	1,65	Відхиляється
d_l	0,94	0,71	2,3	1,65	Відхиляється
d_n	0,7	0,71	0,18	1,65	Не відхиляється
d_m	0,94	0,93	0,1	1,65	Не відхиляється

Величина лівого (x_l) кута між материнським та дочірнім сегментами, величина правого (x_n) кута між материнським та дочірнім сегментами, діаметр лівого ($d_{\bar{e}}$) дочірнього сегмента – незалежні. Гіпотези про незалежність характеристик перевірено за допомогою критерію χ^2 незалежності ознак на рівні значущості 0,05.

Критерій класифікації точки біфуркації як точки росту піддерева (піддерево має напрямок росту ліворуч). Кожну точку біфуркації будемо характеризувати вектором характеристик $x = (x_l, x_n, d_n)$. Установимо зв'язок між ймовірністю класифікації точки біфуркації як точки росту піддерева та значеннями характеристик $x = (x_l, x_n, d_n)$. Будемо виходити з того, що ймовірність $P(x_l, x_n, d_n)$ класифікації точки біфуркації як точки росту піддерева з вектором характеристик $x = (x_l, x_n, d_n)$ є функцією

$$P(y) = P(x_l, x_n, d_n) = 1/(1 + \exp\{-y\})$$

від лінійної комбінації

$$y = \alpha_0 + \alpha_1 x_l + \alpha_2 x_n + \alpha_3 d_n$$

характеристик $x = (x_l, x_n, d_n)$. Коефіцієнти $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ оцінюються за методом максимальної правдоподібності. Оцінки коефіцієнтів $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, знайдені за вибіркою, позначатимемо через $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3$. У нашій задачі значення оцінок виявилися такими

$$\hat{\alpha}_0 = -67,96, \quad \hat{\alpha}_1 = 1,08, \quad \hat{\alpha}_2 = -0,48, \quad \hat{\alpha}_3 = -2,06.$$

Функція ризику набуває вигляду

$$y = -67,96 + 1,08x_l - 0,48x_n - 2,06d_n, \quad (7)$$

а ймовірність класифікації точки біфуркації як точки росту піддерева з вектором характеристик $x = (x_l, x_n, d_n)$ дорівнює

$$P(y) = P(x_l, x_n, d_n) = 1/(1 + \exp\{67,96 - 1,08x_l + 0,48x_n + 2,06d_n\}) \quad (8)$$

На рис. 6 зображено логістичну функцію. На графіку точками позначено значення функції, чисельно рівні ймовірностям $P(y) = P(x_l, x_n, d_n)$ класифікації точки біфуркації з вектором характеристик $x = (x_l, x_n, d_n)$ як точки росту піддерева. Одиницею позначено ймовірності (8) для точок росту піддерева, нулем – для точок розгалуження. За вектором характеристик $x = (x_l, x_n, d_n)$ кожної точки біфуркації можна обчислити значення функції ризику за формулою (7) та ймовірність класифікації точки біфуркації як точки росту піддерева за формулою (8).

Наприклад, для точки біфуркації з вектором характеристик $(x_l, x_n, d_n) = (82,75; 4,76; 1)$ ймовірність класифікації точки біфуркації як точки росту піддерева становить 0,99, а для точки біфуркації з вектором характеристик $(x_l, x_n, d_n) = (53,13; 32,1; 0,45)$ ймовірність класифікації точки біфуркації як точки росту піддерева становить 0,01, тобто майже в 100 разів менша.

Запропонована модель (8) була застосована для класифікації 57 «сумнівних» точок біфуркації. Для кожної точки біфуркації, зокрема, «сумнівної», відомі значення вектора характеристик $x = (x_l, x_n, d_n)$. На рис. 7 зображено логістичну функцію. На графіку точками позначено значення функції, чисельно рівні ймовірностям $P(y) = P(x_l, x_n, d_n)$ класифікації «сумнівної» точки біфуркації як точки росту піддерева з вектором характеристик $x = (x_l, x_n, d_n)$ згідно з моделлю (8).

Очікувана та фактична кількість випадків класифікації точки біфуркації як точки росту піддерева. З'ясуємо, наскільки добре погоджується запропонована модель залежності ймовірності $P(x_l, x_n, d_n)$ від вектора характеристик $x = (x_l, x_n, d_n)$. Для цього впорядкуємо значення функції ризику в точках біфуркації за величиною (від менших до більших) і розіб'ємо всі отримані значення функції ризику на 10 рівних за чисельністю груп. До першої групи ввійдуть 10 % точок біфуркації, значення функції ризику в яких мінімальне, до другої 10 % точок біфуркації, що мають більші значення функції ризику і т. д. і в останню групу включимо 10 % точок біфуркації, значення функції ризику яких набуває максимального значення.

Називатимемо ці групи *децилями*. Для кожної децилі можна обчислити *очікувану кількість* випадків класифікації точки біфуркації як точки росту піддерева (підсумовуючи для цієї групи значення ймовірностей (8)) та *фактичну кількість* випадків класифікації точки біфуркації як точки росту піддерева.

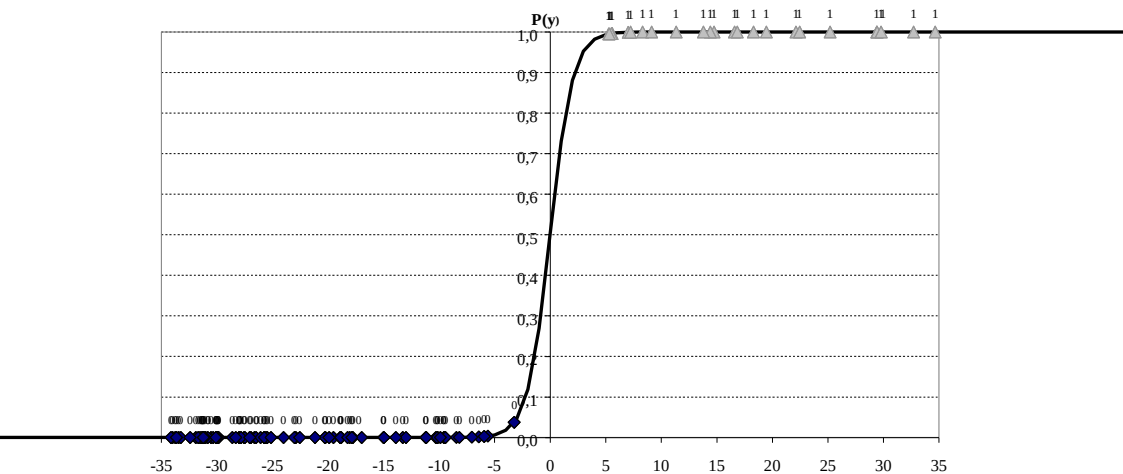


Рис. 6. Значення функції логістичної регресії для 286 напевно точок розгалуження та 24 напевно точок росту піддерев

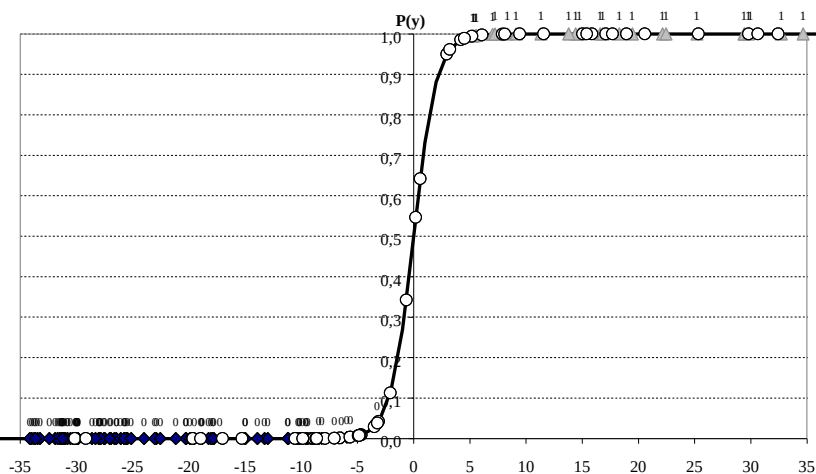


Рис. 7. Значення функції логістичної регресії для 57 «сумнівних» точок біфуркації (позначені порожнім кружечком)

Дійсно, нехай ймовірність класифікації i -тої точки біфуркації як точки росту піддерева дорівнює $P(y_i)$, тоді ймовірність класифікації i -тої точки біфуркації як точки розгалуження сегмента дорівнює $1 - P(y_i)$. Індикатор I_i події « i -та точка біфуркації класифікується як точка росту піддерева» має розподіл

$$\begin{pmatrix} 1 & 0 \\ P(y_i) & 1 - P(y_i) \end{pmatrix} \quad i=1, \dots, n.$$

Математичне сподівання індикатора дорівнює $M I_i = 1 \cdot P(y_i) + 0 \cdot (1 - P(y_i)) = P(y_i)$, $i=1, \dots, n$. Тоді очікувана кількість випадків класифікації точки біфуркації як точки росту піддерева дорівнює

$$M \sum_{i=1}^n I_i = \sum_{i=1}^n M I_i = \sum_{i=1}^n P(y_i), \text{ де } P(y_i) = 1 / (1 + \exp\{-y_i\}).$$

Фактична кількість точок росту піддерев та очікувана кількість випадків класифікації точки біфуркації як точки росту піддерев наведені в таблиці 3. Очікувана та фактична кількість точок росту піддерев збігаються.

Таблиця 3

Очікувана і фактична кількість випадків класифікації точки біфуркації як точки росту піддерев за децилями ризику

Дециль ризику	Кількість випадків класифікації точки біфуркації як точки росту піддерев		Фактичний % точок росту піддерев
	очікувана	фактична	
10	24	24	100%
9	0	0	0%
8	0	0	0%
7	0	0	0%
6	0	0	0%
5	0	0	0%
4	0	0	0%
3	0	0	0%
2	0	0	0%
1	0	0	0%
Разом	24	24	7,73%

Спробуємо оцінити внесок кожної характеристики $x = (x_l, x_n, d_n)$ дендрита до значення ризику прийняття рішення щодо класифікації точки біфуркації як точки росту піддерев дендрита нейрона. Оцінки коефіцієнтів функції ризику

$$y = -67,96 + 1,08x_l - 0,48x_n - 2,06d_n$$

необхідно інтерпретувати так. Збільшення величини лівого (x_l) кута між материнським та дочірнім сегментами на 1 градус призводить до зростання ризику класифікації точки біфуркації як точки росту піддерев (надалі – ризику) на 1,08, в той час як зменшення величини правого (x_n) кута між материнським та дочірнім сегментами на 1 градус дає зростання ризику на 0,48. Зменшення діаметра правого (d_n) дочірнього сегмента на 1 μm збільшує ризик на 2,06.

Критерій класифікації точки біфуркації як точки росту піддерева (піддерево має напрямок росту праворуч). Методика класифікації точки біфуркації як точки росту піддерева (піддерево має напрямок росту праворуч) аналогічна.

Кожну точку біфуркації будемо характеризувати вектором характеристик $x = (x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$. Функція ризику набуває вигляду

$$y = -80,91 - 0,67x_{\bar{i}} + 1,23x_{\bar{n}} + 2,55d_{\bar{i}}, \quad (9)$$

а ймовірність класифікації точки біфуркації як точки росту піддерева з вектором характеристик $x = (x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$ дорівнює

$$P(y) = P(x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}}) = 1/(1 + \exp\{80,91 + 0,67x_{\bar{e}} - 1,23x_{\bar{i}} - 2,55d_{\bar{e}}\}). \quad (10)$$

На рис. 8 зображено логістичну функцію. На графіку точками позначено значення функції, чисельно рівні ймовірностям $P(y) = P(x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$ класифікації точки біфуркації з вектором характеристик $x = (x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$ як точки росту піддерева. Одиницею позначено ймовірності (10) для точок росту піддерева, нулем – для точок розгалуження. За вектором характеристик $x = (x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$ кожної точки біфуркації можна обчислити значення функції ризику за формулою (9) та ймовірність класифікації точки біфуркації як точки росту піддерева за формулою (10).

Наприклад, для точки біфуркації з вектором характеристик $(x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}}) = (3,91; 86,19; 0,5)$ ймовірність класифікації точки біфуркації як точки росту піддерева становить 0,99, а для точки біфуркації з вектором характеристик $(x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}}) = (29,74; 78, 69; 0,4)$ ймовірність класифікації точки біфуркації як точки росту піддерева становить 0,05, тобто майже в 20 разів менша.

Запропонована модель (10) була застосована для класифікації 79 «сумнівних» точок біфуркації. Для кожної точки біфуркації, зокрема, «сумнівної», відомі значення вектора характеристик $x = (x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$. На рис. 9 зображено логістичну функцію. На графіку точками позначено значення функції, чисельно рівні ймовірностям $P(y) = P(x_{\bar{e}}, x_{\bar{i}}, d_{\bar{e}})$ класифікації

«сумнівної» точки біфуркації як точки росту піддерева з вектором характеристик $x = (x_{\bar{e}}, x_i, d_{\bar{e}})$ згідно з моделлю (10).

Фактична кількість точок росту піддерева та очікувана кількість випадків класифікації точки біфуркації як точки росту піддерева наведені в таблиці 4. Очікувана та фактична кількість точок росту піддерева збігаються.

Оцінимо внесок кожної характеристики дендрита до значення ризику прийняття рішення щодо класифікації точки біфуркації як точки росту піддерева дендрита нейрона. Оцінки коефіцієнтів функції ризику

$$y = -80,91 - 0,67x_{\bar{e}} + 1,23x_i + 2,55d_{\bar{e}}$$

необхідно інтерпретувати так. Зменшення величини лівого ($x_{\bar{e}}$) кута між материнським та дочірнім сегментами на 1 градус призводить до зростання ризику класифікації точки біфуркації як точки росту піддерева (надалі – ризику) на 0,67, в той час як збільшення величини правого (x_i) кута між материнським та дочірнім сегментами на 1 градус дає зростання ризику на 1,23. Збільшення діаметра лівого ($d_{\bar{e}}$) дочірнього сегмента на 1 μm збільшує ризик на 2,55.

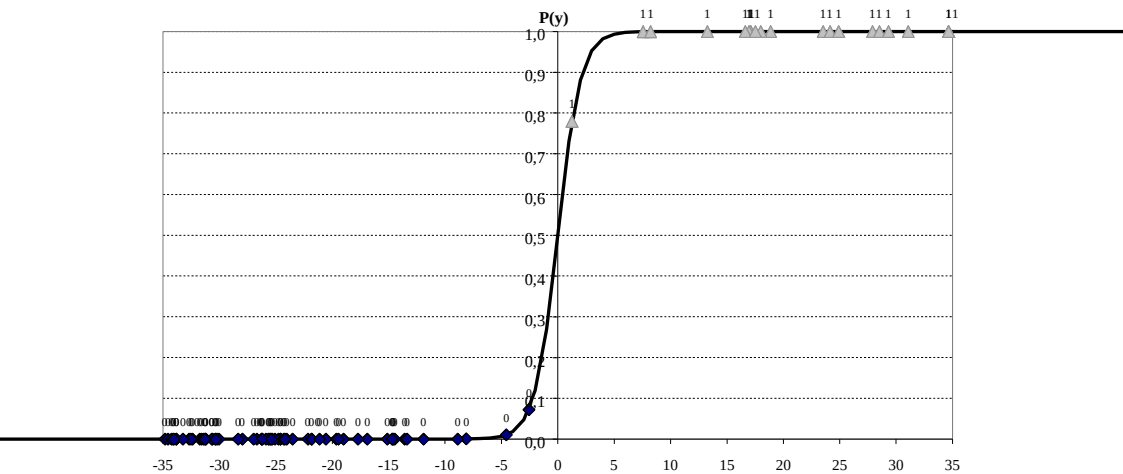


Рис. 8. Значення функції логістичної регресії для 286 напевно точок розгалуження та 21 напевно точок росту піддерева

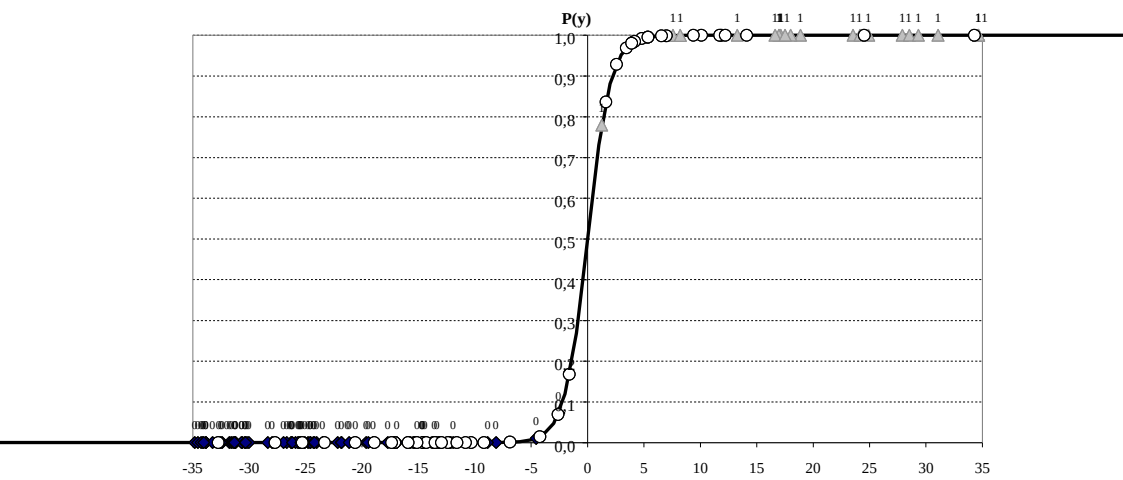


Рис. 9. Значення функції логістичної регресії для 79 «сумнівних» точок біфуркації (позначені порожнім кружечком)

Таблиця 4

Очікувана і фактична кількість випадків класифікації точки біфуркації як точки росту піддерева за децилями ризику

Дециль ризику	Кількість випадків класифікації точки біфуркації як точки росту піддерева		Фактичний % точок росту піддерева
	очікувана	фактична	
10	21	21	100%
9	0	0	0%
8	0	0	0%
7	0	0	0%
6	0	0	0%
5	0	0	0%
4	0	0	0%
3	0	0	0%
2	0	0	0%
1	0	0	0%
Разом	21	21	6.84%

Висновки. Задача створення нейронної мережі, що є перспективною задачею сьогодення, включає дослідження біологічних нейронів, розробку математичної моделі дендритів нейронів та концепції нейронної мережі. У 2004–2006 рр. авторами статті було розроблено теоретико-ймовірнісну модель дендритів нейронів [18–21]. При розробці моделі виявилось, що точки біфуркації сегментів природно розділяються на два класи: перший клас складають точки розгалуження сегментів, другий – точки росту піддерев. Питання щодо характеристик дендрита, які описують точку біфуркації сегмента до сьогодні залишалося відкритим.

Існують певні уявлення в якісному відношенні безумовно правильні, але в кількісному відношенні вельми неоднозначні щодо диференціації точки біфуркації сегмента дендрита нейрона, а саме диференціації точки росту піддерева та точки розгалуження сегмента.

Утворення піддерева дендрита нейрона характеризується асиметричністю структури біфуркації, спеціальними формами кореневого сегмента та продовження материнського сегмента, тощо. Але ці характеристики кількісно не описуються. Тому становлять інтерес інші характеристики, які пов'язані з вищезазначеними та вимірні. Цими характеристиками виявилися: величина лівого (χ_l) та величина правого (χ_n) кутів між материнським та дочірнім сегментами, діаметр правого (d_n) дочірнього сегмента (піддерево має напрямок росту ліворуч), діаметр лівого (d_l) дочірнього сегмента (піддерево має напрямок росту праворуч).

Розроблено критерій класифікації точки біфуркації як точки росту піддерева за векторами характеристик дендрита нейрона. Запропонований критерій дозволяє класифікувати «сумнівні» точки біфуркації дендрита нейрона, а також класифікувати точки біфуркації інших дендритів даного класу нейронів.

Практичне значення побудованого критерія полягає в усуненні суб'єктивного підходу до класифікації точок біфуркації та у повній автоматизації диференціації точки біфуркації як точки росту піддерева дендрита нейрона.

Бібліографічні посилання

1. <http://www.krasnow.gmu.edu/I-Neuron/I-Neuron/index.html>

2. **Hillman D.E.** Neuronal shape parameters and substructures as a basis of neuronal form. In: The neurosciences, fourth study program (ed. F. Schmitt). – Cambridge, 1979. – P. 477–498.
3. **Hillman D.E.** Parameters of dendritic shape and substructure: intrinsic and extrinsic determination? In: Intrinsic determinants of neuronal form and function (ed. R.J.Lasek & M.M. Black). – New York, 1988. – P. 83–113.
4. **Uylings H.B.M.** The metric analysis of three-dimensional dendritic tree pattern: a methodological review / Uylings H.B.M., Ruiz-Marcos A., J. van Pelt // *Journal of Neuroscience Methods*. – 1986. – V. 18. – P. 127–151.
5. **Nowakowski R.S.** Competitive interactions during dendritic growth: a simple stochastic growth algorithm / Nowakowski R.S., Haye N.L., Egger M.D. // *Brain Research*. – 1992. – V. 576. – P. 152–156.
6. **Van Veen M.P.** Terminal and intermediate segment lengths in neuronal trees with finite length / Van Veen M.P., Van Pelt J. // *Bulletin of Mathematical Biology*. – 1993. – V. 55. – P. 277–294.
7. **Verwer R.W.H.** Multifurcations in topological trees: growth models and comparative analysis / Verwer R.W.H., Van Pelt J. // *Acta Stereologica*. – 1987. – Vol. 6/III. – P. 399–404.
8. **Verwer R.W.H., Van Pelt J., Uylings H.B.M.** An introduction to topological analysis. In: Quantitative method in neuroanatomy (ed. M.G. Stewart) – P. 295–323.
9. **Van Pelt J.** Topological properties of binary trees grown with order-dependent branching probabilities / Van Pelt J., Verwer R.W.H. // *Bulletin of Mathematical Biology*. – 1986. – V. 48. – P. 197–211.
10. **Van Pelt J.** Tree asymmetry: a sensitive and practical measure for binary topological trees / Van Pelt J., Uylings H.B.M., Verwer R.J., Pentney R.J., Woldenberg M.J. // *Bulletin of Mathematical Biology*. – 1992. – V. 54. – P. 759–784.
11. **Kliemann W.** A stochastic dynamical model for the characterization of the geometrical structure of dendritic processes // *Bulletin of Mathematical Biology*. – 1987. – V. 49. – P. 135–152.
12. **Uemura E.** Mathematical modeling of dendritic growth in vitro / Uemura E., Carriquiry A., Kliemann W., Goodwin J. // *Brain Research*. – 1995. – Vol. 671. – P. 187–194.
13. **Burke R.E.** A parsimonious description of motoneurons dendritic morphology using computer simulation / Burke R.E., Marks W.B., Ulfhake B. // *The Journal of Neuroscience*. – 1992. – V. 12. – P. 2403–2416.
14. **Ascoli G.A.** L-Neuron: A modeling tool for the efficient generation and parsimonious description of dendritic morphology / Ascoli G.A., Krichmar J.L. // *Neurocomputing* – 2000. – V. 32–33. – P. 1003–1011.
15. **Ascoli G.A.** Computer generation and quantitative morphometric analysis of virtual neurons / Ascoli G.A., Krichmar J.L., Scorcioni R., Nasuto S.J., Senft S.L. // *Journal of Anatomy and Embryology*. – 2001. – Vol. 204. – P. 283–301.
16. **Ascoli G.A.** Progress and perspectives in computational neuroanatomy // *The Anatomical Record*. – 1999. – V. 257. – P. 195–207.

17. **Ascoli G.A.** Generation, description and storage of dendritic morphology data / Ascoli G.A., Krichmar J.L., Nasuto S.J., Senft S.L. // Philosophical Transactions of the Royal Society B: Biological Sciences. – 2001. – V. 356. – P. 1131–1145.
18. **Турчин В.М.** Імовірнісна модель та алгоритм побудови дендрита нейрона / В.М. Турчин, Я.С. Бондаренко // Питання прикладної математики і математичного моделювання. – Дніпропетровськ, – 2004. – С. 187–199.
19. **Бондаренко Я.С.** Алгоритм імітаційної моделі росту дендрита з піддеревами / Я.С. Бондаренко // Вісник Запорізького національного університету. Фізико-математичні науки. Біологічні науки. – 2005. – №1. – С. 5–10.
20. **Бондаренко Я.С.** Імітаційна ймовірнісна модель дендрита нейрона / Я.С. Бондаренко // Математичні машини і системи. – Київ, – 2006. – №1. – С. 13–27.
21. **Бондаренко Я.С.** Статистичне моделювання дендритів нейронів: дис... кандидата фіз.-мат. наук: 01.05.02 / Бондаренко Я.С. – Дніпропетровськ. – 2006. – 198 С.
22. **Бондаренко Я.С.** Статистичне моделювання дендритів нейронів / Я.С. Бондаренко, В.М. Турчин // Математичні машини і системи. – Київ, – 2011. – №1. – С. 29–38.

Надійшла до редколегії 15.04.11

УДК 519.2

4Я.С. Бондаренко, Т.В. Барабанова

Дніпропетровський національний університет імені Олеся Гончара

СТАТИСТИЧНИЙ КРИТЕРІЙ ДИФЕРЕНЦІАЦІЇ РЕЗУЛЬТАТИВНОГО ГРАВЦЯ

Запропоновано універсальний статистичний критерій диференціації результативного гравця за числом «успішних» дій протягом гри. Критерієм можуть користуватися співробітники відділу кадрів при атестації робітників деякого підприємства, при прийнятті робітників на роботу, лікарі клініки при діагностиці деякого захворювання, менеджери при закупівлі партії виробів, дослідники при контролі партії медичного препарату на токсичність, а також математики, які використовують статистичні методи на практиці.

Предложен универсальный статистический критерий дифференциации результативного игрока по числу «успешных» действий в течение игры. Критерием могут пользоваться сотрудники отдела кадров при аттестации работников некоторого предприятия, при принятии на работу, врачи клиники при диагностике некоторого заболевания, менеджеры при покупке партии изделий, исследователи при контроле партии медицинского препарата на токсичность, а также математики, которые используют статистические методы на практике.

In work the universal statistical criterion of differentiation of the productive player on number "of «successful" actions during game is offered. The criterion can be used by the employees of a staff department at certification of the workers of some enterprise and at acceptance on work, by the doctors of clinic at diagnostics of some disease, by the managers at purchase of a party of products, by the researchers at the control of a party of a medical preparation on toxins, and also mathematics, which use statistical methods in practice.

Ключові слова: баскетбол, дія, гравець, випадкова величина, нульова гіпотеза, альтернативна гіпотеза, критична множина, помилка I роду, помилка II роду, потужність критерію.

Вступ. Задачі математичної статистики природно виникають у різноманітних сферах діяльності людини: в наукових дослідженнях з фізики, біології, астрономії; у психології, медицині, зокрема в спорті.

Баскетбол – це командна гра. Будь який тренер сам створює собі команду, відбираючи й підбираючи гравців один для одного. Відбір – обов'язкова складова частина тренувального процесу, і якщо тренер завзято й професійно проводить роботу з відбору спортсменів, то його навіть менш талановиті вихованці можуть досягти відмінних результатів.

Саме проблема відбору ставиться до найважливішої проблеми сучасного баскетболу. Саме вона змушує багатьох фахівців поєднуватися й працювати спільно. Колективна робота таких фахівців як лікарі, тренери, психологи, фізіотерапевти має на увазі тісне співробітництво на всіх етапах роботи. Але все ж таки провідна роль належить тренерові.

Активність і ефективність дій спортсмена оцінюється за кількістю виконаних дій і процентному співвідношенню успішно й невдало виконаних дій. З метою визначення ефективності ігрових дій спортсмена в баскетболі, фахівцями використовується коефіцієнт корисної гри (ККГ), який дорівнює відношенню отриманого гравцем результату до здійснених ним витрат. На сьогоднішній день найпоширенішими методами визначення коефіцієнта корисної гри є

1) варіант запропонований російськими фахівцями [1]:

$$КПИ = (O + AB + 1,4 \cdot PX + 1,2 \cdot BШ + 1,2 \cdot СЩ + 1,4 \cdot ЧЩ + 0,5 \cdot FC - \Phi - (2\text{очк. мимо}) - 1,5 \cdot (3\text{очк. мимо}) - 0,8 \cdot (\text{штраф. мимо}) - 1,4 \cdot ПП - ПТ) / СВ$$

де КПИ – коефіцієнт «полезности» игрока; О – очки; АВ – атакующие передачи; ПХ – перехваты; БШ – блокшоты; СЩ – подборы на своем щите; ЧЩ – подборы на щите соперника; FC – фолы команды соперников; Φ – фолы; ПП – потери при передаче; ПТ – потери технические; СВ – сыгранное время.

2) варіант запропонований європейськими фахівцями [2]:

$$INDEX = (POINTS - MISSED SHOTS + REBOUNDS + ASSISTS + STEALS + BLOCKS - T / OVERS) / TIME PLAYED(7 MINS MIN PLAY)$$

де INDEX – індекс ефективності змагальної діяльності гравця; POINTS – очки; MISSED SHOTS – неточні кидки; REBOUNDS – підбори м'яча; ASSISTS – результативні передачі; STEALS – перехвати; BLOCKS – блокшоти; T/OVERS – помилки (втрати м'яча); TIME PLAYED – час перебування гравця на майданчику (мінімум – 7 хвилин);

3) варіант запропонований сучасними фахівцями з баскетболу Л.В. Костіковою, Ю.М. Портновим [3]:

$$KKГ = ((S^+) - (S^-)) / T,$$

$$(S^+) = P_{K1} \cdot K1 \cdot K_{M1} + P_{K2} \cdot K2 \cdot K_{M2} + P_{K3} \cdot K3 \cdot K_{M3} + \\ + 2,5C + 1,5Ч + 1,3П, \quad (S^-) = K \cdot T / 200 + 1,5П_n,$$

де $KKГ$ – коефіцієнт корисної гри, (S^+) – сума набраних «позитивних» очок, (S^-) – сума «негативних» очок за помилки в нападі та захисті,

P_{K1} – результативність гравця, набрана 1-очковими кидками,

$K1$ – процентне значення влучень 1-очкових кидків, K_{M1} – процентне значення влучень 1-очкових кидків за модельними даними,

P_{K2} – результативність гравця, набрана 2-очковими кидками,

$K2$ – процентне значення влучень 2-очкових кидків, K_{M2} – процентне значення влучень 2-очкових кидків за модельними даними,

P_{K3} – результативність гравця, набрана 3-очковими кидками,

$K3$ – процентне значення влучень 3-очкових кидків, K_{M3} – процентне значення влучень 3-очкових кидків за модельними даними, $П$ – кількість перехоплень м'яча, $С$ – кількість узятих відскоків біля свого щита,

$Ч$ – кількість узятих відскоків біля щита суперника, $П_n$ – кількість втрат м'яча, $К$ – кількість очок, отриманих командою в грі, $Т$ – час участі в грі.

Якщо $KKГ$ гравця віднести до середнього значення $KKГ$ усіх гравців команди, то дістанемо відносний $KKГ$ гравця. Якщо відносний $KKГ$ гравця більше 1, то гравець грає краще гравця середнього рівня. Якщо ж відносний $KKГ$ гравця менше 1, то гравець грає нижче середнього рівня.

Відмітимо, що $KKГ$ показує рівень гри кожного спортсмена незалежно від сили суперника. Така оцінка стимулює завжди грати з повною віддачею і попереджає тенденцію розподілу суперників на слабких та сильних.

У представленій роботі запропоновано принципово новий теоретико-ймовірнісний підхід до визначення ефективності гри кожного спортсмена.

Результативність гравця характеризується низкою числових характеристик: числом «успішних» 2-очкових кидків, числом «успішних» 3-очкових кидків, числом «успішних» штрафних кидків, числом атакуючих передач, що закінчилися влученням, числом перехоплень м'яча, числом блокшотів, числом підборів м'яча на своєму та чужому щиті. Усі ці характеристики є випадковими величинами (зазначимо, що випадкова величина повністю описується своїм розподілом). Тому постановка задачі диференціації результативного гравця як задачі математичної статистики є природною, але до сьогодні формулювання та розв'язання такої задачі в термінах перевірки статистичних гіпотез не відбувалося.

Постановка задачі диференціації результативного гравця. Перед початком сезону тренер деякої баскетбольної команди проводить селекційну роботу, шукаючи сильних гравців для своєї команди та спілкуючись з їхніми агентами.

Один з агентів запропонував тренеру свого підопічного, стверджуючи, що той сильний та результативний гравець, спроможний допомогти команді покращити рівень гри. Перед прийняттям рішення про купівлю гравця, тренеру необхідно перевірити надану інформацію і впевнитися в ефективності гравця. Результативність гравця характеризується числом «успішних» дій, які виконує гравець протягом гри.

Агент гравця стверджує, що частка «успішних» дій гравця становить p_0 і хоче, щоб ці здібності гравця проявлялися з імовірністю не меншою 0,95 (прибуток агента залежить від результативності гравця).

Придбання нового гравця потребує певних фінансових витрат й тренер, природно, відмовляється купувати гравця, якщо частка його «успішних» дій протягом гри становить p_1 ($p_1 < p_0$), наполягаючи, щоб при перегляді такого гравця останнє виявлялося з імовірністю не меншою 0,9.

Необхідно запропонувати статистичний критерій диференціації результативного гравця за числом «успішних» дій протягом гри, при цьому забезпечити виконання перелічених вище умов (під «успішною» дією будемо розуміти влучення; атакуючу передачу, що закінчилась влученням; перехоплення м'яча; блокшот; підбір м'яча на своєму та чужому щиті).

Розподіл числа «успішних» дій протягом гри. Число ξ «успішних» дій гравця протягом матчу складається з числа «успіхів» при кожній дії, тобто

$$\xi = \xi_1 + \xi_2 + \dots + \xi_v,$$

де ξ_k – число «успіхів» при k -тій дії (дорівнює 0 або 1), $k = 1, 2, \dots, v$,
 v – число дій гравця протягом матчу. Випадкові величини ξ_k незалежні й однаково розподілені: кожна з розподілом

$$P\{\xi_k = 0\} = 1 - q, \quad P\{\xi_k = 1\} = q.$$

Генератриса випадкової величини ξ_k дорівнює

$$g(z) = Mz^{\xi_k} = 1 - q + qz.$$

Нехай число дій гравця протягом матчу v має біноміальний розподіл з параметрами (m, u) . Генератриса випадкової величини v дорівнює

$$\pi(z) = Mz^v = \sum_{i=0}^m C_m^i u^i (1-u)^{m-i} z^i = (1-u + uz)^m.$$

Тоді генератриса числа ξ «успішних» дій гравця протягом матчу набуває вигляду

$$\begin{aligned} G(z) &= Mz^\xi = \sum_{i=0}^m M(z^\xi, \{v=i\}) = \sum_{i=0}^m Mz^\xi I\{v=i\} = \\ &= \sum_{i=0}^m Mz^\xi MI\{v=i\} = \sum_{i=0}^m Mz^\xi P\{v=i\} = \sum_{i=0}^m \left(M \prod_{j=1}^i z^{\xi_j} \right) P\{v=i\} = \\ &= \sum_{i=0}^m \left(\prod_{j=1}^i Mz^{\xi_j} \right) P\{v=i\} = \sum_{i=0}^m (g(z))^i P\{v=i\} = \pi(g(z)). \end{aligned}$$

Отже,

$$G(z) = \pi(g(z)) = (1 - qu + quz)^m.$$

Оскільки розподіл випадкової величини однозначно відновляється за генератрисою, то число ξ «успішних» дій гравця протягом гри має біноміальний розподіл з параметрами (m, qu) .

Критерій диференціації результативного гравця за числом «успішних» дій протягом гри. Відносно ефективності гри гравця висуваємо дві гіпотези: H_1 – гравець результативний, H_2 – гравець не результативний. Розглянемо наслідки помилки, яка полягає у відхиленні гіпотези, коли вона справедлива.

Помилка, яка полягає у відхиленні H_1 , коли вона справедлива, призводить до відрахування з команди сильного гравця, що знижує загальний командний результат. Помилка, яка полягає у відхиленні H_2 , коли вона справедлива, призводить до втрати можливості підвищити командний результат, не відрахувавши слабого гравця. Важливіше уникнути першої помилки, тому за нульову гіпотезу розглянемо H_0 : гравець результативний. Альтернативна гіпотеза H_1 : гравець не результативний.

Стохастичний експеримент полягає в аналізі v ($v = n$) дій гравця протягом одного матчу й спостереженні випадкової величини ξ – числа «успішних» дій гравця (ξ може набувати значень $0, 1, 2, \dots, n$).

Якщо гіпотеза H_0 справедлива, то ξ має розподіл

$$P_0\{\xi = k\} = C_n^k p_0^k (1 - p_0)^{n-k}, \quad k = 0, 1, \dots, n, \quad (1)$$

де p_0 – ймовірність «успішної» дії гравця за справедливої H_0 .

Якщо гіпотеза H_1 справедлива, то ξ має розподіл

$$P_1\{\xi = k\} = C_n^k p_1^k (1 - p_1)^{n-k}, \quad k = 0, 1, \dots, n, \quad (2)$$

де p_1 – ймовірність «успішної» дії гравця за справедливої H_1 . Очевидно, що $p_0 \gg p_1$.

Природно вважати, що гравець результативний, якщо $\xi(\omega)$ набуває «великих» значень, і не результативний, якщо $\xi(\omega)$ набуває «малих» значень. Отже, критична множина має вигляд

$$S = \{x : 0 \leq x \leq l\}. \quad (3)$$

Дійсно, якщо $\xi(\omega) \in S$, то гіпотезу H_0 відхиляємо, а якщо $\xi(\omega) \notin S$, то гіпотезу H_0 не відхиляємо. Число дій n та l повинні бути такими, щоб виконувались умови, висунуті тренером та агентом, а саме

Таблиця 1

Імовірності помилок першого та другого роду

Гіпотеза		Справедлива	
		H_0	H_1
Приймається	H_0	$P(\bar{S} H_0) \geq 0,95$	$P(\bar{S} H_1) \leq 0,1$
	H_1	$P(S H_0) \leq 0,05$	$P(S H_1) \geq 0,9$

де $P(S | H_0)$ – імовірність відхилення H_0 , коли H_0 – справедлива, $P(S | H_1)$ – імовірність відхилення H_0 , коли H_0 – не справедлива, $P(\bar{S} | H_0)$ – імовірність не відхилення H_0 , коли H_0 – справедлива, $P(\bar{S} | H_1)$ – імовірність не відхилення H_0 , коли H_0 – не справедлива.

Шукані n та l повинні задовольняти системі нерівностей

$$\begin{cases} P(S | H_0) \leq 0,05; \\ P(S | H_1) \geq 0,9, \end{cases}$$

або,

$$\begin{cases} \sum_{k=0}^l C_n^k p_0^k (1-p_0)^{n-k} \leq 0,05; \\ \sum_{k=0}^l C_n^k p_1^k (1-p_1)^{n-k} \geq 0,9. \end{cases} \quad (4)$$

Підставляючи до системи n, p_0, p_1 знаходимо l . Якщо відоме нам число «успішних» дій $\xi(\omega)$ менше або дорівнює знайденому l , то гіпотезу

H_0 відхиляємо. Якщо ж відоме нам число «успішних» дій $\xi(\omega)$ більше 1, то гіпотезу H_0 не відхиляємо.

Імовірність помилки I роду дорівнюватиме

$$P(S | H_0) = \sum_{k=0}^l C_n^k p_0^k (1 - p_0)^{n-k}.$$

Імовірність помилки II роду дорівнюватиме

$$P(\bar{S} | H_1) = 1 - P(S | H_1) = 1 - \sum_{k=0}^l C_n^k p_1^k (1 - p_1)^{n-k}.$$

Потужність критерію дорівнюватиме

$$P(S | H_1) = \sum_{k=0}^l C_n^k p_1^k (1 - p_1)^{n-k}.$$

Дослідження імовірнісних характеристик гравця залежно від його амплуа. У процесі розвитку баскетболу склалося, що у складі кожної команди є п'ять унікальних ігрових амплуа, кожне з яких має свій набір функцій та задач. Амплуа (від франц. *emploi* – роль) – спеціалізація гравця при виконанні певних функцій та задач на гральному майданчику. Амплуа гравців розподіляються так.

Перше амплуа – розігруючі. Це гравці, що починають майже всі атакуючі комбінації своєї команди. Основна функція розігруючих – виведення своїх партнерів на вигідну позицію, роздача результативних передач, або ж «бомбардування» кільця суперника з дальньої дистанції – з «периметра».

Друге амплуа – атакуючі захисники. Це гравці, які орієнтовані на стрімкий прохід до кошика суперника, скидання м'яча на «периметр», а також завершення «швидких відривів». Їх функції зосереджуються на допомозі розігруючому і «бомбардуванні» кільця здалеку.

Третє амплуа – легкі форварди. Це гравці, які є зв'язковою ланкою між розігруючими та «великими» гравцями. Основна функція – набирання очок, надважлива риса – добрий дриблінг, точний середній кидок, а також достатня швидкість для того, щоб встигати на підбирання м'яча.

Четверте амплуа – важкі форварди. Це гравці, основною функцією яких є захисні дії, підбирання м'яча після відскоку від свого чи чужого кільця, блокування кидків гравців суперника. Обов'язковим є вміння стабільно закидати з середньої дистанції.

П'яте амплуа – центрові. Це найкремезніші гравці, які повинні володіти достатньою швидкістю, щоб за найкоротший час переміщатися від свого щита до чужого і навпаки. Функція центрових – завершення атак з під кільця, підбирання м'яча та блокування кидків суперника, постановка «заслонів».

Спочатку застосуємо запропонований критерій до гіпотетичного гравця з ліги NBA (як відомо, до цієї ліги відбирають найкращих гравців), а потім до гравців однієї з українських баскетбольних команд.

Нехай для сильного гравця ймовірність «успішної» дії дорівнює $p_0 = 0,8$ за справедливої гіпотези H_0 : *гравець результативний*, і дорівнює $p_1 = 0,5$ за справедливої гіпотези H_1 : *гравець не результативний*.

На рисунку 1А по вісі абсцис відкладаємо число дій гравця n протягом матчу, по вісі ординат – число «успішних» дій $\xi(\omega)$ (точки з координатами $(n, \xi(\omega))$ позначено ромбами). Розв'язуючи систему (4) для кожного фіксованого n знаходимо число l – верхню границю критичної множини S (точки з координатами (n, l) позначено трикутниками).

Для точок, що лежать вище верхньої границі критичної множини S , гіпотеза H_0 : *гравець результативний* не відхиляється, а для точок, що лежать нижче верхньої границі критичної множини S , гіпотеза H_0 відхиляється. Оскільки число «успішних» дій гравця з ліги NBA $\xi(\omega)$ набуває «великих» значень ($\xi(\omega) \notin S$), то гіпотеза H_0 не відхиляється. Отже, згідно з запропонованим критерієм гіпотетичний гравець з ліги NBA є сильним гравцем.

Вихідними даними для апробації критерію були характеристики «успішних» дій гравців однієї з українських баскетбольних команд. Проілюструємо роботу критерію для визначення результативності типового гравця з розігруючих.

Виберемо для сильного гравця ймовірність «успішної» дії $p_0 = 0,8$ за справедливої гіпотези H_0 : *гравець результативний (сильний)*, і $p_1 = 0,5$ за справедливої гіпотези H_1 : *гравець не результативний*.

На рисунку 1Б по вісі абсцис відкладаємо число дій гравця n протягом матчу, по вісі ординат – число «успішних» дій $\xi(\omega)$ (точки з координатами $(n, \xi(\omega))$ позначено ромбами). Розв'язуючи систему (4) для кожного фіксованого n знаходимо число l – верхню границю критичної множини S (точки з координатами (n, l) позначено трикутниками).

Для точок, що лежать вище верхньої границі критичної множини S , гіпотеза H_0 : *гравець результативний (сильний)* не відхиляється, а для точок, що лежать нижче верхньої границі критичної множини S , гіпотеза H_0 відхиляється. Оскільки число «успішних» дій гравця з ліги NBA $\xi(\omega)$ набуває «малих» значень, то гіпотеза H_0 відхиляється. Отже, згідно з запропонованим критерієм типовий гравець з розігруючих не є сильним гравцем.

Виберемо для середнього за силою гравця ймовірність «успішної» дії $p_0 = 0,5$ за справедливої гіпотези H_0 : *гравець результативний (середній за силою)*, і $p_1 = 0,2$ за справедливої гіпотези H_1 : *гравець не результативний (слабкий)*.

На рисунку 1Б по вісі абсцис відкладаємо число дій гравця n протягом матчу, по вісі ординат – число «успішних» дій $\xi(\omega)$ (точки з координатами $(n, \xi(\omega))$ позначено ромбами). Розв'язуючи систему (4) для кожного фіксованого n знаходимо число m – верхню границю критичної множини S (точки з координатами (n, m) позначено квадратами).

Для точок, що лежать вище верхньої границі критичної множини S , гіпотеза H_0 : *гравець результативний (середній за силою)* не відхиляється, а для точок, що лежать нижче верхньої границі критичної множини S , гіпотеза H_0 відхиляється. Оскільки число «успішних» дій гравця з ліги NBA $\xi(\omega)$ набуває «великих» значень, то гіпотеза H_0 не відхиляється. Отже, згідно з запропонованим критерієм типовий гравець з розігруючих є середнім за силою гравцем.

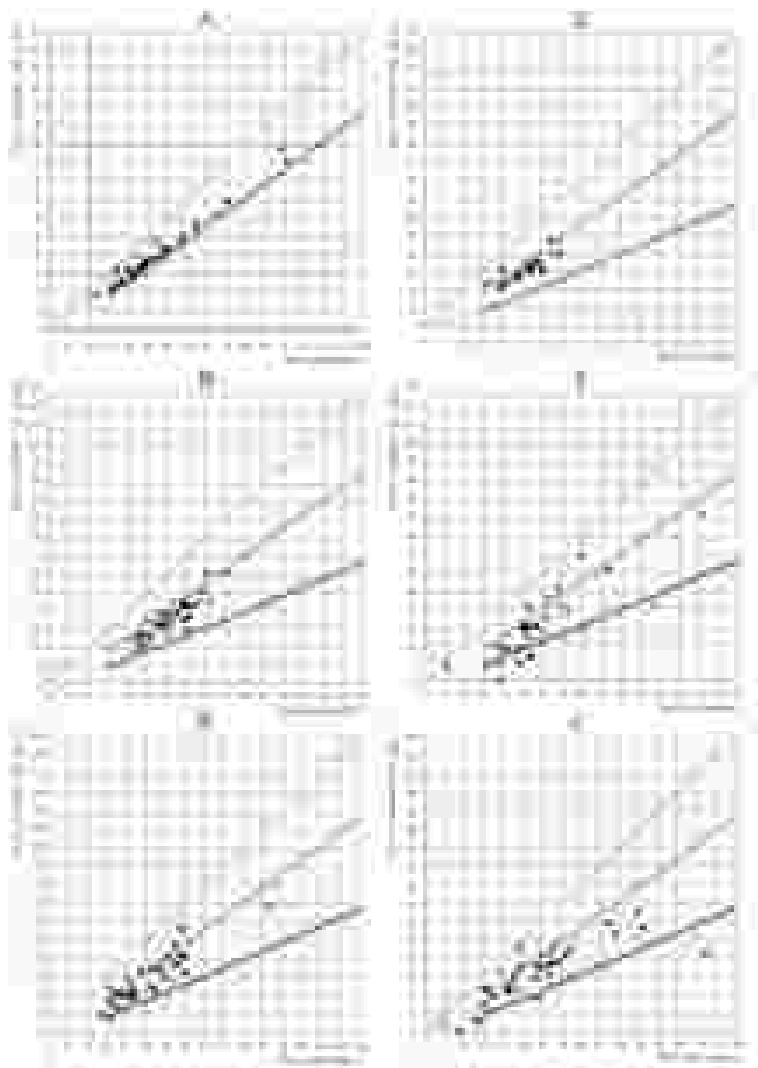


Рис. 1. Зміна верхньої границі l критичної множини S зі змінюю числа дій n . Точки з координатами (n, l) позначено трикутниками при $p_0 = 0,8, p_1 = 0,5$, і квадратами при $p_0 = 0,5, p_1 = 0,2$

Методика диференціації результативних гравців серед атакуючих захисників, легких форвардів, важких форвардів та центрових аналогічна.

Згідно з запропонованим критерієм типовий гравець з атакуючих захисників є середнім за силою (рис. 1В), типовий гравець з легких форвардів є слабким (рис. 1Г), типовий гравець з важких форвардів є середнім за силою (рис. 1Д), типовий гравець з центрових є слабким (рис. 1Є). У команді не виявлено виражених сильних гравців. У цілому – команда складається з гравців середнього рівня.

Висновки. Запропонований критерій значно підвищує використання статистичної інформації про гравців тренером команди будь-якого ігрового виду спорту (баскетболу, волейболу, гандболу, футболу та ін.), а саме дозволяє відбирати до команди результативних (сильних та середніх за силою) гравців, ефективно комплектувати команду на змаганнях, обгрунтовано проводити заміни гравців.

Критерій диференціації результативного гравця, запропонований у роботі, є універсальним. Критерієм можуть користуватися співробітники відділу кадрів при атестації робітників деякого підприємства, при прийнятті робітників на роботу (гравець – робітник, «успішна» дія – правильно виконане завдання), лікарі клініки при діагностиці деякого захворювання (гравець – хворий, «успішна» дія – позитивний аналіз), менеджери при закупівлі партії виробів (гравець – партія, «успішна» дія – виріб бракований), дослідники при контролі партії медичного препарату на токсичність (гравець – партія, «успішна» дія – летальний випадок піддослідної тварини), а також математики, які використовують статистичні методи на практиці.

Бібліографічні посилання

1. www.infobasket.ru
2. www.eurowomen2003.com
3. Баскетбол: учеб. для ин-тов физ. культуры. Под редакцией Портнова Ю.М. – М., 1997.
4. Турчин В.Н. Теория вероятностей и математическая статистика : учеб. / В.Н. Турчин. – Д., 2008. – 656 с.

Надійшла до редколегії 18.05.11

⁵Л.Л. Гарт

Днепропетровский национальный университет имени Олеся Гончара

О ЧИСЛЕННОМ МОДЕЛИРОВАНИИ РЕШЕНИЯ НЕЛИНЕЙНОГО ПАРАМЕТРИЧЕСКОГО УРАВНЕНИЯ ПРОЕКЦИОННО-ИТЕРАЦИОННЫМ МЕТОДОМ

Досліджується питання про існування, область розташування та наближення розв'язку нелінійного операторного рівняння, що залежить від числового параметра, у банаховому просторі за допомогою проекційно-ітераційного методу, оснований на методі Ньютона-Канторовича. Доведено відповідні теореми, отримано оцінки похибки.

Исследуется вопрос о существовании, области расположения и приближении решения нелинейного операторного уравнения, зависящего от числового параметра, в банаховом пространстве при помощи проекционно-итерационного метода, основанного на методе Ньютона-Канторовича. Доказаны соответствующие теоремы, получены оценки погрешности.

The problem of existence, a location domain and approximation of solution of a nonlinear operatoral equation with a numerical parameter in Banach space is investigated with applying the projection-iteration method based on Newton-Kantorovich method. Corresponding theorems are proved, error estimates are obtained.

Ключевые слова: нелинейное операторное уравнение, параметр, решение, метод Ньютона-Канторовича, проекционно-итерационный метод, сходимость, последовательность, приближение, погрешность.

Введение. Метод решения функциональных уравнений, известный под названием метода Ньютона-Канторовича (метода Ньютона или касательных – в случае вещественных уравнений), а также некоторые его модификации являются в настоящее время одними из немногих, применяемых на практике для фактического нахождения решения нелинейного функционального уравнения. Важно отметить и теоретическое значение метода, так как с его помощью можно делать заключение о существовании, единственности и области расположения решения уравнения, не находя самого решения, что подчас не менее важно, чем фактическое знание решения.

Исследованию метода Ньютона Канторовича и его модификаций посвящены работы Л.В. Канторовича [1, 2], И.П. Мысовских [3], Б.А. Вертгейма, С. Фенга и других авторов. Для решения нелинейных функциональных уравнений применяются также линеаризованные одношаговые методы скорейшего спуска, с минимальными невязками, с минимальными ошибками [4], методы касательных парабол и касательных гипербол [5], методы типа Чебышева и др., относящиеся к классу итерационных методов, а также группа проекционных (аппроксимационных) методов.

Впервые общая теория проекционных методов для решения операторных уравнений второго рода в банаховых пространствах была построена Л.В. Канторовичем в [1]. Общая теория проекционных методов решения уравнения первого рода

$$Ax = f \quad (1)$$

с нелинейным оператором A , действующим из нормированного пространства X в нормированное пространство Y , изложена в работах В.В. Петришина [6] и Ф. Браудера. В работе С.Д. Балашовой [7] проведены исследования методов, получивших название проекционно-итерационных, для решения линейных и нелинейных уравнений первого рода, основанных на следующей идее. Исходное операторное уравнение вида (1) аппроксимируется последовательностью приближенных уравнений:

$$A_n x_n = f_n, \quad n=1,2,\dots, \quad (2)$$

где каждый из операторов A_n переводит X_n в Y_n , $\{X_n\}$ и $\{Y_n\}$ – возрастающие последовательности подпространств X и Y соответственно. К решению этих уравнений применяется некоторый итерационный метод, причем для каждого из приближенных уравнений находится по указанному методу лишь несколько приближений $x_n^{(k)}$ ($k=1,2,\dots,k_n$), последнее из которых полагается равным начальному приближению в итерационном процессе для следующего, $(n+1)$ -го приближенного уравнения. В качестве последовательности приближений к решению исходного уравнения принимается последовательность $\{x_n^{(k_n)}\}$. Такой подход к нахождению приближенного решения того или иного уравнения, естественно, устраняет трудности, возникающие при решении того же уравнения обычным проекционным методом и облегчает выбор подходящего начального приближения (особенно в нелинейном случае) по сравнению с решением исходного уравнения итерационным методом.

Постановка задачи. Пусть в уравнении (1) A – дифференцируемый по Фреше нелинейный оператор, переводящий открытое множество Ω банахова пространства X в банахово пространство Y . Предположим, что в Ω содержится решение x^* уравнения (1).

Возьмем произвольный элемент $x^{(0)} \in \Omega$. Предполагая, что производная Фреше $A'(x)$ непрерывна в замкнутом шаре $S(x^{(0)}, R) \subset \Omega$, из формулы конечных приращений получим, что в достаточно малой окрестности $x^{(0)}$ элемент Ax достаточно близок к $Ax^{(0)} + A'(x^{(0)})(x - x^{(0)})$, и, следовательно, решение уравнения $Ax^{(0)} + A'(x^{(0)})(x - x^{(0)}) = f$ достаточно близко к x^* . Это уравнение линейно, и если существует $[A'(x^{(0)})]^{-1}$, то его решение $x^{(1)}$ находится по формуле $x^{(1)} = x^{(0)} - [A'(x^{(0)})]^{-1}(Ax^{(0)} - f)$. Продолжая этот процесс, получим последовательность приближений

$$x^{(k+1)} = x^{(k)} - [A'(x^{(k)})]^{-1}(Ax^{(k)} - f), \quad k = 0, 1, \dots \quad (3)$$

к решению x^* уравнения (1), определяющую итерационный метод Ньютона-Канторовича [2].

Теорема о сходимости метода Ньютона-Канторовича при условиях типа Коши, когда оценка нормы оператора $\Gamma(x) = [A'(x)]^{-1}$ известна не только в точке $x^{(0)}$, но и во всем шаре $S(x^{(0)}, R)$ (что позволяет ослабить ограничения, накладываемые на выбор $x^{(0)}$), доказана в [3]. Обобщением упомянутой теоремы, когда вместо оператора $\Gamma(x) = [A'(x)]^{-1}$ требуется лишь существование в $S(x^{(0)}, R)$ оператора $D(x)$, близкого к $\Gamma(x)$, является следующая теорема.

Теорема 1. Пусть оператор A дифференцируем по Фреше в некотором шаре $S(x^{(0)}, R)$ и его производная на этом шаре удовлетворяет условию Липшица

$$\|A'(x) - A'(y)\| \leq L \|x - y\|, \quad L = \text{const} > 0. \quad (4)$$

Пусть существует линейный оператор $D(x)$, отображающий Y в X и имеющий непрерывный обратный, причем для всех $x \in S(x^{(0)}, R)$

$$\|D(x)\| \leq b, \quad (5)$$

$$\|D(x)A'(x) - I\| \leq \delta < 1. \quad (6)$$

Если начальное приближение $x^{(0)}$ удовлетворяет условиям

$$\|Ax^{(0)} - f\| \leq \eta^{(0)}, \quad (7)$$

$$h^{(0)} = \frac{b^2 \eta^{(0)} L}{(1-\delta)^2} < 2, \quad r = \frac{b \eta^{(0)} G}{1-\delta} \leq R, \quad G = \sum_{i=0}^{\infty} \left(\frac{h^{(0)}}{2} \right)^{2^i - 1}, \quad (8)$$

то уравнение (1) имеет в шаре $S(x^{(0)}, r)$ решение x^* , к которому сходится процесс Ньютона-Канторовича (3) с оценкой погрешности

$$\|x^{(k)} - x^*\| \leq \frac{b \eta^{(0)}}{1-\delta} G \left(\frac{h^{(0)}}{2} \right)^{2^k - 1}.$$

Доказательство теоремы 1 можно найти в [8]. Там же сформулированы условия, на основании которых можно судить о разрешимости уравнения (1) по решению уравнения

$$Ax \equiv Tx + Qx = f,$$

близкого к нему, но уже более простого (также, вообще говоря, нелинейного), где T и Q – некоторые дифференцируемые по Фреше операторы, действующие из X в Y .

Такая ситуация является частным случаем более общей, когда левая часть уравнения (1) зависит от числового или иного параметра, причем при одном значении параметра решение уравнения известно и требуется установить существование решения для значений параметра, близких к начальному. Предположим, что упомянутый параметр входит в уравнение линейно, то есть будем рассматривать операторное уравнение вида

$$Ax(\mu) \equiv Tx + \mu Qx = f, \quad (9)$$

где μ означает линейный оператор в пространстве Y , в частности, μ может быть числовым множителем.

Метод решения. Условия существования и область расположения решения $x^* \equiv x^*(\mu)$ уравнения (9), а также условия сходимости к нему итерационной последовательности метода Ньютона-Канторовича (3) устанавливает следующая теорема.

Теорема 2. Пусть операторы T и Q дифференцируемы по Фреше в некотором шаре $S(x^{(0)}, R) \subset X$ и их производные удовлетворяют на этом шаре условию Липшица с константами K и M соответственно. Пусть $x^{(0)}$ является решением уравнения $Tx = f$, так что

$$Ax^{(0)}(0) \equiv Tx^{(0)} = f \quad (10)$$

и для каждого $x \in S(x^{(0)}, R)$ существует непрерывный линейный оператор $[T'(x)]^{-1}$, причем

$$\|[T'(x)]^{-1}\| \leq b, \quad (11)$$

$$\|Q'(x)\| \leq v \quad (12)$$

для всех $x \in S(x^{(0)}, R)$. Тогда, если начальное приближение $x^{(0)}$ удовлетворяет условию

$$\|Qx^{(0)}\| \leq \zeta^{(0)} \quad (13)$$

и μ таково, что

$$|\mu|bv < 1, \quad h^{(0)} = \frac{b^2|\mu|\zeta^{(0)}(K + |\mu|M)}{(1 - |\mu|bv)^2} < 2, \quad r = \frac{b|\mu|\zeta^{(0)}G}{1 - |\mu|bv} \leq R, \quad (14)$$

где G определено в (8), то уравнение (9) имеет в шаре $S(x^{(0)}, r)$ решение x^* , к которому сходится итерационный процесс (3) с оценкой погрешности

$$\|x^{(k)} - x^*\| \leq \frac{b|\mu|\zeta^{(0)}G}{1 - |\mu|bv} \left(\frac{h^{(0)}}{2} \right)^{2^k - 1}.$$

Доказательство. Утверждение теоремы 2 получается непосредственно из теоремы 1, если положить в ней $D(x) = [T'(x)]^{-1}$. Проверим выполнимость условий теоремы 1.

По свойствам производных [2], из дифференцируемости операторов T и Q в шаре $S(x^{(0)}, R)$ следует дифференцируемость оператора $A = T + Q$ в том же шаре. Для произвольных $x, y \in S(x^{(0)}, R)$ будем иметь:

$$\|A'(x) - A'(y)\| = \|T'(x) + \mu Q'(x) - T'(y) - \mu Q'(y)\| \leq$$

$$\leq \|T'(x) - T'(y)\| + |\mu| \|Q'(x) - Q'(y)\| \leq (K + |\mu| M) \|x - y\|,$$

то есть условие Липшица (4) выполняется с константой $L = K + |\mu| M$.

Далее с учетом (11), (12) и первого из условий (14) получаем

$$\begin{aligned} \|D(x)A'(x) - I\| &= \|[T'(x)]^{-1}(T'(x) + \mu Q'(x)) - I\| = \\ &= |\mu| \|[T'(x)]^{-1}Q'(x)\| \leq |\mu| b\nu < 1 \end{aligned}$$

для всех $x \in S(x^{(0)}, R)$, то есть условие (6) выполняется с константой $\delta = |\mu| b\nu$. Что касается условия (7), то оно, очевидно, выполняется с константой $\eta^{(0)} = |\mu| \zeta^{(0)}$ на основании выполнимости условий (10), (13):

$$\|Ax^{(0)} - f\| \equiv \|Ax^{(0)}(\mu) - f\| = \|Tx^{(0)} + \mu Qx^{(0)} - f\| \leq |\mu| \zeta^{(0)}.$$

Таким образом, все условия теоремы 1 выполнены, из чего следует справедливость утверждения настоящей теоремы.

Теорема доказана.

Проекционно-итерационный метод. Рассмотрим теперь проекционно-итерационный метод, основанный на методе Ньютона – Канторовича, применительно к уравнению (1) с нелинейным оператором A , действующим в банаховом пространстве X , дифференцируемым по Фреше в некотором шаре $S(x_N^{(0)}, R)$ этого пространства.

Согласно общей схеме проекционно-итерационных методов [9] уравнение (1) аппроксимируется последовательностью приближенных уравнений (2), где A_n – нелинейный оператор, действующий в подпространстве X_n исходного пространства ($X_1 \subset X_2 \subset \dots \subset X_n \subset \dots \subset X$), дифференцируемый по Фреше, начиная с некоторого $n \geq N$, на множестве $\Omega_n = X_n \cap S(x_N^{(0)}, R)$; $f_n = P_n f$, P_n – линейный проектор, переводящий X на X_n ($P_n : X \rightarrow X_n$, $P_n x_n = x_n$ для $x_n \in X_n$), причем выполнены условия близости

$$\|A_n x_n - P_n A x_n\| \leq \alpha_n, \quad \|A'_n(x_n) - P_n A'(x_n)\| \leq \alpha'_n, \quad (15)$$

$$\|P_n A x - A x\| \leq \beta_n, \quad \|P_n A'(x) - A'(x)\| \leq \beta'_n, \quad (16)$$

$$\|P_n f - f\| \leq \gamma_n \quad (17)$$

для всех $x_n \in \Omega_n$, $x \in S(x_N^{(0)}, R)$, причем $\alpha_n, \alpha'_n, \beta_n, \beta'_n, \gamma_n \rightarrow 0$ при $n \rightarrow \infty$. Будем считать, что производная $A'_n(x_n)$ на множестве Ω_n удовлетворяет условию Липшица

$$\|A'_n(x_n) - A'_n(y_n)\| \leq L \|x_n - y_n\|, \quad x_n, y_n \in \Omega_n \quad (18)$$

и для всех $x_n \in \Omega_n$ существует непрерывный линейный оператор $\Gamma_n(x_n) = [A'_n(x_n)]^{-1}$, что позволяет применить к каждому из уравнений (2), начиная с номера $n \geq N$, итерационный метод Ньютона-Канторовича и построить проекционно-итерационную последовательность приближений к решению x^* уравнения (1) по формулам:

$$x_n^{(k+1)} = x_n^{(k)} - [A'_n(x_n^{(k)})]^{-1} (A_n x_n^{(k)} - f_n), \quad x_{n+1}^{(0)} = x_n^{(k_n)} \quad (19)$$

$$(k = 0, 1, \dots, k_n - 1; \quad n \geq N; \quad x_N^{(0)} \in \Omega_N \subset X).$$

Теорема о существовании решения x^* уравнения (1), области его расположения, а также сходимости процесса (19) содержится в [10].

Приведем здесь обобщение упомянутой теоремы, когда вместо обратных операторов $\Gamma(x) = [A'(x)]^{-1}$, $x \in S(x_N^{(0)}, R)$ и $\Gamma_n(x_n) = [A'_n(x_n)]^{-1}$, $x_n \in \Omega_n$ требуется существование в указанных областях лишь операторов $D(x)$ и $D_n(x_n)$, близких к $\Gamma(x)$ и $\Gamma_n(x_n)$ соответственно.

Теорема 3. Пусть оператор A дифференцируем по Фреше в некотором шаре $S(x_N^{(0)}, R) \subset X$, а оператор A_n для $n \geq N$ дифференцируем на множестве $\Omega_n = X_n \cap S(x_N^{(0)}, R)$ и его производная $A'_n(x_n)$ удовлетворяет на Ω_n условию Липшица (18). Пусть выполняются условия близости (15)-(17), а также существуют линейные непрерывно обратимые операторы $D(x)$, обладающий свойствами (5), (6) для всех $x \in S(x_N^{(0)}, R)$, и $D_n(x_n)$ такие, что

$$\|D_n(x_n) A'_n(x_n) - I\| \leq \delta_n < 1 \quad (20)$$

для всех $x_n \in \Omega_n$, $n \geq N$. Если начальное приближение $x_N^{(0)} \in \Omega_N$ удовлетворяет условиям

$$\|A_N x_N^{(0)} - f_N\| \leq \eta_N^{(0)}, \quad (21)$$

$$h_N^{(0)} = b_N^2 \eta_N^{(0)} L < 2, \quad r_N = b_N \eta_N^{(0)} H_N \leq R, \quad b_N = \frac{b}{1 - b(\alpha'_N + \beta'_N) - \delta},$$

где

$$H_N = G_N + \sum_{i=N}^{\infty} \left(\frac{h_N^{(0)}}{2} \right)^{2^{Z_i} - 1} < 2 G_N, \quad G_N = \sum_{i=0}^{\infty} \left(\frac{h_N^{(0)}}{2} \right)^{2^i - 1}, \quad (22)$$

$Z_i = \sum_{m=N}^i (k_m - 1)$, то уравнение (1) имеет в шаре $S(x_N^{(0)}, r_N) \subset X$ решение x^* , к которому сходится процесс последовательных приближений $\{x_n^{(k_n)}\}$ определяемый формулами (19), с оценкой погрешности

$$\|x_n^{(k_n)} - x^*\| \leq b_N \eta_N^{(0)} V_n \left(\frac{h_N^{(0)}}{2} \right)^{2^{Z_n} - 1},$$

где

$$V_n = \sum_{i=0}^{\infty} \left(\frac{h_N^{(0)}}{2} \right)^{2^{Z_n} (2^i - 1)} + \sum_{i=n+1}^{\infty} \left(\frac{h_N^{(0)}}{2} \right)^{2^{Z_i} - 2^{Z_n}} < 2 G_N. \quad (23)$$

Доказательство теоремы 3 можно найти в [9].

Рассмотрим теперь следствие из теоремы 3 для случая, когда операторное уравнение (1), заданное в банаховом пространстве X , зависит от числового параметра μ в соответствии с (9), а уравнения (2), заданные в подпространствах $X_n \subset X$, имеют вид

$$A_n x_n(\mu) \equiv T_n x_n + \mu Q_n x_n = f_n. \quad (24)$$

Теорема 4. Пусть операторы T и Q дифференцируемы по Фреше в некотором шаре $S(x_N^{(0)}, R) \subset X$, а операторы T_n и Q_n для $n \geq N$ дифференцируемы на множестве $\Omega_n = X_n \cap S(x_N^{(0)}, R)$ и их производные $T'_n(x_n)$ и $Q'_n(x_n)$ удовлетворяет на этом множестве условию Липшица с константами K и M соответственно. Пусть выполняются условия близости (17),

$$\|T_n x_n - P_n T x_n\| \leq \tau_n, \quad \|Q_n x_n - P_n Q x_n\| \leq \theta_n, \quad (25)$$

$$\|T'_n(x_n) - P_n T'(x_n)\| \leq \tau'_n, \quad \|Q'_n(x_n) - P_n Q'(x_n)\| \leq \theta'_n$$

для всех $x_n \in \Omega_n$, причем $\tau_n, \tau'_n, \theta_n, \theta'_n \rightarrow 0$ при $n \rightarrow \infty$,

$$\begin{aligned} \|P_n T x - T x\| &\leq \chi_n, \quad \|P_n Q x - Q x\| \leq \lambda_n, \\ \|P_n T'(x) - T'(x)\| &\leq \chi'_n, \quad \|P_n Q'(x) - Q'(x)\| \leq \lambda'_n \end{aligned} \quad (26)$$

для всех $x \in S(x_N^{(0)}, R)$, причем $\chi_n, \chi'_n, \lambda_n, \lambda'_n \rightarrow 0$ при $n \rightarrow \infty$; для всех $x \in S(x_N^{(0)}, R)$ существует непрерывный линейный оператор $[T'(x)]^{-1}$, обладающий свойством (11), и выполняется условие (12); для всех $x_n \in \Omega_n$, $n \geq N$ существуют непрерывные линейные операторы $[T'_n(x_n)]^{-1}$ такие, что

$$\|[T'_n(x_n)]^{-1} Q'_n(x_n)\| \leq v_n. \quad (27)$$

Пусть $x_N^{(0)} \in \Omega_N$ является решением уравнения $T_N x_N = f_N$, то есть

$$A_N x_N^{(0)}(0) \equiv T_N x_N^{(0)} = f_N. \quad (28)$$

Тогда, если начальное приближение $x_N^{(0)}$ удовлетворяет условию

$$\|Q_N x_N^{(0)}\| \leq \zeta_N^{(0)} \quad (29)$$

и μ таково, что

$$|\mu| b v < 1, \quad |\mu| v_n < 1, \quad h_N^{(0)} = b_N^2 |\mu| \zeta_N^{(0)} (K + |\mu| M) < 2, \quad (30)$$

$$r_N = b_N |\mu| \zeta_N^{(0)} H_N \leq R, \quad b_N = \frac{b}{1 - b(\tau'_N + \chi'_N + |\mu|(\theta'_N + \lambda'_N + v))},$$

где H_N определено в (22), то уравнение (9) имеет в шаре $S(x_N^{(0)}, r_N) \subset X$ решение $x^* \equiv x^*(\mu)$, к которому сходится проекционно-итерационный процесс (19) с оценкой погрешности

$$\|x_n^{(k_n)} - x^*\| \leq b_N |\mu| \zeta_N^{(0)} V_n \left(\frac{h_N^{(0)}}{2} \right)^{2^{Z_n} - 1},$$

где V_n дается формулой (23), $Z_n = \sum_{m=N}^n (k_m - 1)$.

Доказательство. Утверждение настоящей теоремы получается непосредственно из теоремы 3, если положить в ней $D(x) = [T'(x)]^{-1}$, $D_n(x_n) = [T'_n(x_n)]^{-1}$. Доказательство проводится также, как и доказательство теоремы 2. Отметим лишь, что выполнимость условий близости (25),

(26), очевидно, влечет выполнимость условий (15), (16) соответственно, при этом $\alpha_n = \tau_n + |\mu|\theta_n$, $\alpha'_n = \tau'_n + |\mu|\theta'_n$, $\beta_n = \chi_n + |\mu|\lambda_n$, $\beta'_n = \chi'_n + |\mu|\lambda'_n$ ($n \geq N$).

Теорема доказана.

Выводы. Сформулированные и доказанные в данной работе теоремы могут быть применены к исследованию конкретных функциональных уравнений, в частности, нелинейных интегральных и дифференциальных уравнений, а также операторных уравнений, возникающих в теории возмущений и задачах приближенного построения конформных отображений.

Бібліографічні посилання

1. **Канторович Л.В.** Функциональный анализ и прикладная математика / Л. В. Канторович // УМН. – 1948. – Т. 3, № 6. – С. 89-85.
2. **Канторович Л.В.** Функциональный анализ / Л. В. Канторович, Г. П. Акилов – СПб., 2004. – 816 с.
3. **Мысовских И.П.** О сходимости метода Л.В. Канторовича для решения нелинейных уравнений и его применениях / И. П. Мысовских // Вестник ЛГУ. – 1953. – № 11. – С. 25-48.
4. **Altman M.** Connections between the method of steepest descent and Newton's method / M. Altman // Bull. Acad. Polon. Sci., Cl. III. – 1957. – V. 5, № III. – P. 1031-1036.
5. **Groze S.** A class of iterative methods in Frechet spaces / S. Groze – Prepr.: Res. Sem. – Babes-Bolyai Univ. Fac. Math. and Phys. – 1987. – № 5. – p. 28-35.
6. **Petryshyn W.V.** Projection methods in nonlinear numerical functional analysis / W. V. Petryshyn // J. Math. and Mech. – 1967. – V. 17, № 4. – p. 353-372.
7. **Балашова С.Д.** Проекционно-итерационные методы решения уравнений в нормированных пространствах: дис. на соиск. ученой степ. канд. физ.-мат. наук / С. Д. Балашова – Д., 1974. – 135 с.
8. **Тавадзе Л.Л.** К вопросу о сходимости метода Ньютона при условиях типа Коши / Л. Л. Тавадзе // Вопросы прикладной математики и математического моделирования. – Д., 1997. – С. 152-155.
9. **Тавадзе Л.Л.** Проекционно-итерационные методы решения краевых задач для уравнений эллиптического типа: дис. на соиск. ученой степ. канд. физ.-мат. наук / Л. Л. Тавадзе – Д., 1995. – 152 с.
10. **Балашова С.Д.** Об одном подходе к решению эллиптических нелинейных задач / С. Д. Балашова, Л. Л. Тавадзе // Математические модели и вычислительные методы в прикладных задачах. – Д., 1996. – С. 24-29.

Надійшла до редколегії 1.05.11.

В.А. Громов, А.Н. Шульга

Днепропетровский национальный университет имени. О. Гончара

ПРИМЕНЕНИЕ МЕТОДА «МУРАВЬИНЫХ КОЛОНИЙ» ДЛЯ ПРОГНОЗИРОВАНИЯ ХАОТИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ

Розглянуто новий метод прогнозування часових рядів, що базується на методі «мурашиних колоній». Для цього сформовано мультиграф пошуку характерних послідовностей у часовому ряді та розроблено алгоритм руху «мурах». Представлені результати роботи алгоритму та його модифікації на часовому ряді системи Лоренца.

Рассмотрен новый метод прогнозирования временных рядов, который основывается на методе «муравьиных колоний». Для этого сформирован мультиграф поиска характерных последовательностей во временном ряде и разработан алгоритм движения «муравьев». Представлены результаты работы алгоритма и его модификации на временном ряде системы Лоренца.

New prediction algorithm based on the Ant colony optimization routines is considered in the paper. The multigraph of typical consequences for a given time series is constructed in order to concentrate information. The algorithm to establish “ant” movement is proposed. The basical algorithm and modified one are applied to Lorenz system time series.

Ключевые слова: прогнозирование, временной ряд, мультиграф, «муравей».

Введение. Широкое применение методов прогнозирования временных рядов в различных отраслях науки и техники обусловлено необходимостью предсказания поведения систем различной природы – биологической, экономической, технической. Большинство существующих методов прогнозирования базируются на предположении о стационарности временного ряда либо о весьма простой структуре тренда исследуемого временного ряда. Здесь применяются такие методы как скользящее среднее, АРМА и АРИМА модели [1]. Однако для анализа временных рядов, порождённых существенно нелинейными системами, эти методы зачастую

показывают неудовлетворительные результаты. Ярким примером временных рядов, для которых невозможен прогноз с помощью классических методов математической статистики, являются хаотические временные ряды.

Для прогнозирования временных рядов данного класса обычно используются различные методы прогнозирования, основанные на нейросетевой аппроксимации наиболее характерных тенденций прогнозируемого временного ряда. Так, например, в [9] прогноз осуществлялся на основе перцептрона, в [3] – на основе радиально-базисных сетей, в [11] – на основе адаптивных нейронных сетей с $\Sigma\Pi$ нейронами, в [5] – на основе полиномиальных нейронных сетей. Все эти методы показали, что могут применяться (и применяются) при прогнозировании хаотических рядов, однако для их успешного применения необходимы априорные представления о структуре сети [1].

Другим классом методов, применяемых для прогноза временных рядов, являются нейронечёткие методы. Они основываются на построении набора локальных линейных уравнений для определения глобальной функции аппроксимации, основанной на нечётких правилах. При применении нейронечётких методов преследуются две цели. Во-первых, построение быстро работающего нечёткого метода с простой структурой и улучшенной робастностью, во-вторых – более простая интерпретация полученных результатов [8]. В [4;8] рассмотрены нейронечёткие методы для прогнозирования хаотических временных рядов.

Несмотря на большое количество методов для прогнозирования хаотических рядов, прогнозирование реального хаотического ряда остаётся сложной задачей. В данной работе предлагается новый метод прогнозирования временного ряда с помощью выделения характерных последовательностей из временного ряда. Для этой цели использовался метод «муравьиных колоний» (метод относится к классу методов распределённого искусственного интеллекта). Описание метода можно найти в [6].

Следует отметить, что метод «муравьиных колоний» ранее применялся в различных методах прогнозирования для оптимизации структуры нейронных сетей и статистических моделей [5;10].

Постановка задачи. Рассматривается задача прогнозирования одномерного временного ряда с хаотической динамикой, а именно необходимо по предыдущим значениям временного ряда: $x_1, x_2, \dots, x_{t-1}$ спрогнозировать

значения временного ряда в последующие моменты времени $y[t+1], y[t+2], \dots, y[t+k]$.

Будем предполагать, что прогнозируемый временной ряд отражает поведение некоторой сложной динамической системы, переходные процессы в которой завершились: движение в системе происходит в окрестности некоторого странного аттрактора. Тем самым, можно утверждать, что траектория будет проходить через одни и те же области аттрактора значительное число раз за время наблюдения системы. При этом в каждой области аттрактора можно выделить характерные последовательности значений временного ряда.

Предполагая, что мы находимся в условиях действия теоремы Такенса [2], можно заключить, что в исследуемом временном ряде будут встречаться подобные участки, выделение и анализ которых позволяет осуществлять прогноз временного ряда. Таким образом, задача прогнозирования значения временного ряда в каждый момент времени представима в виде выбора последовательности, первая часть которой описывает предшествующие наблюдения временного ряда, а вторая часть выступает в качестве прогноза. Для выделения характерных для временного ряда последовательностей точек предлагается использовать метод «муравьиных колоний».

Описание алгоритма. Предложенный алгоритм прогнозирования состоит из двух частей: анализа временного ряда с целью выделения характерных последовательностей точек и собственно прогнозирования динамики временного ряда на основе выделенных последовательностей.

Для выделения последовательностей точек во временном ряде используется метод «муравьиных колоний»; для его реализации необходимо представить пространство поиска в виде графа. Собственно говоря, сам метод представляет собой специальный метод поиска в пространстве состояний некоторой системы и заключается в формировании колонии искусственных «муравьёв», которые движутся в пространстве состояний задачи. После окончания движения «муравей» увеличивает вес рёбер, принадлежащих пройденному маршруту (откладывает «феромон» – в терминах метода «муравьиных колоний»). Вероятность выбора «муравьём» определенного маршрута пропорциональна количеству «феромона» на рёбрах этого маршрута.

Пространство поиска. Предположим, что исследуемый ряд нормирован от 0 до 1, разобьем промежуток $[0, 1]$ на подпромежутки возможных значений временного ряда $[x_{i-1}, x_i] = [L_i, R_i]$. Также для нахождения характерных последовательностей точек введем величину D – максимальное допустимое расстояние между соседними элементами последовательности (под расстоянием понимается разница между порядковыми номерами наблюдений, соответствующих элементов данной последовательности). В методе предполагается отыскание корреляций лишь между достаточно близко расположенными наблюдениями.

В качестве пространства поиска было предложено выбрать мультиграф, число вершин которого M , причем любые две вершины мультиграфа связаны с помощью D рёбер. Такой мультиграф кодирует все возможные последовательности точек временного ряда с заданными значениями D и M , причём последовательность не привязана к порядковым номерам своих элементов во временном ряде, но лишь к расстоянию между ними.

Далее отождествляется понятие движения «муравья» по мультиграфу поиска и временному ряду: переход «муравья» с j -й вершины мультиграфа состояний в j_0 по k -му ребру тождественен переходу между наблюдением $y[j]$ и $y[j+k]$ во временном ряде, значения которых принадлежат i -му и j -му промежуткам возможных значений соответственно.

Анализ временного ряда. Предположим, что «муравьи» двигаются по временному ряду, в результате чего выделяют некоторые характерные последовательности переходов, откладывая «феромон» на соответствующие рёбра мультиграфа поиска. Тогда, через определенный промежуток времени количество «феромона» на каждой последовательности рёбер будет пропорционально частоте появления этой последовательности во временном ряде.

Для описания и сравнения таких последовательностей вводится понятие цепочки переходов. Цепочка переходов – это последовательность переходов «муравья» по мультиграфу. В дальнейшем, будем обозначать цепочку переходов как $L(a_1(s_1), a_2(s_2), \dots, a_{n-1}(s_{n-1}), a_n)$, где a_i – номер вершины i -го звена цепочки (т. е. номер промежутка возможных значений, отвечающему i -му звену), s_i – номер ребра i -го перехода в цепочке (т. е. расстоянием между i -м и $i+1$ -м элементами цепочки равно s_i).

Длиной цепочки будем называть количество переходов ней.

Будем считать, что цепочка отвечает части временного ряда, начиная от наблюдения $y[n]$, если:

$$\left| y \left[n + \sum_{j=1}^l \delta_j \right] - \frac{d_{l+1}}{N} \right| < \alpha, \quad \forall l = 0..k,$$

где α – параметр алгоритма (в расчётах использовался $\alpha = \frac{1}{2}$);

k – длина цепочки.

Понятие цепочки переходов обобщает частоту появления той или иной характерной последовательности точек во временном ряде: соответственно, чем большее число раз последовательность данного вида появлялась в исследуемом временном ряде, тем больше «феромона» будет отложено на соответствующей цепочке.

Для построения мультиграфа поиска был использован вспомогательный список цепочек. Если фрагмент временного ряда, по которому сейчас движется «муравей», соответствует цепочке уже существующей в списке цепочек, то на соответствующие этой цепочке рёбра мультиграфа добавляется «феромон», если нет – она добавляется в список. Причем, вероятность выбора одной из возможных цепочек из фрагмента временного ряда прямо пропорциональна количеству «феромона» на соответствующих рёбрах мультиграфа. Если цепочка, присутствующая в списке цепочек, не была найдена ни разу за определенное количество шагов движения «муравьёв», цепочка удаляется из списка.

В качестве параметров алгоритма выступают: количество промежутков возможных значений наблюдений временного ряда – M ; минимальное и максимальное расстояние между двумя соседними точками цепочки – соответственно d_{min}, d_{max} (d_i для каждого звена каждой цепочки принадлежит промежутку $[d_{min}, d_{max}]$); минимальная и максимальная длина цепочки – соответственно k_{min}, k_{max} ; «феромонная» характеристика цепочки – p (т. е. максимально допустимое число шагов движения «муравьёв» по временному ряду для поиска цепочки).

Каждая цепочка характеризуется текущей «феромонной» характеристикой – F_i , т. е. числом шагов движения «муравья» по временному ряду, произошедших с того момента, как данная цепочка была зафиксирована во временном ряде последний раз.

Таким образом, алгоритм поиска характерных последовательностей во временном ряде можно представить следующим образом:

1. Строится мультиграф поиска для заданных значений параметров алгоритма. Все рёбра мультиграфа инициализируются начальным количеством «феромона». Создается пустой список цепочек.
2. «Муравьи» случайным образом размещаются в точки временного ряда. Каждый «муравей» пытается найти цепочку в списке, которая отвечает его положению во временном ряде. Если такая цепочка найдена, то переходим к п. 5.
3. Из фрагмента временного ряда, где находится «муравей», генерируется новая цепочка. Вероятность попадания (i, j, k) -го ребра в цепочку пропорциональна количеству «феромона» на соответствующем ребре мультиграфа.
4. Сгенерированная цепочка добавляется в список цепочек.
5. На рёбра мультиграфа, соответствующие цепочке, добавляется «феромон».
6. Текущая «феромонная» характеристика P_i для данной цепочки увеличивается
7. Текущие «феромонные» характеристики для всех цепочек из списка найденных цепочек уменьшаются. Если для той или иной цепочки текущая «феромонная» характеристика равна нулю, то цепочка удаляется из списка.
8. «Испаряется» «феромон» на всех рёбрах мультиграфа. Переходим к п.2.

В качестве критериев завершения процесса могут быть выбраны следующие:

1. Количество «феромона» на рёбрах мультиграфа изменяется не больше чем на величину P_i за последние k итераций.
2. За последние k итераций не было сгенерировано ни одной новой цепочки.

3. Количество итераций N превышает некоторый порог N_{max} .

Алгоритм прогнозирования значений временного ряда. На основе проведённого анализа строится прогноз значений временного ряда:

1. Рассматриваются последние $\delta_{max} \cdot k_{max}$ наблюдений временного ряда.
2. Для каждого из этих наблюдений находятся все цепочки, для которых d_1 (промежуток возможных значений для первой точки цепочки) соответствует промежутку возможных значений для текущего наблюдения t , и первые n звеньев цепочки соответствуют фрагменту временного ряда, который начинается с текущего наблюдения.
3. Рассматриваются следующие k -н переходов (k – длина выбранной цепочки). Если соответствующие точки находятся на уже известной части ряда, увеличиваем n и переходим к п.2.
4. Из цепочек, отобранных в п.3 выбирают ту, для которой значение величины r минимально:

$$r = \frac{\sum_{i=1}^n \left(\frac{d_{i+1}}{M} - y[t + \sum_{j=1}^i \delta_j] \right)^2}{n}.$$

Отношение r характеризует меру уклонения реально наблюдаемых величин для этой части ряда от цепочки.

5. По оставшимся точкам выбранной цепочки строится прогноз:

$$y[t + \sum_{i=1}^j \delta_i] = \frac{d_{j+1}}{M}, j = \overline{n, k}$$

j – номер перехода, который привел в эту точку;

k – длина цепочки.

Для работы с большими флуктуациями в наблюдениях метод может быть модифицирован следующим образом:

- На рёбра цепочек, с большой амплитудой откладывается больше «феромона». Цепочка считается цепочкой с большой амплитудой, если выполняется следующее соотношение:

$$\exists i: |d_i - d_{i+1}| > \frac{2M}{3},$$

где M — количество промежутков возможных значений наблюдений временного ряда;

d_i — номер промежутка возможных значений для i -й точки цепочки.

- При нахождении цепочки на очередном фрагменте временного ряда, для всех наблюдений временного ряда внутри каждого звена цепочки рассчитывались и запоминались максимальный и минимальный номера промежутков возможных значений этих наблюдений — $N_i^{\max}, N_i^{\min}, i = \overline{1, K}$, где K — длина цепочки. При прогнозе, для наблюдений внутри каждого звена найденной части цепочки таким же образом определялись величины — $N_i^{\max}, N_i^{\min}, i = \overline{1, K}$, где K — количество звеньев в найденной части цепочки. По цепочке строился прогноз только тогда, когда:

$$[N_i^{\max}, N_i^{\min}] \in [N_i^{\max}, N_i^{\min}], i = \overline{1, K}.$$

Анализ полученных результатов. Работа предложенного алгоритма протестирована на классическом временном ряде, полученном из системы Лоренца. Анализ осуществлялся на первых 5000 наблюдений, проверка качества прогнозирования — на последующих 2000. Максимальная длина прогнозируемого промежутка составляла 40.

Оптимальными (для ряда Лоренца) можно считать следующие параметры: $M = 20$, $d_{\max} = 2$, $d_{\min} = 10$, $k_{\max} = 5$, $k_{\min} = 8$.

На рис. 1 представлены результаты применения алгоритма при 1000 шагов движения «муравьёв» и 1000 «муравьях», на рис. 2 — при 5000 шагов движения «муравьёв» и 1000 «муравьях». Здесь и на последующих рисунках сплошной линией отмечены реальные значения наблюдений, а линией с кружочками — прогнозные значения. Наблюдается существенное улучшение качества прогноза при увеличении числа шагов движения «муравьёв» на первом этапе алгоритма.



Рис. 1. Результаты работы алгоритма при 1000 шагов движения «муравьёв» и 1000 «муравьях»



Рис. 2. Результаты работы алгоритма при 5000 шагов движения «муравьёв» и 1000 «муравьях»

На рис.3 представлены результаты применения модификации алгоритма при 1000 шагов движения «муравьёв» и 1000 «муравьях», на рис.4 – при 5000 шагов движения «муравьёв» и 1000 «муравьях». Наблюдается существенное улучшение качества прогноза по сравнению с обычной версией алгоритма.

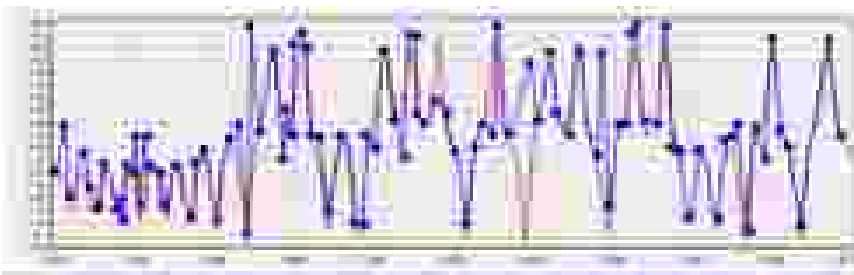


Рис. 3. Результаты работы модифицированной версии алгоритма при 1000 шагов движения «муравьёв» и 1000 «муравьях»



Рис. 4. Результаты работы модифицированной версии алгоритма при 1000 шагов движения «муравьёв» и 5000 «муравьях»

На рис. 5 и 6 представлены графики зависимости процента числа точек для которых возможно построить прогноз и среднеквадратической ошибки в зависимости от количества шагов движения «муравьёв». С ростом числа итераций количество прогнозов растёт, среднеквадратическая ошибка уменьшается.

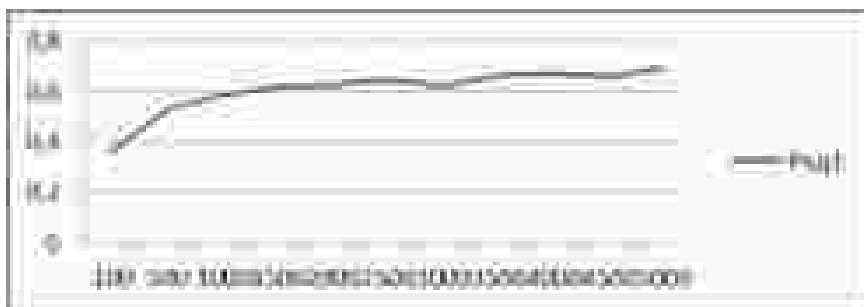


Рис. 5. Зависимость процента числа точек для которых возможно построить прогноз от числа шагов движения «муравьёв»

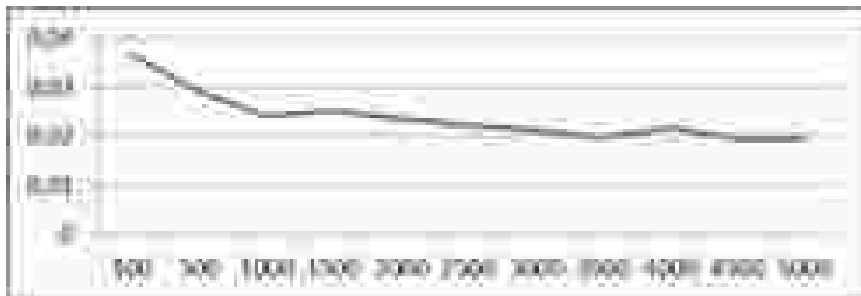


Рис. 6. Зависимость среднеквадратической ошибки от числа шагов движения «муравьёв»

Выводы. В работе предлагается новый метод прогнозирования хаотических временных рядов, основанный на использовании алгоритма «муравьиных колоний».

Анализируется качество работы метода, определяется набор оптимальных параметров.

Эффективность работы метода проверяется на классическом хаотическом временном ряде системы Лоренца. Алгоритм продемонстрировал удовлетворительные результаты прогноза.

Библиографические ссылки

1. **Дуброва Т.А.** Статистические методы прогнозирования /Т.А. Дуброва – М., 2003.
2. **Малинецкий Г.Г.** Современные проблемы нелинейной динамики / Г.Г. Малинецкий, А.Б. Потапов – М., 2000.
3. **Abbas Erfanian Omidvar.** Configuring radial basis function network using fractal scaling process with application to chaotic time series prediction / Abbas Erfanian Omidvar // Chaos Solitons and Fractals. – 2004. – № 22. – P. 757–766.
4. **Cheng-Jian Lin.** A self-adaptive neural fuzzy network with group-based symbiotic evolution and its prediction applications / Cheng-Jian Lin, Yong-Ji Xu // Fuzzy Sets and Systems. – 2006. – №157. – P. 1036–1056.
5. **Dongxiao Niu.** Power load forecasting using support vector machine and ant colony optimization / Dongxiao Niu, Yongli Wang, Desheng Dash Wu // Expert Systems with Applications. – 2010. – №37, P. 2531–2539.

6. **Dorigo M.** Ant System: Optimization by a Colony of Cooperating Agents / Dorigo M, V. Maniezzo & A. Colomi // IEEE Transactions on Systems, Man, and Cybernetics-Part B, - 1996. - № 26(1). - P. 29-41.
7. **Gomez-Ramirez E.** Forecasting time series with a new architecture for polynomial artificial neural network / E. Gomez- Ramirez, K. Najim, E. Ikonen // Applied Soft Computing. - 2007. - №7. - P. 1209-1216.
8. **Hong Gu.** Fuzzy prediction of chaotic time series based on singular value decomposition / Hong Gu, Hongwei Wang // Applied Mathematics and Computation. - 2007. - №185. - P. 1171-1185.
9. **Sanjay Vasant Dudul.** Prediction of a Lorenz chaotic attractor using two-layer perceptron neural network / Sanjay Vasant Dudul // Applied Soft Computing. - 2005. - № 5. - с. 333-355.
10. **Wei-Chiang Hong.** Forecasting urban traffic flow by SVR with continuous ACO / Wei-Chiang Hong, Yucheng Dong, Feifeng Zheng, Chien-Yuan Lai // Applied Mathematical Modelling. - 2011. - №35. - P. 1282-1291.
11. **Wong W.K.** Adaptive neural network model for time-series forecasting / W.K. Wong, Min Xia, W.C. Chu // European Journal of Operational Research. - 2010. - № 207. - P. 807-816.

Надійшла до редколегії 24.06.11

УДК 519.246.8

⁷Т.В. Гурьянова

Донецкий национальный технический университет

ПРИМЕНЕНИЕ КРИТЕРИЯ КЕЛЛИ ДЛЯ УПРАВЛЕНИЯ КАПИТАЛОМ В СЛУЧАЕ УСЕЧЕННОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДОХОДНОСТИ ТОРГОВОЙ СИСТЕМЫ

Проведено аналіз алгоритмів динамічного управління капіталом, що ґрунтуються на критерії Келлі. Установлені загальні закономірності залежності величини φ^* Р.Вінса від параметрів торговельної системи, що оптимізується, для випадку обмеженого нормального закону розподілу прибутковості торговельної системи.

Проведен анализ алгоритмов динамического управления капиталом, основанных на использовании критерия Келли. Установлены общие закономерности поведения величины «оптимального φ » Р.Винса при изменении параметров оптимизируемой торговой системы для случая усеченного нормального распределения доходности торговой системы.

The paper is devoted the analysis of money management's dynamics algorithms based on Kelly's criterion. A general dependency «optimal φ » R.Vince on trading system parameters case Gaussian distribution of profitableness was calculated. The results can be used in investigation of real trading system.

Ключевые слова: реинвестирование, критерий Келли, управление капиталом.

Введение. Начало современной дискуссии об оптимальном управлении капиталом было положено Дж. Л. Келли. В [1] он показал, что для достижения максимального роста своего благосостояния при реинвестировании, игрок должен максимизировать ожидаемую величину логарифма темпа роста своего капитала. В настоящее время, известно большое количество алгоритмов динамического управления капиталом (ДУК), являющихся развитием критерия Келли.

В [2] Р. Винс предложил оригинальный алгоритм ДУК – теорию «оптимального f ». На недостатки применения этого эмпирического метода, связанного с высокой чувствительностью к величине наибольшего проигрыша, было указано в [2], и для устранения этого недостатка Р. Винсом был предложен параметрический метод нахождения «оптимального f ». Однако подробного анализа зависимости величины «оптимального f » от параметров оптимизируемой торговой системы не приводится.

Данная работа посвящена изучению связи величины f^* – части капитала, которую нужно реинвестировать в каждую сделку для максимизации капитала игрока, и средней геометрической доходности G с параметрами торговой системы для случая нормального закона распределения доходности системы.

Модель функционирования торговой системы. Любая торговая система может быть описана одномерным законом распределения вероятностей ее выходных случайных величин $\{x_i\}$ на интервале стационарности их реализации τ_{cm} . В большинстве практических случаев это нормальный закон, который вытекает из центральной предельной теоремы теории вероятностей (когда на систему действует одновременно большое количество разнонаправленных случайных факторов). В этом случае на интервале стационарности τ_{cm} система характеризуется двумя числовыми характеристиками: математическим ожиданием M и дисперсией σ^2 , которые являются постоянными и не зависят от времени.

В то же время, для реальной торговой системы на значения $\{x_i\}$ накладываются некоторые ограничения, которые в нашем исследовании сформулированы следующим образом: максимальный проигрыш системы не может быть больше 100 %, максимальный выигрыш не может превышать значения 100 %. Таким образом, $\{x_i\} \in [-1; +1]$. Например, если система при j -й сделке выиграла 10 % от вложенного капитала, то $x_j = 0,1$, если проиграла на k -й сделке 90 % от вложенного капитала, то $x_k = -0,9$. Следовательно, после n сделок капитал игрока можно найти по формуле

$$K_n = K_0 \cdot \prod_{i=1}^n (1 + x_i \cdot f_i), \quad (1)$$

где K_n – капитал игрока после n сделок, K_0 – первоначальный капитал

игрока, x_i – выигрыш (проигрыш) системы в i -й сделке, f_i – часть капитала, вложенная в i -ю сделку. Тогда коэффициент приумножения капитала за n сделок равен

$$\frac{K_n}{K_0} = \exp\left(n \ln\left(\frac{K_n}{K_0}\right)^{\frac{1}{n}}\right) = e^{n \ln G}, \quad (2)$$

где G – среднее геометрическое доходности или темп роста капитала.

Таким образом, по критерию Келли для достижения максимального роста своего благосостояния игроку необходимо максимизировать $\ln G$. Для этого ему нужно реинвестировать в каждую сделку такую часть своего капитала f^* , которая соответствует максимальному значению $\ln G$ для данной торговой системы.

Алгоритм динамического управления капиталом. В [2] Р. Винсом предложен параметрический метод определения оптимальной части капитала f при реинвестировании. В этом случае для нахождения f – оптимальной доли капитала для реинвестирования, необходимо максимизировать логарифм темпа роста системы при реинвестировании, что может быть выражено уравнением

$$\ln G|_{\max} = \int_{-1}^1 \ln(1 + f \cdot x) \cdot w(x) \cdot dx, \quad (3)$$

где $w(x)$ – заданная функция плотности вероятности закона распределения доходности системы.

Анализ результатов. Уравнение (2) решалось с использованием численных методов – метода Ньютона [3]. Для заданного закона распределения $w(x)$ методом Симпсона [3] рассчитывалось значение интеграла (3), затем выбиралось начальное приближение значения f (как правило $f = 0,5$), далее проводилось уточнение решения. Расчеты значений f проводилось с точностью до 10^{-6} , поиск решения проводился на отрезке $f \in [0; +1]$. Численное моделирование проводилось в среде табличного процессора MS Excel.

Расчеты значений «оптимального f » (f^*) проводилось для $M \in [0,02; 1]$ с шагом 0,02, $\sigma \in [0,02; 1]$ с шагом 0,02.

Из анализа проведенных расчетов следует, что в этом случае существует достаточно широкая область значений параметров M и σ , для которых значение f почти не отличается от 1 (так, при любых значениях M и $\sigma \leq 0,14$ значение $f \geq 0,999$). В таких торговых системах применение критерия Келли почти, наверное, приведет к разорению. Таким образом, торговые системы, где $\sigma \leq 0,14$, не представляют интереса для нашего исследования.

В таблице 1 приведены значения зависимости $f^*(M, \sigma)$ при $M \in [0,02;1]$ с шагом 0,04, $\sigma \in [0,02;1]$ с шагом 0,1. В таблице 2 приведены значения зависимости $G(M, \sigma)$ при $M \in [0,02;1]$ с шагом 0,04, $\sigma \in [0,02;1]$ с шагом 0,1.

На рис. 1 приведена поверхность, построенная по результатам расчета зависимости $f^*(M, \sigma)$. На рис. 2 приведен график зависимости $G(M, \sigma)$.

Из анализа проведенных расчетов следует, что для случая ограниченного нормального закона распределения зависимость $G(M)$ является практически линейной в широком диапазоне значений σ . Причем, при $M \leq 0,5$ и $\sigma \leq 0,2$ значение $G \approx M + 1$ (с точностью до 1 %). При дальнейшем увеличении вариабельности величины выигрыша торговой системы зависимость $G(M)$ становится существенно нелинейной.

Выводы. Проведенный анализ позволяет сделать вывод об общих закономерностях зависимостей $f^*(M, \sigma)$ и $G(M, \sigma)$ от параметров торговой системы и выбрать оптимальные параметры моделей торговых систем для анализа эффективности алгоритмов ДУК.

Очевидно, что для реальной задачи ДУК торговой системы (с ее ограниченным набором реализаций $\{x_i\}$, на котором проводится оптимизация доли реинвестирования капитала) результаты будут отличаться от значений, полученных на неограниченном числе реализаций, однако общие закономерности сохранятся. При этом, с увеличением периода анализа интервала стационарности для оценки последующего поведения ряда они будут выполняться с увеличивающейся степенью точности.

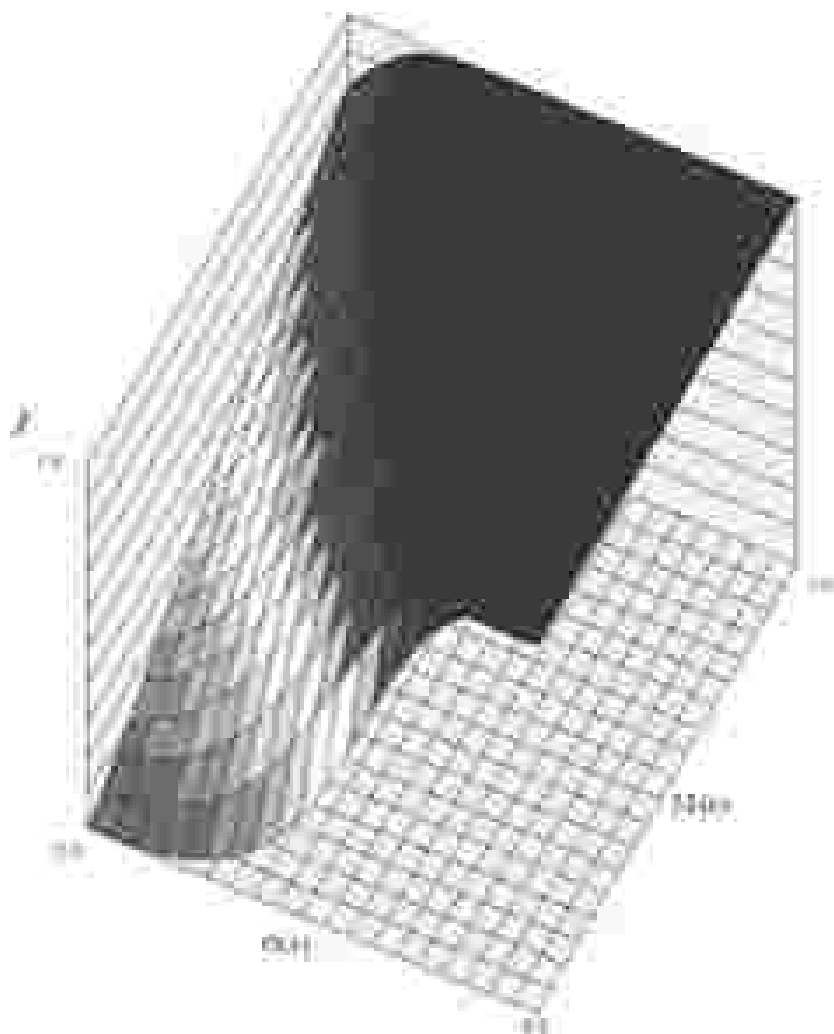


Рис. 1. Зависимость $f^*(M, \sigma)$ для случая усеченного нормального закона распределения

Таблиця 1

Результаты расчетов f^* для торговых систем при $M \in [0,02;1]$ с шагом 0,04, $\sigma \in [0,02;1]$ с шагом 0,1

$\sigma \backslash M$	0.02	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.02	1.00	1.00	0.49	0.22	0.12	0.08	0.06	0.04	0.03	0.02	0.02
0.06	1.00	1.00	1.00	0.64	0.37	0.24	0.17	0.12	0.09	0.07	0.06
0.10	1.00	1.00	1.00	0.97	0.59	0.39	0.27	0.20	0.16	0.12	0.10
0.14	1.00	1.00	1.00	1.00	0.79	0.53	0.38	0.28	0.22	0.17	0.14
0.18	1.00	1.00	1.00	1.00	0.93	0.66	0.48	0.36	0.28	0.22	0.18
0.22	1.00	1.00	1.00	1.00	1.00	0.78	0.57	0.43	0.34	0.27	0.22
0.26	1.00	1.00	1.00	1.00	1.00	0.87	0.66	0.50	0.39	0.31	0.26
0.30	1.00	1.00	1.00	1.00	1.00	0.94	0.74	0.57	0.45	0.36	0.29
0.34	1.00	1.00	1.00	1.00	1.00	0.98	0.81	0.64	0.50	0.41	0.33
0.38	1.00	1.00	1.00	1.00	1.00	1.00	0.87	0.70	0.56	0.45	0.37
0.42	1.00	1.00	1.00	1.00	1.00	1.00	0.92	0.75	0.61	0.49	0.41
0.46	1.00	1.00	1.00	1.00	1.00	1.00	0.95	0.80	0.65	0.53	0.44
0.50	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.84	0.70	0.57	0.48
0.54	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.88	0.74	0.61	0.51
0.58	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.92	0.78	0.65	0.54
0.62	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.94	0.81	0.69	0.58
0.66	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.85	0.72	0.61
0.70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.88	0.75	0.64
0.74	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.90	0.78	0.67
0.78	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.92	0.81	0.70
0.82	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.94	0.84	0.72
0.86	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.86	0.75
0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	0.88	0.77
0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.90	0.80
0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.92	0.82
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.93	0.83

Таблиця 2

Результаты расчетов G для торговых систем при $M \in [0,02;1]$ с

шагом 0,04, $\sigma \in [0,02;1]$ с шагом 0,1

$\sigma \backslash M$	0.02	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.02	1.02	1.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.06	1.06	1.06	1.04	1.02	1.01	1.01	1.00	1.00	1.00	1.00	1.00
0.10	1.10	1.10	1.08	1.05	1.03	1.02	1.01	1.01	1.00	1.00	1.00
0.14	1.14	1.14	1.12	1.09	1.05	1.03	1.02	1.01	1.01	1.00	1.00
0.18	1.18	1.18	1.16	1.14	1.09	1.05	1.03	1.02	1.01	1.01	1.00
0.22	1.22	1.22	1.20	1.18	1.12	1.07	1.04	1.03	1.02	1.01	1.01
0.26	1.26	1.26	1.24	1.22	1.16	1.10	1.06	1.03	1.02	1.01	1.01
0.30	1.30	1.30	1.28	1.26	1.20	1.13	1.08	1.05	1.03	1.02	1.01
0.34	1.34	1.34	1.32	1.30	1.24	1.16	1.10	1.06	1.04	1.02	1.02
0.38	1.38	1.38	1.36	1.33	1.27	1.19	1.12	1.07	1.05	1.03	1.02
0.42	1.42	1.42	1.40	1.37	1.31	1.22	1.14	1.09	1.06	1.04	1.03
0.46	1.46	1.46	1.44	1.41	1.34	1.25	1.17	1.10	1.07	1.04	1.03
0.50	1.50	1.50	1.48	1.44	1.37	1.28	1.19	1.12	1.08	1.05	1.04
0.54	1.54	1.54	1.52	1.47	1.40	1.31	1.22	1.14	1.09	1.06	1.04
0.58	1.58	1.58	1.56	1.51	1.43	1.34	1.24	1.16	1.10	1.07	1.05
0.62	1.62	1.62	1.60	1.54	1.46	1.36	1.27	1.18	1.12	1.08	1.05
0.66	1.66	1.66	1.63	1.57	1.49	1.39	1.29	1.20	1.13	1.09	1.06
0.70	1.70	1.70	1.66	1.59	1.51	1.41	1.31	1.22	1.15	1.10	1.07
0.74	1.74	1.74	1.69	1.62	1.53	1.44	1.34	1.24	1.16	1.11	1.08
0.78	1.78	1.77	1.72	1.64	1.56	1.46	1.36	1.26	1.18	1.12	1.08
0.82	1.82	1.81	1.75	1.67	1.58	1.48	1.38	1.28	1.19	1.13	1.09
0.86	1.86	1.84	1.77	1.69	1.60	1.50	1.40	1.30	1.21	1.15	1.10
0.90	1.90	1.87	1.79	1.71	1.62	1.52	1.42	1.32	1.23	1.16	1.11
0.94	1.94	1.89	1.81	1.73	1.64	1.54	1.44	1.34	1.24	1.17	1.12
0.98	1.97	1.91	1.83	1.74	1.65	1.56	1.46	1.35	1.26	1.18	1.13
1.00	1.98	1.92	1.84	1.75	1.66	1.57	1.46	1.36	1.27	1.19	1.13

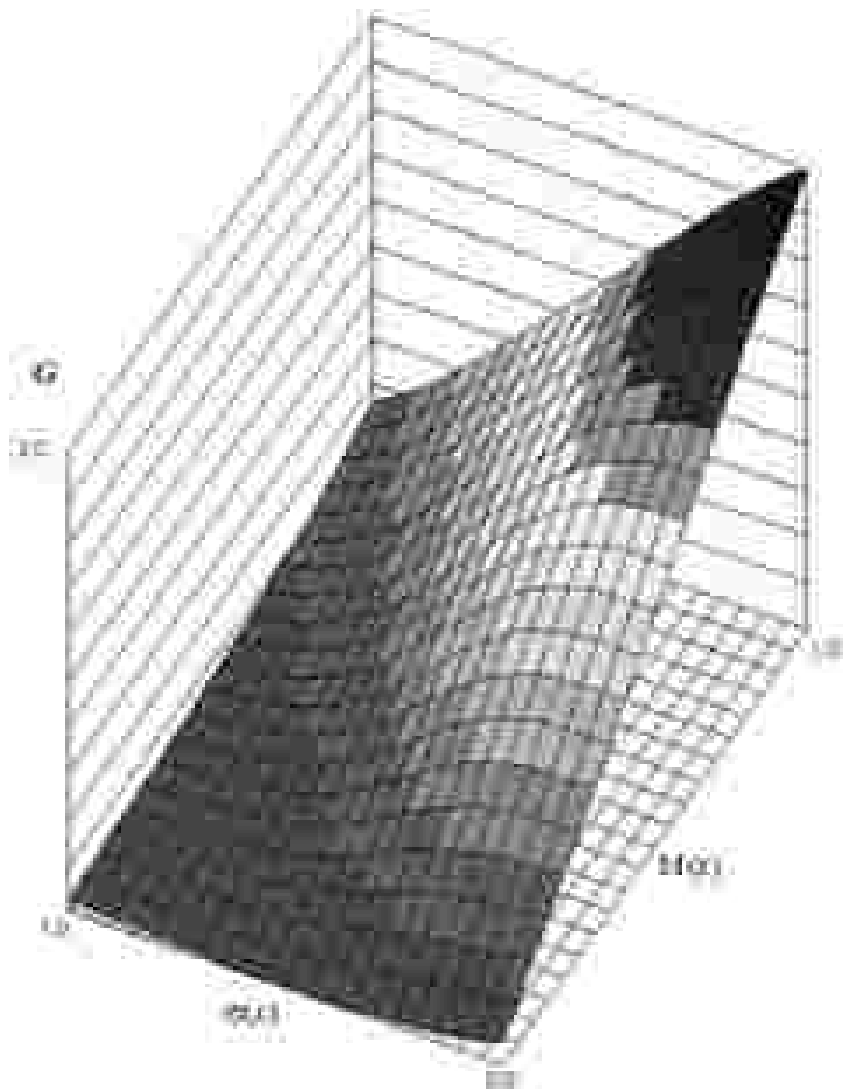


Рис. 2. Зависимость $G(M, \sigma)$ для случая усеченного нормального закона распределения

Библиографические ссылки

1. **J. L. Kelly Jr.** A New Interpretation of Information Rate. The Bell System Technical Journal, 35(4), 1956. – P. 917-926.

2. **Винс Р.** Математика управления капиталом. Методы анализа риска для трейдеров и портфельных менеджеров: Пер. с англ. / Р. Винс – М., 2001. – 400 с.

3. **Бахвалов Н.С.** Численные методы – 6-е изд. / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. – М., 2008. – 636 с.

Надійшла до редколегії: 10.11.10

⁸А.А. Дегтярёв, Т.А. Зайцева

Днепропетровский национальный университет имени Олеся Гончара

**ЛИНГВОСТАТИСТИЧЕСКАЯ МОДЕЛЬ ОЦЕНКИ
РЕПРЕЗЕНТАТИВНОСТИ ТЕКСТОВОГО ФРАГМЕНТА С
ПРИМЕНЕНИЕМ АДАПТАЦИОННЫХ МЕХАНИЗМОВ
ОПРЕДЕЛЕНИЯ ГРАНИЦ ПРЕДЛОЖЕНИЯ**

Пропонується підхід до оцінки лінгвістичної узгодженості текстового фрагмента як синтаксично оформленого речення. Особливо детально проводиться аналіз тієї його частини, що виконує нечітку оцінку семантичної зумовленості речення. В аспекті визначення границь текстового фрагменту, запропонований розв'язок також використовує декілька простих евристичних допоміжних підходів, адаптованих під особливості веб_контенту.

Предлагается подход к оценке лингвистической целостности текстового фрагмента как синтаксически оформленного предложения. Особенно детально проводится анализ выполнения нечеткой оценки семантической обусловленности предложения. В аспекте определения границ фрагмента, предложенное решение использует несколько простых эвристических вспомогательных подходов, адаптированных под особенности веб_контента.

An approach for linguistic integrity estimation of a textual fragment as a sentence is proposed. It is based mainly on statistical model of part-of-speech tags distribution over the sentence. Special detailed analysis of fuzzy estimation of sentence semantics is given. The approach also uses several simple heuristic co-approaches adapted for web-content peculiarities, which deal with sentence boundary determining.

Ключевые слова: нечеткая оценка, веб _контент, компьютерная лингвистика, текст, предложение.

Введение. Многие прикладные задачи из области компьютерной лингвистики зачастую сводятся к оценке входного текстового потока на предмет соответствия некоторым наперед заданным характеристикам. Это может быть, например, определение степени принадлежности полученно-

го на вход контента некоторой тематике. К таким задачам относится углубленный анализ некоторого языкового корпуса на предмет частоты встречаемости словосочетаний заданного шаблона, например, *глагол+предлог+существительное, существительное+прилагательное* и т. п., где на каждый элемент шаблона также могут накладываться некоторые условия. Во всех задачах данного направления присутствует выделение текстовых фрагментов и их последующий комплексный лингвистический анализ, в свою очередь, состоящий из синтаксического, морфологического и семантического. Как правило, анализу подвергаются фрагменты, которые могут быть расценены как предложение, и первичным этапом обработки текстового потока является определение границ предложения [1]. Внутри предложения проводится вторичный анализ, выполняющийся в соответствии со спецификой исходной задачи. Анализ предложения может быть относительно простым, например, включать только поиск некоторых ключевых слов в контексте определения принадлежности ресурса определенной тематике, либо быть более углубленным, и задействовать морфологическую информацию входящих в него слов и/или разбор зависимостей между словами в предложении [2].

Постановка задачи. В контексте определения того, насколько представленный входной поток данных можно считать осмысленным текстом, сформулируем задачу:

1. Построить алгоритм разбиения входного текстового потока на значимый и незначимый контент;
2. Построить алгоритм расчета функции числовой оценки качества фрагмента, отнесенного к значимому контенту, как хорошо оформленного предложения с точки зрения синтаксиса, морфологии и семантики.

В начале, необходимо провести первичное разбиение входного непрерывного потока на фрагменты, которые являются кандидатами в хорошо оформленные предложения. Для этого можно использовать разбиение текста на фрагменты на основании разметки входного документа, характер которой зависит от формата документа. Например, в контексте поисковой системы по Интернет_ресурсам, основными элементами разметки являются теги HTML. В других случаях, разметку определяют другие формат-зависимые служебные элементы документа. Кроме того, в качестве дополнительного критерия разбивки может быть использован регистр первых букв входных слов в сочетании со знаками пунктуации. Таким образом, фрагменты, обрамленные парными тегами HTML или начинающиеся с капитализированного слова и заканчивающиеся, например, точкой, бу-

дуг отнесены к кандидатам в значимые фрагменты, к которым будет применен вторичный анализ. Остальной текст, соответственно, автоматически причисляется к незначимой части. Так как решение первой задачи сводится к выделению формат _зависимых паттернов и представляется относительно тривиальным, рассмотрим более детально вторую задачу, которая заключается в проведении анализа полученных фрагментов _кандидатов на предмет численной оценки их лингвистического качества.

Расчет числовой оценки лингвистической ценности фрагмента.

В основу алгоритма оценки фрагмента легли две составляющие:

1. попарная оценка согласованности словоформ, входящих во фрагмент;
2. оценка равномерности расположения разных частей речи внутри фрагмента.

Первая составляющая имеет в своей основе морфологию языка, и, очевидно, более применима к языкам с высокой флексией, к которым относится, например, русский или украинский. В виду разнообразия грамматических категорий, которыми обладают эти языки, каждое слово имеет сравнительно большое количество ассоциированных словоформ (за счет изменений падежа, рода, числа и т. п.), что, в свою очередь, требует соблюдения определенных правил их сочетания при наличии определенной связи между соответствующими словами в предложении. Если во фрагменте_кандидате для большинства таких пар слов выполняются эти правила согласования, то качество такого фрагмента должно быть оценено как высокое. И наоборот – если во фрагменте много несогласованных пар, то велика вероятность того, что в рассмотрение попал бессвязный шум, и его качество необходимо оценить как низкое. Исходя из описанного принципа, морфологическую составляющую оценки фрагмента можно выразить как

$$\mu = \frac{N'}{N}, \quad (1)$$

где N – общее количество рассматриваемых во фрагменте пар слов, N' – количество морфологически согласованных пар слов.

Очевидно, что организация расчета морфологической составляющей требует реализации соответствующего аппарата, который обеспечивал бы удобное представление и анализ морфологических свойств словоформ. Кроме того, необходим языковой словарь с хорошим наполнением, что обеспечивало бы максимальное покрытие контента и минимизацию неизвестных, а значит и не участвующих в анализе, слов. С этой целью разра-

ботана иерархия классов частей речи на языке C++, и реализован соответствующий вызывающий модуль, который обеспечивает эффективную проверку согласованности словоформ в предложении. В качестве словарной базы для русского языка использовался модифицированный словарь А. Зализняка [3] с применением подходов, описанных в [4]. Следует отметить, что целесообразно ограничиться рассмотрением только тех слов в парах, которые принадлежат к т. н. главным частям речи (существительное, глагол, прилагательное и наречие).

Рассмотренный принцип оценки предложения по количеству согласованных пар слов является информативным, однако не может давать полное представление о предложении как языковой единице. Очевидно, что можно подобрать контент таким образом, чтобы он сплошь состоял из морфологически согласованных пар слов, но в целом представлял бы из себя не предложение, а набор, например, словосочетаний, которые никак не связаны друг с другом. В условиях неопределенности данных в интернет_контенте такая ситуация встречается относительно часто в контексте так называемой SEO _оптимизации интернет _ресурса. В этом случае на ресурс помещаются автоматически формируемые, хорошо согласованные пары слов, быть может с хорошей пунктуационной разметкой по фрагменту в целом, среди которых оптимизируемое для поисковых систем слово встречается чаще, чем другие. Рассмотрим пример:

Дифференциальные модели, дискретная динамическая система, решение уравнений в частных производных.

Этот фрагмент является не связанным набором слов, а показатель лингвистической целостности для него должен быть мал. Однако, определение показателя целостности предложения по количеству согласованных пар слов даст здесь высокий показатель. Для выхода из сложившейся ситуации могут быть в разных комбинациях использованы три подхода:

- 1.1. Анализ синтаксической структуры предложения (определение подчиненности фрагментов, определение членов предложения и установление связей между ними);
- 1.2. Определение конкретных технологий SEO_оптимизации предложения, их распознавание и понижение показателя целостности в зависимости от проведенного анализа;
- 1.3. Статистический анализ расположения и порядка чередования частей речи, входящих в предложение.

Реализация точного синтаксического анализа предложения существенно зависит от особенностей конкретного языка, а именно , от особенностей порядка слов в предложении и правил чередования членов предложе-

ния [4], хороший подход к этому вопросу для английского и корейского языков рассмотрен в [5]. Кроме того, точный синтаксический анализ в условиях интернет-контента может оказаться плохо применимым, так как хорошо оформленные, корректные тексты встречаются реже, чем тексты с нарушенной структурой и написанные без строго соблюдения языковых правил (сообщения на форумах, в блогах, статьи плохой верстки и т. д.). Поэтому оценка лингвистической целостности фрагмента, основанная на точном синтаксическом анализе, будет давать для большей части контента весьма низкие значения и не будет обеспечивать необходимую дифференциацию контента по его качеству.

Отдельного внимания заслуживает подход к вычленению SEO-оптимизированных фрагментов из контента. В силу развития SEO-технологий, особенно актуальной становится разработка и реализация алгоритмов, определяющих, насколько контент является качественным с точки зрения человека и насколько контент удовлетворяет требованиям и специфике конкретной поисковой системы. Что касается именно лингвостатистических аспектов определения наличия SEO-оптимизации на контенте, то стоит выделить следующие универсальные подходы:

1. определение наличия длинных последовательностей повторяющихся слов;
2. вычисление частоты слов на контенте и определение наличия слов, имеющих частоту очень высокую по отношению к размеру контента, при этом лучшая точность такой проверки обеспечивается при выполненной нормализации и синонимизации слов, т. е. при расчете частоты только по базовым синонимам в нормальной форме;
3. определение отношения количества слов, не попавших в предложения, к общему количеству слов на контенте, при этом предложения для этой проверки могут определяться по теговой и пунктуационной разметке с учетом регистра символов и других языковых особенностей обозначения начала и конца предложения.

Построение лингвостатистической модели оценки качества предложения, основанной на морфологической информации входящих в него слов, имеет свои особенности. Морфологическая согласованность словоформ может не давать хорошей оценки, особенно для языков с нестрогим порядком слов в предложении, так как можно подобрать осмысленное правильное предложение, в котором попарно взятые согласованные словоформы не будут давать желаемого значения соотношения (1). Рассмотрим

схему расчета другого показателя, основанного на морфологии слов, который базируется на оценке распределения тегов частей речи по всему предложению. Рассчитаем величину, которую назовем дисперсией части речи в предложении

$$D(P) = \frac{1}{N_p + 1} \sum_{i=0}^{N_p} \left(s_i - \frac{N - N_p}{N_p + 1} \right)^2, \quad (2)$$

где $\boxed{N}$ – общее количество слов в предложении, N_p – количество слов с установленным тегом части речи $\boxed{P}$, $\boxed{S_i}$ – расстояния между двумя последовательными появлениями части речи $\boxed{P}$, измеренные в словах. При этом при подсчете расстояний кроме частей речи учитываются также метки начала и конца оцениваемого фрагмента, а измерение проходит без включения слов, между которыми рассчитываются расстояния. Проводя оценку заведомо корректных предложений и производя расчет по формуле (2) для каждой из присутствующих во фрагменте части речи, можно заключить, что для высокого значения показателя лингвистической целостности фрагмента следует требовать как можно большей суммарной равномерности распределения частей речи по этому фрагменту. Лучшая равномерность обеспечивается при меньшем значении дисперсии, поэтому целесообразно рассчитать величину, которую назовем обратной дисперсией части речи во фрагменте

$$R(P) = 1 - kD(P). \quad (3)$$

При этом $\boxed{k}$ – коэффициент нормализации $D(P)$ к интервалу $(0; 1)$, который легко определяется из формулы (2), раскрытой для худшего случая с точки зрения равномерности распределения части речи (максимального значения дисперсии), когда все N_p слов с тегами части речи $\boxed{P}$ стоят подряд в начале или в конце предложения, по следующей формуле

$$\begin{aligned} k &= \left(\frac{1}{N_p + 1} \left(N_p \left(\frac{N - N_p}{N_p + 1} \right)^2 + \left(\frac{N_p (N - N_p)}{N_p + 1} \right)^2 \right) \right)^{-1} = \\ &= \left(\frac{N_p + 1}{N_p (N - N_p)} \right)^2. \end{aligned} \quad (4)$$

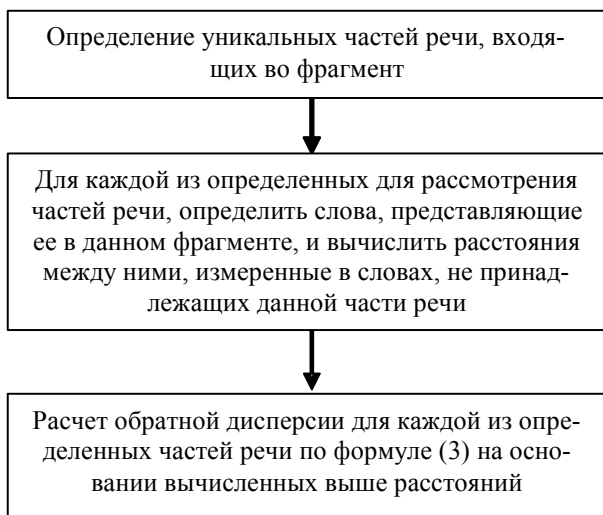
Получив значения обратной дисперсии для каждой из частей речи, для которых $N > N_p \geq 1$, можно рассчитать итоговую оценку распределения частей речи фрагмента по следующей формуле

$$\Psi = \frac{1}{p} \sum_{i=1}^p N_i R(P_i). \quad (5)$$

Здесь p – количество классов частей речи, слова, принадлежащие которым, хотя бы один раз появляются во фрагменте; N_i – количество слов, помеченных тегом i -ой части речи; $R(P_i)$ – обратная дисперсия i -ой части речи, рассчитанная согласно (3).

Следует отметить, что данная оценка позволяет сделать вывод о том, насколько именно распределение частей речи во фрагменте может соответствовать лингвистически корректному случаю. Для всестороннего оценивания фрагмента полезно комбинировать этот показатель с рассмотренной в (1) простейшей морфологической оценкой.

Рассмотрим пример. Проведем оценку следующего фрагмента: *расчет семантической нагруженности текстового фрагмента по четко определенной процедуре*. Воспользуемся следующим алгоритмом:



Очевидно, что данный фрагмент является хорошо оформленным предложением и должен давать значение семантической нагрузки близкое к единице. Для рассматриваемого нами в качестве примера фрагмента, расположение тегов частей речи выглядит



Здесь пунктиром показаны связи между морфологически согласованными словоформами, а соответствующие обозначенными схематически словами (С – существительные, П – прилагательные, Пр – предлог, Н – наречие) показано расположение частей речи, рассчитанное расстояние между которыми входит в расчет расстояний для вычисления лингвистической целостности фрагмента по существительным согласно формуле (3).

Рассматривая распределение существительных, можем составить следующую условную схему их расположения во фрагменте ($w_i, i = 1, 4$ – каждое из четырех существительных в рассматриваемом предложении):

$$w_1 \rightarrow +1 \rightarrow w_2 \rightarrow +1 \rightarrow w_3 \rightarrow +2 \rightarrow w_4.$$

Здесь $+n$ – это инкремент расстояния между рассматриваемыми частями речи (существительными). Приведенное выше распределение означает, что мы имеем последовательность из 5 расстояний, а именно (0, 1, 1, 2, 0). Следует заметить, что первый и последний элементы учитывают расстояние от начала и от конца фрагмента до ближайшего слова, которое принадлежит рассматриваемой части речи. Кроме того, при подсчете расстояний и количества слов целесообразно не учитывать второстепенные части речи.

Имея последовательность расстояний, вычисляем обратную дисперсию существительных по формуле (3) с использованием (4) и (2)

$$R(P_{Noun}) = 1 - kD(P_{Noun}) = 1 - \left(\frac{4+1}{4*(8-4)} \right)^2 \frac{1}{4+1} \sum_{i=0}^4 \left(s_i - \frac{8-4}{5} \right)^2 \approx 0.96.$$

Полученное значение можно считать достаточно высоким, чтобы считать исходный фрагмент хорошо обусловленным по распределению существительных. Проводя аналогичную оценку по прилагательным, можно прийти к выводу, что данный фрагмент хорошо обусловлен также по рас-

пределению прилагательных. Кроме того, оценка количества согласованных морфологически пар словоформ по формуле (1) также дает хороший результат, – пример выбора согласованных пар приведен на схеме распределения частей речи по данному фрагменту в виде пунктирных линий – любая система морфологического анализа может легко выделить эти пары по совпадению, например, именных характеристик (падеж и род). Как видно из проведенного анализа, приведенное в качестве примера предложение достаточно хорошо обусловлено с точки зрения распознающей системы.

Выводы. Для оценки лингвистической репрезентативности текстового фрагмента и его целостности как хорошо оформленного предложения рассмотрен метод, который учитывает при анализе основные естественно-языковые аспекты: синтаксис, семантику и морфологию. Разработан соответствующий алгоритм и реализующее его программное обеспечение, с помощью которого проводились эксперименты на реальном контенте.

Библиографические ссылки

1. **Lee D.-G.** Towards Language-Independent Sentence Boundary Detection. // Lee D.-G., Rim H.-C. – Computational Linguistics and Intelligent Text Processing, V. 2945/2004 – Springer Berlin, Heidelberg, 2004 – P. 142-145
2. **Hoojung C.** Unlexicalized Dependency Parser for Variable Word Order Languages Based on Local Contextual Pattern. // C. Hoojung, R. Hae-Chang – Computational Linguistics and Intelligent Text Processing, – V. 2945/2004 – Springer Berlin, Heidelberg, 2004 – P. 112-123
3. Построение полных парадигм слов русского языка на базе грамматического словаря А. А. Зализняка – <http://www.jurta.org/ru/nlp/rus/zaliz>
4. **Bolshakov I.** Computational linguistics: Models, Resources, Applications. / I.A. Bolshakov, A. Gelbukh - IPN-UNAM-FCE, 2004 – 186 P.
5. Russian dictionaries and morphology – <http://starling.rinet.ru/morpho.php?lan=en>

Надійшла до редколегії 18.05.11

Э.Э. Заманова

Государственный университет информатики и искусственного интеллекта

МОДЕЛЬ СИСТЕМЫ ПРИНЯТИЯ РЕШЕНИЙ ПРИ СОСТАВЛЕНИИ УЧЕБНОГО РАСПИСАНИЯ

Розглядається проблема створення системи прийняття рішень при складанні розкладу у ВНЗ. Пропонується початкова модель системи прийняття рішень при складанні розкладу, що використовує методи побудови розкладу на основі принципу обчислення свободи розташування окремого заняття в отриманому розкладі.

Рассматривается проблема создания системы принятия решений при составлении расписания в ВУЗе. Предлагается начальная модель системы принятия решений при составлении расписания, которая использует методы построения расписания на основе принципа вычисления свободы расположения отдельного занятия в полученном расписании.

The paper is devoted to the problem of a decision support system for scheduling a university. It is proposed to model the initial decision-making system for scheduling, which uses the methods of construction schedule based on the principle of freedom of calculating the location of the individual studies in the resulting schedule.

Ключевые слова: модель принятия решений, расписание, принцип вычисления свободы.

Введение. Целью данной работы является разработка начальной модели системы принятия решений при составлении учебного расписания, а также использование и дополнение методов построения расписания на основе принципа вычисления свободы расположения отдельного занятия.

Постановка задачи. Одной из сфер применения современных компьютерных технологий является автоматизированное составление расписаний для ВУЗа. Качество функционирования расписания во многом определяется решениями, принимаемыми на этапах календарного планирова-

ния и оперативного управления. Наряду с улучшением качества плановых решений все более жесткими становятся требования к сокращению сроков их выработки, повышению оперативности и гибкости управления.

С понятием расписания (календарного плана, распорядка дня, плана графика, режима работы и т.д.) каждый человек знакомится с самого начала осознанной жизни. Расписания движения самолетов, поездов, автобусов. Распорядки работы магазинов, мастерских, учреждений. Программы радио- и телепередач, расписание работы кинотеатров. Производственные планы на смену, сутки, неделю, квартал, год.

Описание алгоритма составления расписания занятий . В основе предлагаемого алгоритма составления расписания был положен принцип вычисления свободы расположения отдельного занятия в полученном расписании. То есть, существуют занятия, для проведения которых требуются специальные условия: аудитория большой вместимости, оборудование компьютерами – такие занятия имеют значительно меньшие возможности расположения в расписании. Кроме того, некоторые занятия могут проводиться только в определенный день или даже в определенное время, в которое его может провести преподаватель, приходящий из другого ВУЗа.

Из выше изложенного следует, что при составлении расписания в первую очередь следует сформировать список занятий, которые будут проводиться в ВУЗе. Для этого используются данные об учебной нагрузке всех потоков учащихся на требуемый семестр.

После составления списка занятий, необходимо произвести оценку свободы их расположения в расписании на основании особенностей и требований к их проведению, таких как: общее количество аудиторий в которых может быть проведено данное занятие, общее количество занятий отдельного потока (например СР_06) в неделю, а также общее количество занятий в неделю, проводимых преподавателем этого занятия. Такая оценка позволит сравнить вероятность невозможности проведения занятия при некотором стечении обстоятельств: занят преподаватель, заняты все подходящие аудитории, для группы в данный момент проводится другое занятие.

Полученные оценки используются для сортировки занятий по возрастанию их оценок свободы расположения.

При формировании расписания, в него в первую очередь добавляются занятия с наименьшим значением оценки свободы расположения. Таким образом, решается проблема «перекрытия» занятий, которая выражается в

том, что провести их невозможно, так как: занята группа, заняты требуемые аудитории, либо занят преподаватель.

При добавлении занятий в расписание производится поиск наиболее выгодной аудитории и времени для ее проведения. Для этого производится полный перебор вариантов проведения занятия в пространстве(аудитории) и времени(номер пары, день недели). В первую очередь происходит проверка возможности проведения занятия:

- а) не происходит «перекрывтия» занятий, в случае если оно произошло, то занятие не может быть проведено;
- б) аудитория оборудована всем необходимым для проведения занятия, например компьютерами, стендами для проведения экспериментов, проектором, и т. д.;
- в) количество рабочих мест в аудитории не меньше количества учащихся в группе.

Таким образом, происходит отсеивание заранее неверных вариантов расписания, когда для какого-либо занятия требуется аудитория оборудованная компьютерами, но эта аудитория уже занята другим занятием, которое не требует наличия компьютеров.

В случае, когда первые три условия выполняются, происходит оценка качества расположения занятия по нескольким критериям:

- а) появление окна в расписании группы студентов;
- б) появление окна в расписании преподавателя;
- в) избыточность количества мест в аудитории по отношению к количеству учащихся;
- г) проведение занятия в неудачное время, например четвертым или пятым по счету в этот день для этой группы студентов;
- д) исчезновение окна в расписании группы студентов;
- е) исчезновение окна в расписании преподавателя;
- ж) исчезновение окна в расписании использования аудитории.

Оценка качества расположения занятия по каждому из критериев может быть использована для получения некоторой общей оценки, необходимой в дальнейшем для выбора максимально выгодного времени и места проведения занятия.

Ход дальнейших разработок . Уже на начальной стадии находится программный модуль для диспетчера. Он будет иметь следующий вид рис.2. Такая организация программы позволяет оперативно вносить изменения в необходимых таблицах.

Необходимо отметить, что при разработке приближенных методов решения задач теории расписаний, весьма нетривиальными являются во-

просы оценки близости построенного расписания к искомому оптимальному. Поэтому, в дальнейших исследованиях этому вопросу будет уделено больше времени и внимания.

Не смотря на обилие программ, предназначенных для составления расписаний, только некоторые из них имеют функции автоматического составления наиболее оптимального расписания занятий. Это объясняется, прежде всего, сложностью решения данной задачи. Тем не менее такие программные продукты пользуются спросом, благодаря удобному интерфейсу пользователя, позволяющему существенно упростить работу диспетчера при составлении нового расписания занятий или изменения уже существующего, кроме того, в таких программных продуктах предусмотрена возможность анализа составленного расписания и составления отчетов, необходимых в работе диспетчера.

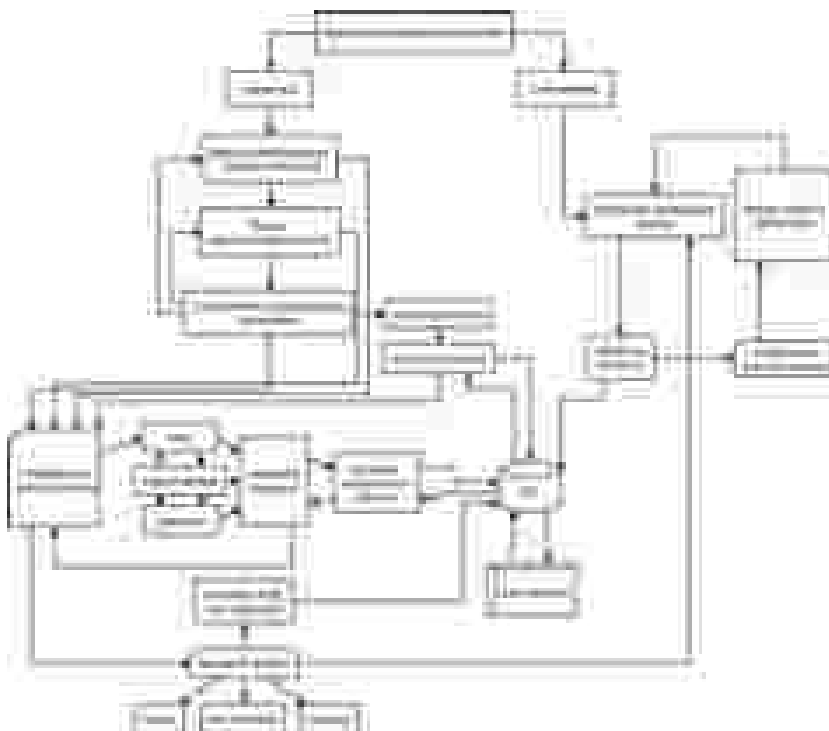


Рис.1. Структурная схема взаимодействия модулей программы

Модель данных – это совокупность соглашений о способах и средствах формализованного описания объектов и их связей, имеющих отношение к автоматизации процессов системы. Вид модели и используемые в ней типы структур данных отражают концепцию организации и обработки данных, используемую в СУБД, поддерживающей модель, или в языке системы программирования, на котором создается прикладная программа обработки данных.

В рамках решения поставленной задачи необходимо создание такой модели данных, при которой объем вспомогательной информации был бы минимальным, существовала принципиальная возможность многопользовательского доступа к данным и был бы обеспечен высокий уровень защиты данных.

Оптимальность построения распределения . Как это ни странно, в самой системе (заданной семеркой объектов) содержатся элементы, которые могут непосредственно влиять на оптимальность (в некоторых случаях даже на субъективную оптимальность) построенного распределения.

За счет выбора оператора частичного упорядочивания, заданного на множестве временных интервалов, можно влиять на характер распределения заданий по процессорам. Это непосредственно следует из алгоритма построения расписания.

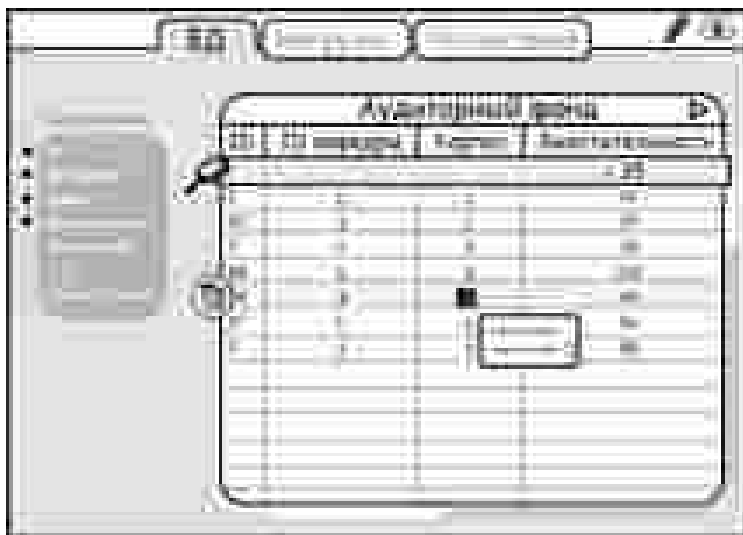


Рисунок 2 – Модель модуля программы составления учебного расписания

По ходу алгоритма, когда ищется допустимая ячейка трехмерной матрицы для задания, процедура построения последовательно проверяет все временные интервалы в составе матрицы, соответствующие предварительно заданному множеству R временных интервалов. На множестве временных интервалов задано отношение частичного порядка, с помощью которого оно и упорядочено. Следовательно, ячейки матрицы (в размерности временных ячеек) соответствуют временным интервалам в том порядке, котором они располагаются в упорядоченном множестве R .

Поэтому с помощью выбора отношения частичного порядка на множестве R можно упорядочить временные интервалы таким образом, чтобы при построении расписания они заполнялись необходимым образом.

Выводы. В заключении можно сказать, что система автоматического составления расписания значительно упрощает процесс составления расписания, влияет на производительность обучения и является на сегодняшний день одной из наиболее востребованных.

В ходе работы еще раз подтвердилась актуальность задачи оптимизации распределения ресурсов на примере составления расписания для ВУ-За, исходя из того, что на данный период оптимизированная система не существует, а лишь существуют предпосылки ее возникновения и многочисленные исследования ученых многих стран по поиску оптимального решения распределительных задач.

Также можно сказать, что существует необходимость в универсальной модели распределения ресурсов для задач различного типа, так как достаточно большое количество мелких задач из множества распределения ресурсов уже были решены, но только для определенных и конкретных узкоспециализированных задач.

В последующем требуется разработка унифицированного набора входных параметров такой модели для широкого класса задач и разработка методов оптимизации выходных параметров универсальной модели распределения ресурсов.

Бібліографічні посилання

1. **Коффман Э.Г.** Теория расписаний и вычислительные машины. — М., 1984. — С. 102–184.
2. **Танаев В.С., Шкурба В.В.** Введение в теорию расписаний. — М., — 1975. — С. 56–98.
3. **Левин В.И.** Структурно-логические методы в теории расписаний. — Пенза, 2006. — С. 85–94.

4. **Конвей Р.В.**, Максвелл В.Л., Миллер Л.В. Теория Расписаний. – М., 1975. – С. 201– 234.
5. **Сергиенко И.В.** Математические модели и методы решения задач дискретной оптимизации. – Киев, – 1988. – С. 75– 86.

Дата надходження до редколегії 31.12.10.

УДК 664.123.4:004.67

¹⁰О.М. Зігунов, В.Д. Кишенько

Національний університет харчових технологій

ПАТЕРН-АНАЛІЗ ВХІДНОЇ ІНФОРМАЦІЇ У ПІДСИСТЕМІ ТЕХНОЛОГІЧНОГО МОНІТОРИНГУ ДИФУЗІЙНОГО ВІДДІЛЕННЯ ЦУКРОВОГО ЗАВОДУ

Розглядаються методи обробки і аналізу даних, отриманих підсистемою технологічного моніторингу дифузійного відділення цукрового заводу. Показана можливість ефективного розпізнавання патернів типу «пик» і «повільна хвиля» за вейвлет-спектрами часових рядів.

Рассматриваются методы обработки и анализа данных, полученных подсистемой технологического мониторинга диффузионного отделения сахарного завода. Показана возможность эффективного распознавания паттернов типа «пик» и «медленная волна» по вейвлет-спектрам временных рядов.

Methods of processing and data analysis obtained by technological monitoring subsystem of diffusion department of sugar factory are considered. The possibility of effective recognition of patterns such as “peak” and “slow wave” by wavelet spectra of time series is shown.

Ключові слова: патерн-аналіз, система технологічного моніторингу, вейвлет-спектр, технологічний процес.

Вступ. У сучасних умовах одним з найважливіших факторів науково-технічної революції є широка інформатизація і автоматизація людської діяльності, створення і розвиток відповідних автоматизованих і інформаційних систем. Під автоматизованою системою розуміється [3] система, що складається з персоналу і комплексу засобів автоматизації його діяльності, що реалізує інформаційну технологію виконання встановлених функцій.

Сучасний рівень розвитку технологій породжує не тільки складні виробничі цикли взаємозалежних технологічних процесів (ТП), але і значне ускладнення самих ТП, ефективне керування якими неможливо з викорис-
--

танням стандартних SCADA_систем. Підвищення ефективності моніторингу стану складних промислових об'єктів можливо на основі розширення аналітичних інформаційних систем на рівень технологічного керування підприємством і введення додатково до SCADA _системи нової - аналітичної інформаційної системи, що перетворює загальний технологічний інформаційний потік до оптимального вигляду для ситуаційного аналізу.

Постановка задачі. Система технологічного моніторингу (СТМ) для забезпечення функціонування системи управління складними технологічними комплексами може виділити зміни, спричинені в результаті функціонування технологічного процесу, коли потрібна детальна інформація про різні коливання і зміни. Моніторинг включає до себе такі основні напрямки діяльності:

1. Спостереження за факторами, які впливають на довкілля та на стан технологічного процесу.
2. Оцінку фактичного стану технологічного процесу.
3. Прогноз стану технологічного процесу та оцінка цього стану.

Отже, технологічний моніторинг є багатоцільовою інформаційною системою. Його основні задачі: спостереження за станом об'єкта, оцінка та прогноз його стану; визначення інтенсивності різних впливів, виявлення факторів та джерел таких впливів.

У підсистемі технологічного моніторингу дифузійного відділення цукрового заводу використовуються датчики температури, маси, витрати, тиску, рівня, властивостей речовин, струмового навантаження, об'єднані в єдину мережу.

Варто також врахувати можливість впливу зовнішніх факторів на показання датчиків, що приводить до перекрученого представлення про стан об'єкта.

Метод розв'язування. Використання в СТМ мережі різнотипних датчиків змушує створювати складну систему збору, обробки і аналізу даних (рис. 1), що повинна:

- здійснювати прийом і форматування сигналів датчиків;
- формувати вектор характеристик стану об'єкта з використанням вектора сигналів датчиків. З метою усунення надмірності вектор характеристик оптимізується;
- зберігати історію об'єкта (дані, зібрані за весь період спостереження за об'єктом) у вигляді «знімків» (кадрів) поточного стану протягом декількох сезонів роботи цукрового заводу із використанням мето-

дів компактного представлення даних при збереженні можливості ретроспективного аналізу.

- «учитися» розпізнавати риси нормальної поведінки (зміни у стані) об'єкта і аномальних змін. Процес «навчання» повинен охоплювати природні життєві цикли об'єкта, пов'язані зі змінами зовнішніх впливів (температура, вологість і ін.) та режимів експлуатації. Навчання і аналіз змін — це два паралельних взаємозалежних процеси. На початковому етапі експлуатації системи неминучі фіктивні тривоги;
- аналізувати зміни з метою локалізації відхилень і оцінки їхньої значущості. Для класифікації необхідно створити і вести базу знань, що враховує досвід експлуатації аналогічних об'єктів, результати моделювання, думки експертів та ін.

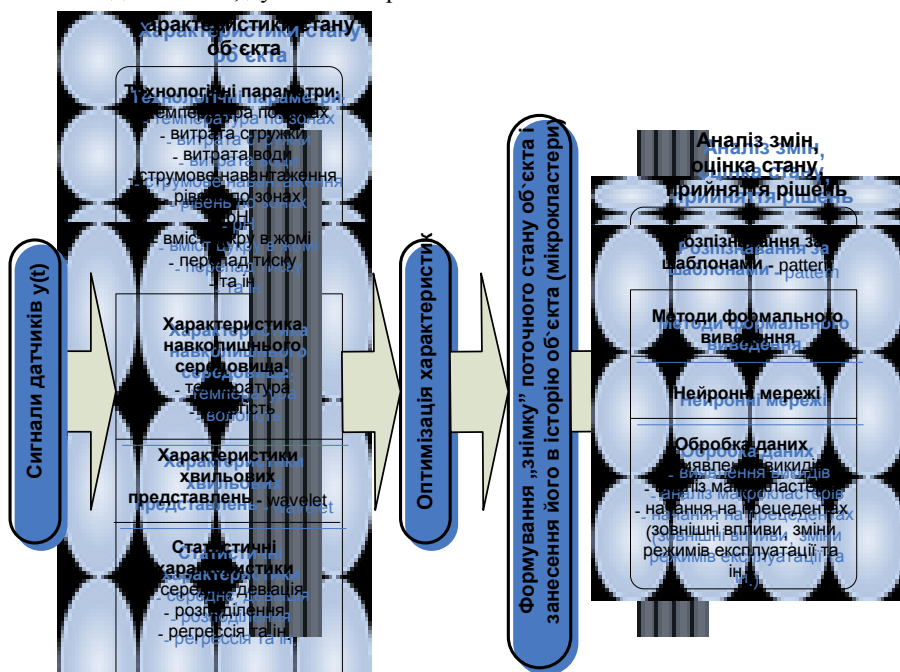


Рис. Підприємство з моніторингу дифузійного відділення цукрового заводу

Унаслідок недосконалості датчиків у їхніх показаннях є шумові складові, тому дані, що надходять до системи, повинні сприйматися в статистичному плані. Виходячи з того, що параметри розподілів сигналів заздалегідь невідомі, то варто намагатись використовувати універсальні оцінки, одержувані, наприклад, у вигляді декомпозиції вейвлетів (wavelets decomposition) [1] або ієрархії кластерів [2].

Для скорочення обсягу оброблюваних і збережених даних як характеристики стану об'єкта використовують агреговані, «сплавлені» показання декількох датчиків. Підставою для агрегування можуть служити функціональні зв'язки, виявлені при моделюванні роботи об'єкта, кореляційні зв'язки, зафіксовані у процесі спостережень за поведінкою об'єкта та ін. Мета агрегування – скорочення розмірностей у багатомірному просторі даних моніторингу. Відомими прийомами скорочення розмірностей є PCA (Principal Component Analysis) [4], MDS (Multidimensional Scaling) [5], SOM (Self Organizing Maps) [6], процес візуальної інтерактивної редукції розмірностей VHDR (Visual Hierarchical Dimension Reduction).

Історія об'єкта потрібна для виявлення змін у його стані і навчання СТМ. Вона може бути представлена у вигляді «знімків» – записів бази даних, що містять оцінки поточного стану об'єкта. Наприклад, це може бути набір мікрокластерів, модифікованих під ієрархічну кластеризацію. «Знімок» можна представити у вигляді кортежу $S = (D, H, T)$, де D – дата формування «знімка», H – інформація про стан об'єкта на час формування «знімка», T – додаткова інформація.

Відповідно до завдання стиснення масиву даних, що зберігаються, знімки можна зібрати в «піраміду» (pyramidal time frame), у якій пізніші знімки частіше заповнюють часову вісь, чим ранні. Можна використовувати логарифмічну шкалу, де часові інтервали між знімками змінюються за законом $\Delta t = \alpha^i$, де α – деяке ціле число $\alpha \geq 1$; i – ранг інтервалу.

Якщо інтервал часу T від початку процесу спостереження до поточного моменту t_c виражений у дискретах тактового сигналу, то він буде населений знімками рангів від 0 до $\log_{\alpha}(T)$.

Якщо зберігати тільки $(\alpha^m + 1)$ останніх знімків кожного рангу, то загальна кількість збережених знімків складе не більше $(\alpha^m + 1)\log_{\alpha}(T)$. Для цілей аналізу важлива точність апроксимації стану об'єкта в момент часу, що цікавить. Найбільш розрідженими є інтервали на початку процесу. Помилка апроксимації тут буде найбільшою. Оцінки можна одержати з наступних міркувань.

Серед збережених кадрів знайдуться знімки рангу r_i , які задовольняють співвідношення

$$\text{int}(\log_a T) - m = r_i,$$

де $\text{int}(\log_a T)$ – найближче найбільше ціле.

Це означає, що інтервал T розбитий на a^m інтервалів розміром не більше a^{r_i} тактів, і максимальна похибка при апроксимації буде в межах

$$\delta = 0,5(a^{r_i}/T).$$

Можливо подальше скорочення кількості знімків, що зберігаються. Так, недоцільно зберігати знімки, що відносяться до перехідних процесів, пов'язаних зі швидкою зміною температури навколишнього середовища. У найпростішому випадку можна відбирати знімки за мінімумом варіації показань датчиків у підбраному змінному вікні, однак можлива процедура, за якою аналізується мінливість мікрокластерів, що становлять знімок. Це «промивання» більш високого, глобального рівня (макроаналіз).

Вейвлет_спектр містить повну інформацію про сигнал. Характерною рисою вейвлет_спектра є те, що він дозволяє виявляти не тільки частотні характеристики сигналу, але і їхню локалізацію за часом. При цьому за зовнішнім виглядом вейвлет_спектру можна судити про подібність ділянок сигналу материнському вейвлету.

Вейвлет_аналіз знайшов широке застосування при діагностуванні станів об'єктів керування, оскільки його використання дозволяє адекватним чином досліджувати масштабно-інваріантну динаміку складних технічних систем. Перевагою вейвлет_аналізу також є можливість локальної оцінки різномасштабних частотних характеристик часових рядів, що особливо цінно при вирішенні задач розпізнавання порушень роботи об'єкта.

Діагностика на основі аналізу вхідної інформації у підсистемі технологічного моніторингу вимагає виявлення певних характерних рис – патернів. Їхня кількість і вид дозволяють судити про наявність або відсутність відхилень. Кожний патерн визначається набором ознак: частотою, амплітудою, тривалістю, повторюваністю.

Аналіз одержаних результатів. Розглянемо виявлення патернів типу «пік» і «повільна хвиля» за вейвлет_спектрами.

В якості вихідних даних використовувалися дані від датчиків системи автоматизації дифузійного відділення цукрового заводу, до складу якої входить підсистема технологічного моніторингу.

Розрахунок вейвлет_спектрів виконувався з використанням пакету Interactive Wavelets, розташованого в мережі Internet за адресою <http://ion.researchsystems.com/IONScript/wavelet/>. Оцінювалася можливість

виявлення особливостей вейвлет-спектрів, що відповідають різним патернам.

Характерним для ряду станів є поява в часовому ряді комбінацій «пік» – «повільна хвиля». Типовий запис патерну «пік» – «повільна хвиля» представлений на рис. 2. Також на рис. 2 наведений вейвлет-спектр цієї ділянки, обчислений для вейвлету Морле (1), який має аналітичний вигляд [1]:

$$\Psi(t) = e^{-\frac{t^2}{2}} e^{i2\pi t}. \quad (1)$$

Розрахунок вейвлет-коефіцієнтів сигналу $\{x(t), t=1, T\}$, відбувається за формулою (2)

$$W(a, b) = \frac{1}{\sqrt{a}} \sum_{t=0}^{T-1} x(t) \int_{-\infty}^{\infty} \psi^* \left(\frac{t-b}{a} \right) d\tau, \quad (2)$$

де ΔT – інтервал дискретності, a, b – параметри масштабу і часового зсуву відповідно, $\psi(t)$ – материнська вейвлет-функція, оператор $*$ позначає комплексне спряження.

На рис. 2 виділені ділянки сигналу з характерною рисою типу «пік» і ділянки вейвлет-спектра, що їм відповідають. Патерни типу «пік» мають високу частоту і тому відповідні їм ділянки вейвлет-спектру розташовані в області малих значень масштабного коефіцієнту a , який визначає величину розтягання вейвлету.

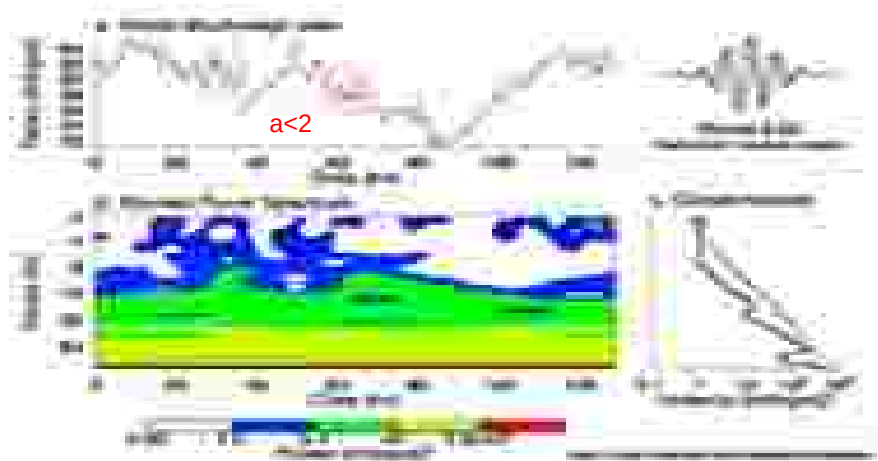


Рис. 2. Аналізований сигнал витрати дифузійного соку та його вейвлет-спектр, а) вихідні дані; б) вейвлет-спектр; в) глобальний вейвлет-спектр

Очевидно, що використання вейвлет-спектру може істотно спростити пошук піків у вхідних даних.

Підвищення наочності може бути досягнуте шляхом побудови ліній рівня вейвлет-спектру. Математично це означає, що при певному рівні здійснюється перетин поверхні в тривимірному просторі, що являє вейвлет-перетворення, площиною. У результаті формуються лінії рівня – ізолінії.

Іншим найважливішим типом патерна, який необхідно визначати в досліджуваних часових рядах – є повільна хвиля. Використання різних материнських вейвлетів показує, що найкращі результати по виявленню повільних хвиль досягаються при використанні Симлет-вейвлета 4-го порядку.

На рис. 3 наведено результати визначення повільних хвиль.

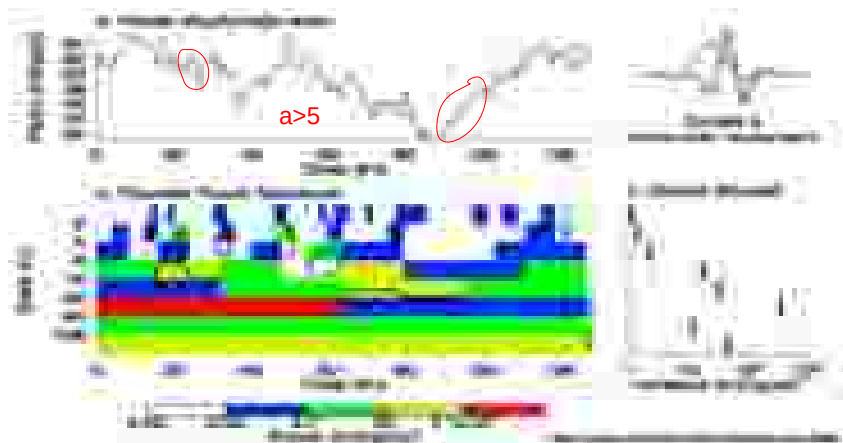


Рис. 3. Виявлення патернів типу «повільна хвиля» за допомогою Симлет-вейвлету.
а) вихідні дані; б) вейвлет-спектр; с) глобальний вейвлет-спектр

Як видно з рис. 3, запропонований у роботі підхід до виявлення патернів і певні параметри вейвлет-спектрів дозволяють здійснювати часову локалізацію патернів і обчислювати їхню кількість на заданому часовому інтервалі.

Висновки. У даній роботі показана принципова можливість ефективного розпізнавання патернів типу «пік» і «повільна хвиля» за вейвлет-спектрами часових рядів. Також показано, що перехід від вейвлет-спектру до ізоліній, які представляють собою лінії перетину поверхні в тривимірному просторі, що зображує вейвлет-спектр, площинами, дозволяє виявити зазначені патерни, здійснити їхню часову локалізацію і визначити кількість. Крім того, були визначені чисельні значення масштабних коефіцієнтів, необхідні для автоматизації розпізнавання кожного з аналізованих патернів. Подальшим напрямком роботи є аналіз інших патернів, таких як «гострі хвилі», а також розробка спеціального програмного забезпечення для обробки часових рядів на основі вейвлет-перетворення.

Бібліографічні посилання

1. **Витязев В.В.** Вейвлет-анализ временных рядов: Учеб. пособие / В.В. Витязев – СПб., 2001. – 58 с.
2. **Мандель И.Д.** Кластерный анализ. / И. Д. Мандель – М., 1988. – 176 с.

3. **Рябинин. И. А.** Надёжность и безопасность структурно сложных систем / И. А. Рябинин. – СПб., 2000. – 351 с.
4. **Jolliffe J.** Principal Component Analysis. –New York: Springer-Verlag, 1986. – 487 p.
5. **Mead A.** Review of the development of multidimensional scaling methods // The Statistician. – San Francisco: Wiley, 1992. - № 33. – p. 27-39.
6. **Kohonen T.** Self-Organizing Maps. - New York: Springer-Verlag, 1997. - 378 p.

Надійшла до редколегії 10.11.10

¹⁴О.М. Кісельова, О.О. Жильцова

Дніпропетровський національний університет імені Олеся Гончара

НЕОБХІДНІ ТА ДОСТАТНІ УМОВИ ОПТИМАЛЬНОСТІ В ЗАДАЧАХ ОПТИМАЛЬНОГО РОЗБИТТЯ МНОЖИН В ТЕРМІНАХ ФУНКЦІЙ МНОЖИН. ДВОЇСТА ЗАДАЧА

Наведено загальну постановку задачі мінімізації функції розбиття множини за умови обмежень у вигляді нерівностей. На базі отриманих раніше необхідних умов оптимальності, сформульовано та доведено достатні умови оптимальності. Також проведено аналіз аналога класичного функціоналу Лагранжа для поставленої задачі, доведено теореми двоїстості, описана двоїста задача.

Приведена общая постановка задачи минимизации функции разбиения множества при ограничениях в виде неравенства. На базе ранее полученных необходимых условий оптимальности, сформулированы и доказаны достаточные условия оптимальности. Также проведен анализ аналога классического функционала Лагранжа для поставленной задачи, доказаны теоремы двойственности, описана двойственная задача.

In the article, the general statement of the problem of set partitioning function minimization with the restrictions in the form of inequalities is presented. For this problem, on the base of the previously obtained necessary optimality conditions, the sufficient optimality conditions are formulated and proved. Also, the analysis of the analogue of the classical Lagrange functional is performed for the stated problem; the duality theorems are proved; the dual problem is described.

Ключові слова: оптимізація, розбиття множини, функція множин, необхідні умови оптимальності, достатні умови оптимальності, теорія двоїстості, функціонал Лагранжа.

Загальні відомості. Постановка задачі. Нехай Ω – неперервна, обмежена, вимірна за Лебегом множина з n -вимірною евклідовою простору $\mathbb{R}^n$.

Розбиттям множини $\Omega \subset \mathbb{R}^n$ на N підмножин назвемо систему її підмножин $\Omega_1, \dots, \Omega_N$, для яких виконуються три умови:

- 1) $\Omega_i \subseteq \Omega, \forall i = 1, \dots, N$;
- 2) $\bigcup_{i=1}^N \Omega_i = \Omega$;
- 3) $\text{mes}(\Omega_i \cap \Omega_k) = 0, \forall i \neq k, i, k = 1, \dots, N$,

де $\text{mes}(\cdot)$ – міра Лебега.

Позначимо через $\mathcal{R}_\Omega^N$ клас (простір) усіх можливих розбиттів множини Ω на N підмножин. При цьому розбиття, що відрізняються лише на множини нульової міри, вважають рівними елементами $\mathcal{R}_\Omega^N$.

В [1] було побудовано метричний простір на базі $\mathcal{R}_\Omega^N \subset \Sigma^N$, де $\Sigma^N = \Sigma \times \dots \times \Sigma$ – декартовий добуток N алгебр множини Ω , та вдалося встановити такі його властивості як обмеженість, щільність у собі, хаусдорфовість, замкненість.

Для цього простору було сформульовану наступну задачу [2].

Нехай функції $F, G_1, \dots, G_m: \mathcal{R}_\Omega^N \rightarrow \mathbb{R}$ – псевдонеперервні функції множин.

Задача 1. Знайти мінімум функції множин F на

$U = \{(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N: G_i(\Omega_1, \dots, \Omega_N) \leq 0, i = 1, \dots, m\}$, тобто

$$F(\Omega_1, \dots, \Omega_N) \rightarrow \min$$

при умовах

$$\begin{aligned} G_i(\Omega_1, \dots, \Omega_N) &\leq 0, i = 1, \dots, m, \\ (\Omega_1, \dots, \Omega_N) &\in \mathcal{R}_\Omega^N. \end{aligned}$$

Також було доведено [2] необхідну умову оптимальності для задачі 1.

Теорема 1 (Необхідна умова оптимальності для задачі 1). Нехай (Ω, Σ, μ) – скінчений безатомний простір з мірою і $F, G_1, \dots, G_m: \mathcal{R}_\Omega^N \rightarrow \mathbb{R}$ – диференційовані функції множин на $(\Omega_1^*, \dots, \Omega_N^*) \in \mathcal{R}_\Omega^N$. Якщо $(\Omega_1^*, \dots, \Omega_N^*)$ – локальний мінімум задачі 1, то існують числа $\lambda_0, \lambda_1, \dots, \lambda_m$ такі, що

$$\sum_{i=1}^N (X_i D F_{\Omega_i, \dots, \Omega_N}^i + \sum_{j=1}^m \lambda_j^i D G_{\Omega_i, \dots, \Omega_N}^{ij} \cdot X_{\Omega_i} - X_{\Omega_i}) \geq 0 \quad (3)$$

для усіх $(\Omega_1, \dots, \Omega_N) \in \mathbb{R}_\Omega^N$,

$$X_i G_i(\Omega_1^i, \dots, \Omega_N^i) = 0, \quad i = 1, \dots, m, \quad (4)$$

$$X_0, X_1, \dots, X_m \geq 0, \quad (5)$$

$$G_i(\Omega_1^i, \dots, \Omega_N^i) \leq 0, \quad i = 1, \dots, m, \quad (6)$$

$$\lambda_j^i \text{ не всі дорівнюють нулю}, \quad (7)$$

де $D G_{\Omega_i, \dots, \Omega_N}^{ij}$ – частинна похідна функції множин G_i за i -м аргументом на $(\Omega_1^i, \dots, \Omega_N^i)$, $i = 1, \dots, N$, $j = 1, \dots, m$.

Будемо казати, що $(\Omega_1^i, \dots, \Omega_N^i)$ задовольняє умову регулярності, якщо існує $(R_1, \dots, R_N) \in \mathbb{R}_\Omega^N$, для якого

$$G_j(\Omega_1^i, \dots, \Omega_N^i) + \sum_{i=1}^N (D G_{\Omega_i, \dots, \Omega_N}^{ij} \cdot X_{R_i} - X_{\Omega_i}) < 0, \quad j = 1, \dots, m.$$

Наслідок 1. Нехай виконані всі умови теореми 1 і $(\Omega_1^i, \dots, \Omega_N^i)$ задовольняє умову регулярності. Тоді існують $\lambda_1^i, \dots, \lambda_m^i$, які в сукупності з $\lambda_0^i = 1$ задовольняють умови (3)–(7).

Достатні умови оптимальності для задачі 1. Для формулювання достатніх умов оптимальності введемо та дослідимо поняття опуклості функції μ множин.

Наступне означення базується на дослідженнях щодо опуклості комбінації (суміші) множин, проведених у [3].

Нехай $F: \Sigma^N \rightarrow \mathbb{R}$. Будемо казати, що F є опуклою функцією множин, якщо для будь-яких $\lambda \in [0; 1]$, $(R_1, \dots, R_N), (S_1, \dots, S_N) \in \Sigma^N$ і для будь-яких послідовностей множин $\{R_k^i\} \subset R_i \setminus S_i$ і $\{S_k^i\} \subset S_i \setminus R_i$, $k = 1, 2, \dots$, таких, що $X_{R_k^i} \xrightarrow{\text{df}} \lambda X_{R_i \setminus S_i}$ і $X_{S_k^i} \xrightarrow{\text{df}} \lambda X_{S_i \setminus R_i}$, $i = 1, \dots, N$, виконується нерівність

$$\overline{\lim_{n \rightarrow \infty}} F[R_1^n \cup S_1^n \cup (R_1 \cap S_1), \dots, R_N^n \cup S_N^n \cup (R_N \cap S_N)] \leq \lambda F(R_1, \dots, R_N) + (1 - \lambda) F(S_1, \dots, S_N) \quad (8)$$

Позначення $\xrightarrow{w}$ відповідає слабкій* збіжності елементів у $L_\infty(\Omega, \Sigma, \mu)$.

У роботі [3] доведено наступну теорему.

Теорема 2. Нехай (Ω, Σ, μ) – скінченний безатомний простір з мірою і функція множин $F: \Sigma^N \rightarrow \mathbb{R}$ диференційована на Σ^N . Якщо $L_1(\Omega, \Sigma, \mu)$ сепарабельно і F опукла, то для всіх $(R_1, \dots, R_N), (S_1, \dots, S_N) \in \Sigma^N$

$$\sum_{i=1}^N (DF_{R_1, \dots, R_N}^i X_{S_i} - X_{R_i}) \leq F(S_1, \dots, S_N) - F(R_1, \dots, R_N). \quad (9)$$

Очевидним наслідком з цієї теореми є наступне твердження.

Наслідок. Нехай (Ω, Σ, μ) – скінченний безатомний простір з мірою і функція множин $F: \Sigma^N \rightarrow \mathbb{R}$ диференційована на Σ^N . Якщо $L_1(\Omega, \Sigma, \mu)$ сепарабельно і F опукла, то для всіх $(R_1, \dots, R_N), (S_1, \dots, S_N) \in \mathcal{R}_N^N$

$$\sum_{i=1}^N (DF_{R_1, \dots, R_N}^i X_{S_i} - X_{R_i}) \leq F(S_1, \dots, S_N) - F(R_1, \dots, R_N). \quad (10)$$

Доведемо наступну теорему.

Теорема 3 (Достатні умови оптимальності для задачі 1). Нехай (Ω, Σ, μ) – скінченний безатомний простір з мірою і сепарабельним $L_1(\Omega, \Sigma, \mu)$. Припустимо, що функції множин $F, G_1, \dots, G_m: \Sigma^N \rightarrow \mathbb{R}$ диференційовані на $(\Omega_1^i, \dots, \Omega_N^i) \in \mathcal{R}_N^N$ і опуклі. Якщо існують числа $\lambda_1, \dots, \lambda_m$, такі, що $(\Omega_1^i, \dots, \Omega_N^i)$ і $\lambda_1, \dots, \lambda_m$ в сукупності з $\lambda_0 = 1$ задовольняють умови (3)–(7), то $(\Omega_1^i, \dots, \Omega_N^i)$ є глобальним мінімумом для задачі 1.

Доведення. Припустимо, що деякий елемент $(\Omega_1, \dots, \Omega_m) \in \mathcal{P}_0^N$ задовольняє умову $G_i(\Omega_1, \dots, \Omega_m) \leq 0, i = 1, \dots, m$. Тоді з опуклості F і наслідку з теореми 2 маємо:

$$F(\Omega_1, \dots, \Omega_m) \geq F(\Omega_1^*, \dots, \Omega_m^*) + \sum_{i=1}^m \langle DF_{\Omega_1^* \dots \Omega_m^*}^i \cdot X_{\Omega_i} - X_{\Omega_i^*} \rangle. \quad (11)$$

Враховуючи нерівність (3) з $X_{\Omega_i} = 1$, отримаємо

$$F(\Omega_1, \dots, \Omega_m) \geq F(\Omega_1^*, \dots, \Omega_m^*) + \sum_{j=1}^m \sum_{i=1}^n \lambda_j^i \langle DG_{\Omega_1^* \dots \Omega_m^*}^{ij} \cdot X_{\Omega_i} - X_{\Omega_i^*} \rangle. \quad (12)$$

Далі, використовуючи в останній нерівності опуклість G_j і умову (4), отримаємо

$$F(\Omega_1, \dots, \Omega_m) \geq F(\Omega_1^*, \dots, \Omega_m^*) + \sum_{j=1}^m \sum_{i=1}^n \lambda_j^i G_j(\Omega_1, \dots, \Omega_m).$$

Остаточно, (5) і той факт, що $G_i(\Omega_1, \dots, \Omega_m) \leq 0, i = 1, \dots, m$, дають нам

$$F(\Omega_1, \dots, \Omega_m) \geq F(\Omega_1^*, \dots, \Omega_m^*),$$

що і потрібно було довести.

Теорема доведена.

Двоїста задача. Подані умови оптимальності можуть бути використані для знаходження розв'язку задачі 1, задля цього вони навіть були конкретизовані [2]. Однак залишається питання про знаходження чисел $\lambda_1^*, \dots, \lambda_m^*$. Аналогічно загальній теорії розв'язання екстремальних задач для класичних функцій та функціоналів [4], побудуємо та проведемо дослідження двоїстої задачі для задачі 1.

Побудуємо функціонал, аналогічний функціоналу Лагранжа, для задачі 1:

$$L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)) = F(\Omega_1, \dots, \Omega_N) + \sum_{i=1}^m \lambda_i G_i(\Omega_1, \dots, \Omega_N) \quad (13)$$

Сідловою точкою функціонала (13) назвемо точку

$$(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_0,$$

де $\Lambda_0 = \{(\lambda_1, \dots, \lambda_m); \lambda_1 \geq 0, \dots, \lambda_m \geq 0\}$, якщо виконується:

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1, \dots, \lambda_m)) \leq L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) \leq L((\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*)) \quad (14)$$

для будь-яких $(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N$ та $(\lambda_1^*, \dots, \lambda_m^*) \in \Lambda_0$.

Лема 1. Для того, щоб точка $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_0$ була сідловою точкою функціонала (13), необхідно і достатньо, щоб виконувались наступні умови:

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) \leq L((\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*)), \forall (\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N, \quad (15)$$

$$\lambda_i^* G_i(\Omega_1^*, \dots, \Omega_N^*) = 0, \quad i = 1, \dots, m, \quad (16)$$

$$G_i(\Omega_1^*, \dots, \Omega_N^*) \leq 0, \quad i = 1, \dots, m. \quad (17)$$

Доведення. Необхідність. Нехай $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_0$ - сідлова точка функціонала (13). Тоді (15) є правою частиною нерівності (14). Покажемо, що виконуються умови (16) та (17). Розглянемо ліву частину нерівності (14):

$$F(\Omega_1^*, \dots, \Omega_N^*) + \sum_{i=1}^m \lambda_i^* G_i(\Omega_1^*, \dots, \Omega_N^*) \leq F(\Omega_1^*, \dots, \Omega_N^*) + \sum_{i=1}^m \lambda_i^* G_i(\Omega_1^*, \dots, \Omega_N^*),$$

$$\forall (\lambda_1, \dots, \lambda_m) \in \Lambda_0.$$

Звідси

$$\sum_{i=1}^m (\lambda_i^* - \lambda_i) G_i(\Omega_1^*, \dots, \Omega_n^*) \geq 0, \forall (\lambda_1, \dots, \lambda_m) \in \Lambda_0 \quad (18)$$

Покажемо, що виконується (17). Для цього візьмемо $(\lambda_1, \dots, \lambda_m)$, де $\lambda_j = \lambda_j^* + 1$ при деякому j , $1 \leq j \leq m$, та $\lambda_i = \lambda_i^*$ для усіх $i \neq j$. Вочевидь, обрана точка належить Λ_0 . З (18) отримаємо $(-1)G_j(\Omega_1^*, \dots, \Omega_n^*) \geq 0$, тобто $G_j(\Omega_1^*, \dots, \Omega_n^*) \leq 0$ для усіх $1 \leq j \leq m$.

Візьмемо точку $(\lambda_1, \dots, \lambda_m)$ з координатами $\lambda_j = 0$ при деякому j , $1 \leq j \leq m$, та $\lambda_i = \lambda_i^*$ для усіх $i \neq j$. Вочевидь, обрана точка належить Λ_0 . З (18) отримаємо $0 \leq \lambda_j^* G_j(\Omega_1^*, \dots, \Omega_n^*)$. Але ж $\lambda_j^* \geq 0$, $G_j(\Omega_1^*, \dots, \Omega_n^*) \leq 0$ при $j = 1, \dots, m$, тому остання нерівність можлива лише при $\lambda_j^* G_j(\Omega_1^*, \dots, \Omega_n^*) = 0$, $i = 1, \dots, m$. Таким чином, продемонстроване виконання усіх співвідношень (15)-(17).

Достатність. Нехай для деякої точки $(\Omega_1^*, \dots, \Omega_n^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathbb{R}_+^n \times \Lambda_0$ виконуються співвідношення (15)-(17). Покажемо, що тоді ця точка є сидловою. З (15) випливає виконання правої частини нерівності (14). Покажемо виконання лівої частини (14).

З умов (16),(17) випливає, що

$$(\lambda_i^* - \lambda_i) G_i(\Omega_1^*, \dots, \Omega_n^*) = 0 \quad (19)$$

для всіх тих i , для яких $G_i(\Omega_1^*, \dots, \Omega_n^*) = 0$, а якщо $G_i(\Omega_1^*, \dots, \Omega_n^*) < 0$ при деякому i , то з (16) отримаємо, що $\lambda_i^* = 0$. Отже в останньому випадку $(\lambda_i^* - \lambda_i) G_i(\Omega_1^*, \dots, \Omega_n^*) = -\lambda_i^* G_i(\Omega_1^*, \dots, \Omega_n^*) \geq 0$. Просумуємо ці нерівності із рівностями (19) та отримаємо

$$\sum_{i=1}^m (\lambda_i^* - \lambda_i) G_i(\Omega_1^*, \dots, \Omega_n^*) \geq 0, \forall (\lambda_1, \dots, \lambda_m) \in \Lambda_0.$$

Додамо до обох частин цієї нерівності $F(\Omega_1^*, \dots, \Omega_n^*)$ та отримаємо ліву частину нерівності (14). Доведено.

Лема 2. Нехай функції $F, G_1, \dots, G_m: \Sigma^N \rightarrow \mathbb{R}$ диференційовані та опуклі функції. Тоді для того, щоб $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_0$ була сідловою точкою функціонала (13), необхідно та достатньо, щоб виконувалися умови (3)-(7) для $\lambda_1^*, \dots, \lambda_m^*$ у сукупності із $\lambda_0 = 1$.

Доведення. Дійсно, за таких умов функціонал $L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)) = F(\Omega_1, \dots, \Omega_N) + \sum_{i=1}^m \lambda_i G_i(\Omega_1, \dots, \Omega_N)$ є опуклим при даних $\lambda_1^*, \dots, \lambda_m^*$. Для нього можна сформулювати задачу безумовної мінімізації на базі задачі 1. Вочевидь, для неї буде виконуватися умова регулярності. Тоді необхідні та достатні умови леми є прямим наслідком теорем 1 та 3, застосованих для згаданої задачі безумовної оптимізації.

Отже, для того щоб виконувалась нерівність

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) \leq L((\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*)), \forall (\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N,$$

необхідно та достатньо виконання умов (3)-(7). Ця нерівність у поєднанні із умовами (4) та (6) є, в свою чергу, за лемою 1, необхідною та достатньою умовою того, що $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_0$ є сідловою точкою функціонала (13).

Теорема 4. Нехай $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_0$ сідлова точка функціонала (13). Тоді $L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) = F(\Omega_1^*, \dots, \Omega_N^*), (\Omega_1^*, \dots, \Omega_N^*)$ є розв'язком задачі 1.

Доведення. З умов (17) випливає, що $(\Omega_1^*, \dots, \Omega_N^*)$ належить до припустимої множини задачі 1. З умови (16) легко бачити, що $L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) = F(\Omega_1^*, \dots, \Omega_N^*)$. Тоді нерівність (15) набуває вигляду:

$$\begin{aligned} F(\Omega_1^*, \dots, \Omega_N^*) &\leq L(\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*) \\ &= F(\Omega_1, \dots, \Omega_N) + \sum_{i=1}^m \lambda_i^* G_i(\Omega_1, \dots, \Omega_N). \end{aligned}$$

$$\forall (\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N.$$

Але якщо $(\Omega_1, \dots, \Omega_N)$ належить припустимій множині, то

$$\sum_{i=1}^m \lambda_i^* G_i(\Omega_1, \dots, \Omega_N) \leq 0.$$

Отже

$$F(\Omega_1^*, \dots, \Omega_N^*) \leq L(\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*) \leq F(\Omega_1, \dots, \Omega_N),$$

тобто $(\Omega_1^*, \dots, \Omega_N^*)$ є розв'язком задачі 1.

Теорема 5. Нехай функції $F, G_1, \dots, G_m: \Sigma^N \rightarrow \mathbb{R}$ диференційовані та опуклі функції, та для точки $(\Omega_1^*, \dots, \Omega_N^*) \in \mathcal{H}_\Omega$ у відповідній задачі 1 виконуються умови регулярності. Тоді, якщо $(\Omega_1^*, \dots, \Omega_N^*) \in \mathcal{H}_\Omega$ мінімум функції F у задачі 1, то існують невід'ємні числа $\lambda_1^*, \dots, \lambda_m^*$ такі, що $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{H}_\Omega^N \times \Lambda_0$ є сідловою точкою функціонала (13).

Доведення. З теореми 1 випливає, що існують невід'ємні числа $\lambda_1^*, \dots, \lambda_m^*$ для яких виконуються умови (3) -(7). У свою чергу, за лемою 2, виконання цих умов для пари $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{H}_\Omega^N \times \Lambda_0$ означає, що вона є сідловою точкою функціонала (13). Доведено.

Введемо у розгляд функцію

$$X(\Omega_1, \dots, \Omega_N) = \max_{\substack{(\lambda_1, \dots, \lambda_m) \in \Lambda_0 \\ (\lambda_1, \dots, \lambda_m) \in \mathcal{H}_\Omega^N}} L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)), (\Omega_1, \dots, \Omega_N) \in \mathcal{H}_\Omega.$$

Позначимо $U = \{(\Omega_1, \dots, \Omega_N) \in \mathcal{H}_\Omega^N: G_i(\Omega_1, \dots, \Omega_N) \leq 0, i = 1, \dots, m\}$ – припустима множина для задачі 1.

Фактично

$$X(\Omega_1, \dots, \Omega_N) = \begin{cases} F(\Omega_1, \dots, \Omega_N), \forall (\Omega_1, \dots, \Omega_N) \in U, \\ \infty, \forall (\Omega_1, \dots, \Omega_N) \in \mathcal{H}_\Omega^N \setminus U. \end{cases}$$

Це означає, що

$$\min_{\mathcal{H}_\Omega^N} X(\Omega_1, \dots, \Omega_N) = \min_U F(\Omega_1, \dots, \Omega_N).$$

Задачу 1, таким чином, можна переписати у рівносильному вигляді:

$$X(\Omega_1, \dots, \Omega_N) \rightarrow \min; (\Omega_1, \dots, \Omega_N) \in \mathcal{H}_\Omega^N. \quad (20)$$

Будемо вважати, що функція $F(\Omega_1, \dots, \Omega_N)$ є обмеженою знизу на припустимій множині та множина розв'язків задачі 1 не порожня. За цих умов множини розв'язків задачі 1 та задачі (20) співпадають.

Розглянемо також функцію

$$\Psi(\lambda_1, \dots, \lambda_m) = \min_{\substack{(\Omega_1, \dots, \Omega_N) \in \mathcal{N}_N^N \\ (\lambda_1, \dots, \lambda_m) \in \Lambda_\Psi}} L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)).$$

Та розглянемо задачу

$$\Psi(\lambda_1, \dots, \lambda_m) \rightarrow \max; \quad (\lambda_1, \dots, \lambda_m) \in \Lambda_\Psi \quad (21)$$

Як заведено, назвемо задачу (21) **двоїстою задачею** для основної задачі 1, а змінні $\lambda_1, \dots, \lambda_m$ - **двоїстими змінними**.

Позначимо

$$\begin{aligned} \max_{\Lambda_\Psi} \Psi(\lambda_1, \dots, \lambda_m) &= \Psi^*, \\ \Lambda^* &= \{(\lambda_1, \dots, \lambda_m) \in \Lambda_\Psi; \Psi(\lambda_1, \dots, \lambda_m) = \Psi^*\}. \end{aligned}$$

Для розглянутих величин виконуються нерівності:

$$\begin{aligned} \Psi(\lambda_1, \dots, \lambda_m) &\leq \Psi^* \leq F_* \leq X(\Omega_1, \dots, \Omega_N), \\ (\Omega_1, \dots, \Omega_N) &\in \mathcal{N}_N^N, (\lambda_1, \dots, \lambda_m) \in \Lambda_\Psi. \end{aligned} \quad (22)$$

Дійсно,

$$\begin{aligned} \Psi(\lambda_1, \dots, \lambda_m) &= \min_{(\Omega_1, \dots, \Omega_N) \in \mathcal{N}_N^N} L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)) \\ &\leq L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)) \end{aligned}$$

при будь-яких $(\Omega_1, \dots, \Omega_N) \in \mathcal{N}_N^N$ та $(\lambda_1, \dots, \lambda_m) \in \Lambda_\Psi$. Тому

$$\begin{aligned} \Psi^* = \max_{\Lambda_\Psi} \Psi(\lambda_1, \dots, \lambda_m) &\leq \max_{\Lambda_\Psi} L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)) \\ &= X(\Omega_1, \dots, \Omega_N), \\ &\quad \forall (\Omega_1, \dots, \Omega_N) \in \mathcal{N}_N^N. \end{aligned}$$

Переходячи до мінімуму по $(\Omega_1, \dots, \Omega_N) \in \mathcal{N}_N^N$ в останній нерівності, отримаємо $\Psi^* \leq F_*$, звідки випливають нерівності (22).

Теорема 6. Для того, щоб задача 1 та задача (21) мали розв'язки, та $F_* = \Psi^*$, необхідно та достатньо, щоб функція $L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m))$, $(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N, (\lambda_1, \dots, \lambda_m) \in \Lambda_m$ мала сідлову точку на $\mathcal{R}_N^N \times \Lambda_m$. Множина сідлових точок цієї функції на $\mathcal{R}_N^N \times \Lambda_m$ співпадає із декартовим добутком множин розв'язків задач 1 та (21).

Доведення. Необхідність. Нехай задача 1 та задача (21) мають розв'язки, та $F_* = \Psi^*$. Візьмемо довільну пару розв'язків $(\Omega_1^*, \dots, \Omega_N^*)$ та $(\lambda_1^*, \dots, \lambda_m^*)$ та покажемо, що $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)$ - сідлова точка. Маємо

$$\begin{aligned} \Psi^* = \Psi(\lambda_1^*, \dots, \lambda_m^*) &= \min_{(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N} L((\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*)) \\ &\leq L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) \\ &\leq \max_{(\lambda_1, \dots, \lambda_m) \in \Lambda_m} L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1, \dots, \lambda_m)) = X(\Omega_1^*, \dots, \Omega_N^*) \\ &= F_*. \end{aligned}$$

За умовою $F_* = \Psi^*$, отже всі попередні нерівності стають рівностями:

$$\begin{aligned} L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) &= \min_{(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N} L((\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*)) \\ &= \max_{(\lambda_1, \dots, \lambda_m) \in \Lambda_m} L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1, \dots, \lambda_m)) = F_*. \end{aligned}$$

З цього випливає виконання нерівностей (14), тобто $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)$ сідлова точка за означенням. Таким чином, ми продемонстрували, що декартовий добуток множин розв'язків задач 1 та (21) належать множині сідлових точок функціонала $L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m))$ на $(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N, (\lambda_1, \dots, \lambda_m) \in \Lambda_m$.

Достатність. Нехай $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_N^N \times \Lambda_m$ - сідлова точка функціонала $L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m))$ на $(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_N^N, (\lambda_1, \dots, \lambda_m) \in \Lambda_m$. Згідно з нерівностями (14) це означає, що

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1, \dots, \lambda_m)) \leq L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)), (\lambda_1, \dots, \lambda_m) \in \Lambda_m.$$

Крім того,

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) \leq L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)), (\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N.$$

Отже

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) = \min_{(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N} L((\Omega_1, \dots, \Omega_N), (\lambda_1^*, \dots, \lambda_m^*)) = \Psi(\lambda_1^*, \dots, \lambda_m^*).$$

З останнього співвідношення та нерівностей (22) випливає

$$L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)) = \Psi(\lambda_1^*, \dots, \lambda_m^*) \leq \Psi^* \leq F_* \leq X(\Omega_1^*, \dots, \Omega_N^*) = L((\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*)),$$

тобто, в останньому записі всі знаки «менш або дорівнює» можна замінити на знаки рівності. Отже тим самим встановлено, що множина сідловин точок функціонала $L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m))$ на $(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N, (\lambda_1, \dots, \lambda_m) \in \Lambda_\lambda$ належить декартовому добутку множин розв'язків задач 1 та (21).

Наслідок 1. Наступні чотири твердження є рівносильними:

- 1) $(\Omega_1^*, \dots, \Omega_N^*), (\lambda_1^*, \dots, \lambda_m^*) \in \mathcal{R}_\Omega^N \times \Lambda_\lambda$ - сідова точка функціонала $L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m))$ на $(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N, (\lambda_1, \dots, \lambda_m) \in \Lambda_\lambda$;
- 2) Задача 1 та задача (21) мають розв'язки, та $F_* = \Psi^*$;
- 3) Існують точки $(\Omega_1^*, \dots, \Omega_N^*) \in \mathcal{R}_\Omega^N$ та $(\lambda_1^*, \dots, \lambda_m^*) \in \Lambda_\lambda$ такі, що $X(\Omega_1^*, \dots, \Omega_N^*) = \Psi(\lambda_1^*, \dots, \lambda_m^*)$;

4) Справедливою є рівність

$$\max_{(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N} \min_{(\lambda_1, \dots, \lambda_m) \in \Lambda_\lambda} L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)) = \min_{(\Omega_1, \dots, \Omega_N) \in \mathcal{R}_\Omega^N} \max_{(\lambda_1, \dots, \lambda_m) \in \Lambda_\lambda} L((\Omega_1, \dots, \Omega_N), (\lambda_1, \dots, \lambda_m)).$$

Задача (21), переписана у вигляді

$$-\Psi(\lambda_1, \dots, \lambda_m) \rightarrow \min; (\lambda_1, \dots, \lambda_m) \in \Lambda_\lambda,$$

є задачею опуклого програмування. Її розв'язок допоможе знайти розв'язок основної задачі згідно необхідних та достатніх умов оптимальності, встановлених в теоремах 1 та 3. Зауважимо, що ці умови були конкретизовані у попередніх дослідженнях [2].

Бібліографічні посилання:

1. **Киселева Е.М.** Некоторые свойства пространства разбиений непрерывного множества / Е. М. Киселева, А.А. Жильцова // *Питання прикладної математики та математичного моделювання* : Зб. наук. праць –2008. – С. 84-91.
2. **Киселева Е.М.** Необходимые условия оптимальности для непрерывных задач разбиения множества в терминах теории функций множеств / Е.М. Киселева, А.А. Жильцова // *Проблемы управления и информатики* – 2008, №6. – С. 55-66.
3. **Кисельова О.М.** Елементи теорії функцій множин: Навч. посіб. / О.М. Кисельова, Л.Л. Гарт. – Дн.ск, 2006. – 176 с.
4. **Васильев Ф.П.** Численные методы решения экстремальных задач. /Ф.П. Васильев – М., 1988.

Надійшла до редколегії 20.05.2011

¹²Е.М. Киселева, Е.П. Заблоцкая

Днепропетровский национальный университет имени Олеся Гончара

НАХОЖДЕНИЕ МИНИМАЛЬНОГО РАДИУСА ПОКРЫТИЯ И ОПТИМАЛЬНЫХ МЕСТ РАСПОЛОЖЕНИЯ ТОЧЕК W_i - F_i ДЛЯ НЕЛИНЕЙНЫХ ФУНКЦИЙ СПРОСА

Розглядається вплив вибору різних функцій попиту на значення мінімального радіусу дії базових станцій W_i - F_i . Приведен алгоритм рішення неперервної задачі оптимального покриття множин із застосуванням методів оптимального розбиття і r -алгоритма Н. З.Шора для декількох неперервних функцій попиту.

Рассматривается влияние различных функций спроса на значение минимального радиуса действия базовых станций W_i - F_i . Приведен алгоритм решения непрерывной задачи оптимального покрытия множеств с использованием методов оптимального разбиения и r -алгоритма Н. З.Шора для нескольких непрерывных функций спроса.

The effect of different demand functions on value of the minimum radius of action of base W_i - F_i stations is considered. The algorithm of solving continuous set covering problem using methods of set partitioning and Shor's r -algorithm for a few continuous demand functions is cited.

Ключевые слова: оптимальное разбиение множеств, оптимальное покрытие множества, r -алгоритм Шора, непрерывная функция спроса.

Введение. Двадцать первый век характерен развитием высоких технологий и средств связи и передачи данных. Одним из таких серьезных прорывов стало объединение всех возможных научных достижений в области обработки радиосигнала, микропроцессорной техники, систем передачи и приема радиосигналов в единый унифицированный стандарт известный под названием W_i - F_i (IEEE 802.11 a\b\g). Установка W_i - F_i невозможна без грамотного выбора оптимальных мест расположения точек доступа и прочего оборудования. Чаще всего установка точки доступа W_i - F_i применяется для решения следующих задач:

- организация беспроводной локальной сети предприятия или офиса;
- предоставление доступа мобильных устройств (ноутбуки, принтеры и т.д.) к уже существующей проводной локальной сети;
- создание так называемых «горячих точек» для бесплатного пользования сетью Интернет в кафе, ресторанах, библиотеках, аэропортах, вокзалах и т. д.

Кроме того, беспроводные точки доступа Wi-Fi часто используются как вспомогательная технология для автоматизации складских бизнес-процессов, а также при создании локальных сетей в жилых домах и отдельных квартирах.

Адекватной математической моделью для данной задачи является задача об оптимальном покрытии множества произвольной структуры конечной системой подмножеств. Так как потребители расположены неравномерно, целесообразно представлять рыночную область как непрерывную функцию спроса. В настоящей работе получены результаты работы алгоритма решения непрерывной задачи оптимального покрытия множеств с использованием методов оптимального разбиения и Γ -алгоритма Н. З. Шора для нескольких непрерывных функций спроса [1].

Постановка задачи. Пусть Ω – ограниченное, замкнутое, измеримое по Лебегу множество в n -мерном евклидовом пространстве E_n .

С шаром радиуса R с центром в точке τ_i из E_n будем называть множество вида $B(\tau_i, R) = \{x \in E_n : c(x, \tau_i) \leq R\}$, где $c(x, \tau_i)$ – некоторая квазиметрика [2].

Скажем, что совокупность центров $\tau_1, \dots, \tau_N$ задаёт шаровое покрытие множества Ω с радиусом R , если

$$\Omega \subseteq \bigcup_{i=1}^N B(\tau_i, R).$$

Ясно, что радиус R покрытия множества Ω , задаваемого центрами $\tau_1, \dots, \tau_N$ (вектором τ^N), имеет вид

$$R(\tau^N) = \sup_{x \in \Omega} \min_{i=1, \dots, N} c(x, \tau_i),$$

где $\tau^N = (\tau_1, \dots, \tau_N) \in E_n \times \dots \times E_n = E_n^N$ (либо, в частном случае,

$$\tau^N = (\tau_1, \dots, \tau_N) \in \underbrace{\Omega \times \dots \times \Omega}_N = \Omega^N).$$

Очевидно, что покрытие множества Ω , задаваемого вектором $\tau^N = (\tau_1, \dots, \tau_N)$ с радиусом $R(\tau^N)$, является минимальным c -шаровым покрытием, генерируемым вектором τ^N , т.е. никакие c -шары меньшего радиуса с центрами в $\tau_1, \dots, \tau_N$ не покрывают Ω .

Покрытие минимального радиуса назовём оптимальным.

Таким образом, для отыскания оптимального покрытия надо найти величину

$$R(\tau^N) = \inf_{\tau^N} \sup_{x \in \Omega} \min_{i=1, \dots, N} c(x, \tau_i),$$

называемую радиусом оптимального покрытия, и вектор $\tau_*^N = (\tau_1^*, \dots, \tau_N^*)$, при которых достигается значение, равное нижней грани.

Сформулируем непрерывную задачу об оптимальном c -шаровом покрытии множества из E_n в следующем виде.

Пусть Ω – компактное (ограниченное, замкнутое), измеримое по Лебегу множество в n -мерном евклидовом пространстве E_n .

При заданном числе N центров $\tau_1, \dots, \tau_N$ найти их размещение в области Ω , осуществляющее c -шаровое покрытие множества Ω с минимальным радиусом, т.е. найти величину

$$R(\tau_*^N) = \inf_{(\tau_1, \dots, \tau_N) \in \Omega^N} \sup_{x \in \Omega} \min_{i=1, \dots, N} c(x, \tau_i),$$

и вектор $\tau_*^N = (\tau_1^*, \dots, \tau_N^*) \in \Omega^N \subset E_n^N$, при которых в $R(\tau_*^N)$ достигается значение, равное нижней грани.

Здесь функция $c(x, \tau_i)$ определяется как евклидова метрика:

$$c(x, \tau_i) = \|x - \tau_i\|_2 = \sqrt{\sum_{j=1}^n (x^j - \tau_i^j)^2}$$

Поскольку, как отмечалось в [1], для компактного множества Ω из E_n и непрерывной функции c оптимальное покрытие множества Ω заданным числом N c -шаров одинакового радиуса существует, то данную задачу можно переписать следующим образом:

$$R(\tau_1, \dots, \tau_N) = \max_{x \in \Omega} \min_{i=1, \dots, N} c(x, \tau_i) \rightarrow \min_{(\tau_1, \dots, \tau_N) \in \Omega^N}$$

Для случая евклидовой нормы $c(x, \tau_i)$, т.е. $c(x, \tau_i) = \|x - \tau_i\|_{E_n}$, функция $R(\tau_1, \dots, \tau_N)$ – непрерывна и ограничена на Ω^N . Отметим также, что функция $R(\tau_1, \dots, \tau_N)$ является липшицевой на Ω^N , т.е. выполняется неравенство

$$\left| \max_{x \in \Omega} \min_{i=1, \dots, N} c(x, \tau_i) - \max_{x \in \Omega} \min_{i=1, \dots, N} c(x, \bar{\tau}_i) \right| \leq L \cdot \|\tau^N - \bar{\tau}^N\|_{E_n^N},$$

где $\tau^N = (\tau_1, \dots, \tau_N)$, $L > 0$ – константа Липшица, не зависящая от выбора точек τ^N и $\bar{\tau}^N$ из Ω^N .

Отыскание решения данной задачи можно предложить следующим образом:

1) найти разбиение Дирихле Вороного $\Omega_1, \dots, \Omega_N$, генерируемое вектором $\tau^N = (\tau_1, \dots, \tau_N)$ из Ω^N , имеющее для $i = 1, \dots, N$ и почти всех $x \in \Omega$ вид:

$$\lambda_i(x) = \begin{cases} 1, & c(x, \tau_i) = \min_{j=1, \dots, N} c(x, \tau_j), x \in \Omega_i, \\ 0, & \text{в остальных случаях} \end{cases}.$$

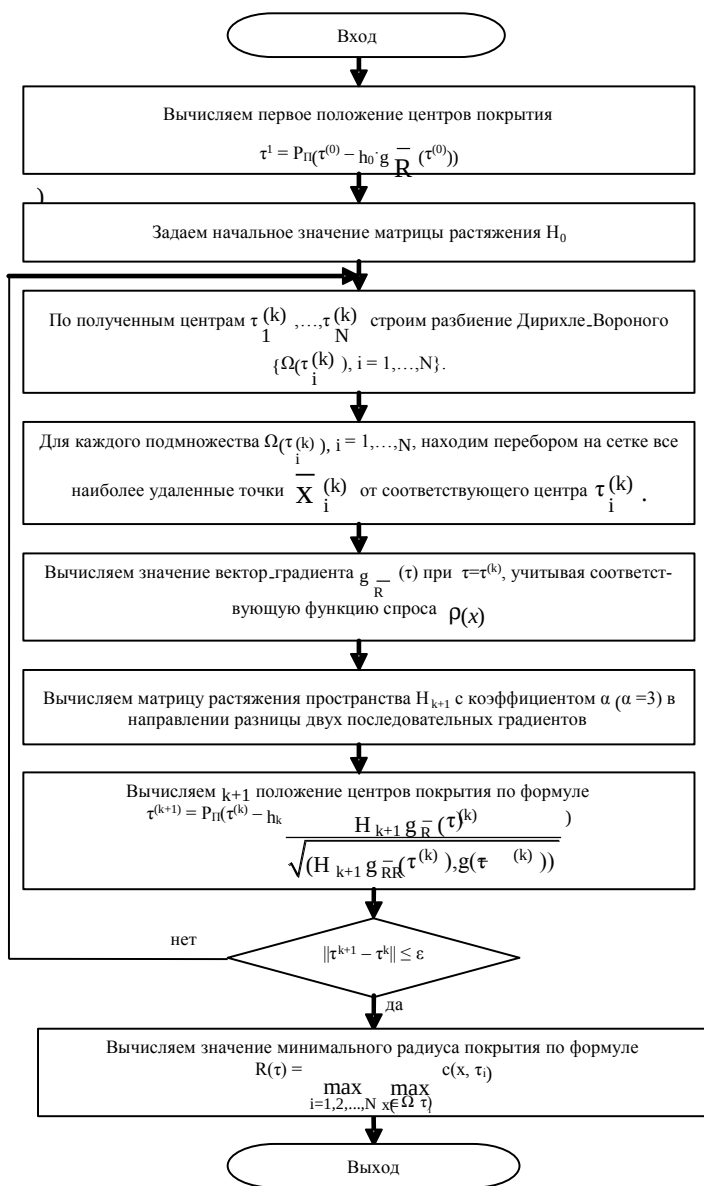
2) для каждого подмножества Ω_i найти максимальное расстояние от точки τ_i до его границы;

3) среди этих максимальных расстояний выбрать наибольшее и принять его в качестве радиуса $R(\tau^N)$ покрытия;

4) из всех наборов $(\tau_1, \dots, \tau_N) \in \Omega^N$ выбрать тот, (обозначим его $(\tau_1^*, \dots, \tau_N^*)$), для которого радиус $R(\tau^N)$ будет минимальным.

На рис. 1 представлен алгоритм решения предложенной задачи в виде блок-схемы.

Результаты работы. Для вычислительных экспериментов использовали нелинейную выпуклую функцию спроса (NLD-1)[3] и функцию спроса Ньюлинга (Newling) (NLD-7) – известной как функции распределения населения в большинстве городов [4]. Графики и уравнения рассматриваемых функций представлены на рис. 2.



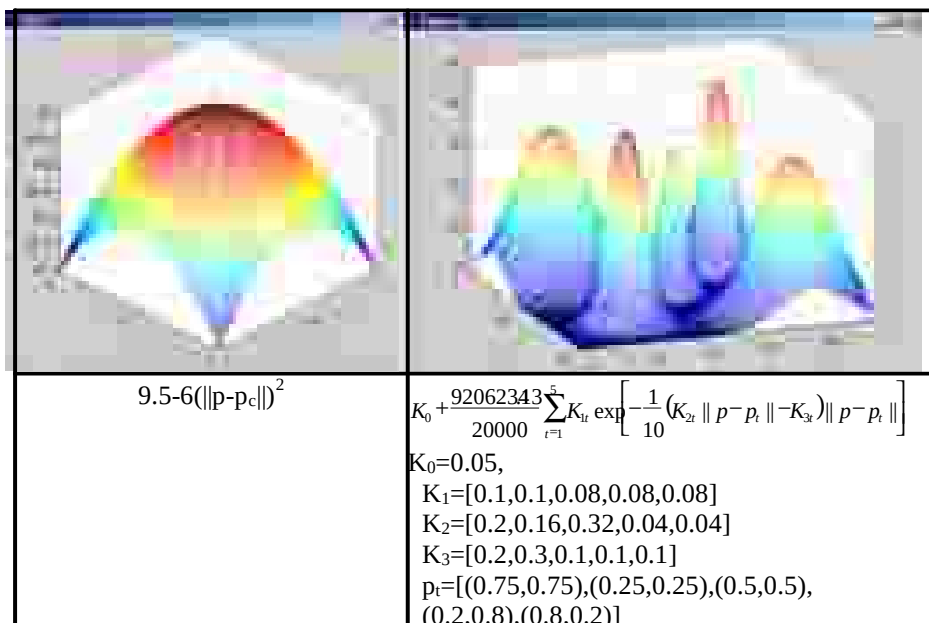


Рис. 2. Графики функций спроса NLD-1 и NLD-7

Поскольку функции $R_i(\tau_i)$ недифференцируемы, то для решения задачи использовался один из вариантов метода обобщенного градиентного спуска с растяжением пространства в направлении разности двух последовательно обобщенных градиентов (Γ -алгоритм Н. З. Шора) [5]. Исследуемое множество представляет собой единичный квадрат. Эксперименты проводились для трех различных количеств точек доступа n ($n=10, n=15, n=20$). Для достоверности результатов было проведено 10 запусков для каждой функции спроса. Далее представлены графики разбиений (рис. 3) и таблица, содержащая минимальный радиус $R(\tau_*^N)$ и набор точек-центров подмножеств $(\tau_1^*, \dots, \tau_N^*)$.

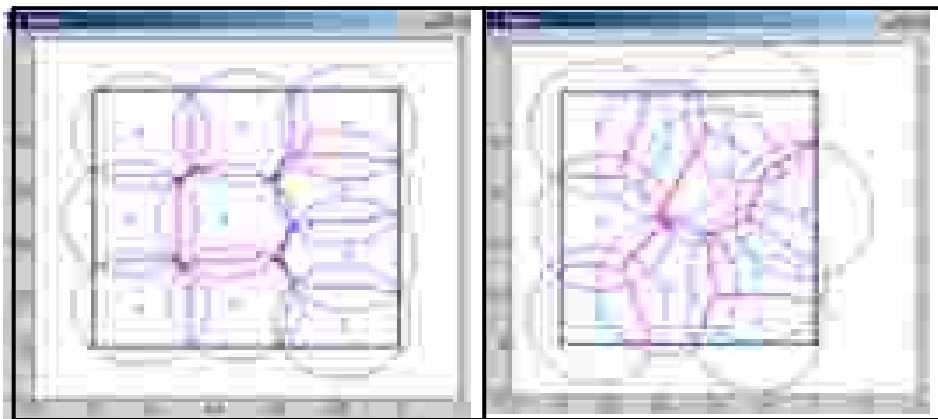


Рис. 3. Решение задачи ОПМ для функций спроса NLD-1 и NLD-7, $n=10$, $m=50$, $\text{eps}=0.0001$

Таблица 1

Результаты вычислительного эксперимента для функций спроса
NLD-1 и NLD-7, $n=10$, $m=50$, $\text{eps}=0.0001$

	$R(\tau_*^N)$	$(\tau_1^*, \dots, \tau_5^*)$	$(\tau_6^*, \dots, \tau_{10}^*)$
NLD-1	0.2206	0.8302 0.8703 0.8105 0.6200 0.8100 0.1000 0.4800 0.8529 0.4602 0.1699	0.1500 0.8382 0.1500 0.1512 0.8337 0.3600 0.1169 0.5000 0.4296 0.4934
NLD-7	0.2714	0.1633 0.4852 0.4204 0.2297 0.1402 0.1165 0.1501 0.8611 0.7553 0.9202	0.4066 0.7719 0.9546 0.5324 0.7774 0.2810 0.5907 0.6643 0.7700 0.0888

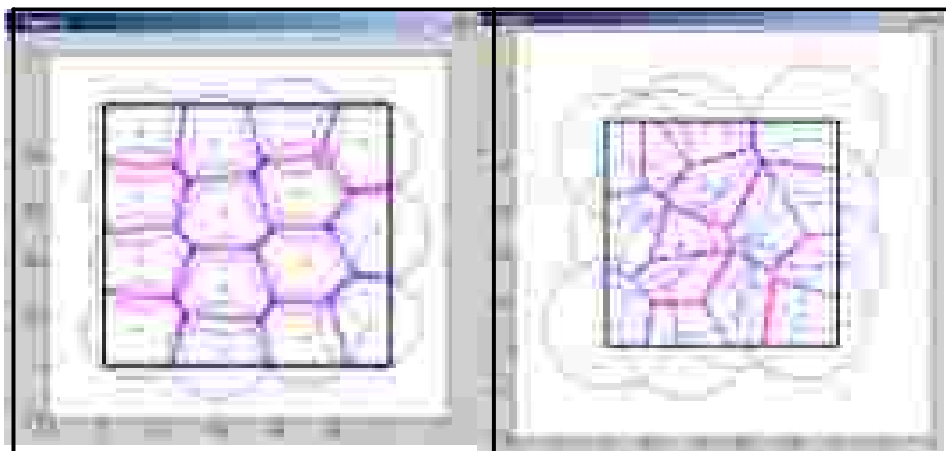


Рис. 4. Решение задачи ОПМ для функций спроса NLD-1 и NLD-7, $n=15$, $m=100$, $\text{eps}=0.0001$

Таблица 2

Результаты вычислительного эксперимента для функций спроса
NLD-1 и NLD-7, $n=15$, $m=100$, $\text{eps}=0.0001$

	$R(\tau_*^N)$	$(\tau_1^*, \dots, \tau_5^*)$	$(\tau_6^*, \dots, \tau_{10}^*)$	$(\tau_{11}^*, \dots, \tau_{15}^*)$
NLD-1	0.1825	0.9536 0.5148	0.1349 0.8957	0.4199 0.3034
		0.6847 0.6680	0.9092 0.8485	0.4038 0.0596
		0.1205 0.6508	0.9289 0.1680	0.6569 0.9251
		0.3923 0.8626	0.4198 0.5917	0.1339 0.3995
		0.7093 0.1080	0.6995 0.3915	0.1134 0.1343
NLD-7	0.2856	0.3457 0.0600	0.4222 0.6962	0.3240 0.4608
		0.5638 0.1867	0.2474 0.8497	0.9198 0.3555
		0.0726 0.8464	0.7368 0.5438	0.0169 0.1026
		0.3385 0.3282	0.3525 0.9198	0.1284 0.5279
		0.9169 0.9644	0.8470 0.1421	0.8887 0.6328

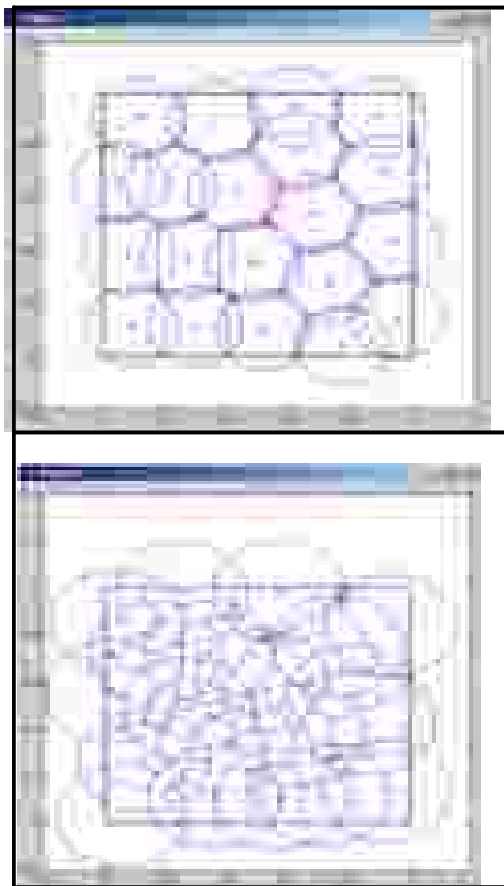


Рис. 5. Решение задачи ОПМ для функций спроса NLD-1 и NLD-7, $n=20$, $m=50$, $\epsilon=0.0001$

Таблица 3

Результаты вычислительного эксперимента для функций спроса

NLD-1 и NLD-7, $n=20$, $m=50$, $\epsilon=0.0001$

	$R(\tau_*^N)$	$(\tau_1^*, \dots, \tau_5^*)$	$(\tau_6^*, \dots, \tau_{10}^*)$	$(\tau_{11}^*, \dots, \tau_{15}^*)$	$(\tau_{16}^*, \dots, \tau_{20}^*)$
NLD-1	0.1501	0.0773 0.3846 0.8851 0.9114 0.2177 0.6513 0.6284 0.7824 0.4230 0.6285	0.9002 0.6947 0.6789 0.5429 0.9501 0.1416 0.0901 0.1156 0.3590 0.8975	0.2702 0.3716 0.6166 0.9515 0.9357 0.4361 0.7688 0.0403 0.1148 0.9106	0.7275 0.2729 0.2949 0.1160 0.0492 0.6658 0.5128 0.0981 0.4826 0.3558
NLD-7	0.2161	0.3586 0.1109 0.0191 0.5579 0.0513 0.0309 0.3974 0.3461 0.9231 0.1410	0.5124 0.6071 0.1233 0.2656 0.0368 0.8594 0.3421 0.7294 0.8699 0.4233	0.1622 0.7251 0.7209 0.1195 0.6268 0.3783 0.2505 0.9794 0.5727 0.1066	0.5703 0.9871 0.8112 0.5397 0.2682 0.4848 0.7641 0.7513 0.9404 0.8402

Согласно полученным данным следует, что минимальный радиус покрытия зависит не только от вида функции спроса, но и от оптимального количества точек доступа. Для 15 точек минимальный радиус для функции Ньюлинга в 1.5649 раза больше, в то время как для 20 точек минимальные радиусы покрытия для выпуклой функции спроса и для функции спроса Ньюлинга отличаются в 1.4397 раза. Это отличие согласуется с тем фактом, что чем больше потребителей сосредотачивается в определенной области, тем большая потребность в наличии нескольких источников сигнала. Такое явление наблюдается на практике, когда точка доступа Wi-Fi служит для автоматизации управления складом и применяется для создания так называемых «горячих точек» – областей, в радиусе которых, клиенты ресторанов, отелей, торговых центров и тому подобных бизнес структур получают бесплатный выход в сеть интернет одновременно. В состав беспроводных локальных сетей обычно входят сразу несколько точек доступа Wi-Fi, каждая из которых может функционировать при подключении к ней более двух сотен рабочих узлов.

Выводы. Реализован алгоритм решения непрерывной задачи оптимального покрытия множества из E_n , разработанного на основе теории оптимального разбиения множества с использованием метода обобщенного градиентного спуска с растяжением пространства в направлении двух последовательных субградиентов (Γ -алгоритм Н. З.Шора) для выпуклой функции спроса (NLD-1) и функции спроса Ньюлинга. Результаты получены для различного количества точек доступа. Сделаны выводы относи-

тельно зависимости минимального радиуса покрытия от вида функции спроса и количества точек доступа. Определены координаты оптимальных мест расположения базовых станций Wi-Fi, в основную задачу которых входит обеспечение беспроводного доступа к уже существующей сети или организация новой беспроводной сети.

Библиографические ссылки

1. **Киселева Е. М.** Непрерывные задачи оптимального разбиения множества: теория, алгоритмы, приложения: Монография. / Е. М. Киселева, Н. З. Шор – К., 2005, С. 430–443
2. **Сухарев А. Г.** Минимаксные алгоритмы в задачах целочисленного анализа. / А.Г. Сухарев – М., 1989.- 300 с.
3. **Murat A., Verter V., Laporte G.** A Continuous Analysis Framework for the Soution of Location-Allocation Problems with Dense Demand, 2008, P. 18–22
4. **Newling BE.** The spatial variation of urban population densities. Geographical Review, 1969, P. 242–252
5. **Шор Н. З.** Методы минимизации недифференцируемых функций и их приложение. / Н.З. Шор – К., 1979.- 200 с.

Надійшла до редколегії 15.06.11

УДК 519.6

¹³О.М. Кісельова, Л.С. Коряшкіна, О.В. Правдивий
Дніпропетровський національний університет імені Олеся Гончара

ІДЕНТИФІКАЦІЯ ДИНАМІЧНОЇ СИСТЕМИ З ДЕКІЛЬКОМА РЕЖИМАМИ ФУНКЦІОНУВАННЯ

Представлена задача ідентифікації «багатозонної» динамічної системи, яка в процесі реальної експлуатації може змінювати стан функціонування. Задача полягає в отриманні або уточненні за експериментальними даними невідомих параметрів моделі об'єкта, а також границь між зонами функціонування. Запропоновано підхід до розв'язання задачі, заснований на чисельному диференціюванні часового ряду, що спостерігається, і відновленні старшої похідної в класі кусково-неперервних функцій.

Представлена задача идентификации «многозонной» динамической системы, которая в процессе реальной эксплуатации может изменять свое состояние функционирования. Задача состоит в отыскании или уточнении по экспериментальным данным неизвестных параметров модели объекта, а также границы переключения режимов. Предложен подход для решения задачи, основанный на численном дифференцировании наблюдаемого временного ряда и восстановлении старшей производной в классе кусочно-непрерывных функций.

The identification problem, «a multi-zone» of a dynamic system, which is in actual operation may change its state of functioning, is considered. The problem is finding or clarification from the experimental data of unknown parameters of the model object as well as the border zones of switching operation. An approach to solving the problem based on numerical differentiation of the observed time series, and the restoration of the highest derivative in the class of piecewise continuous functions, is proposed.

Ключові слова: ідентифікація, динамічна система, часовий ряд, звичайне диференціальне рівняння з розривною правою частиною.

Вступ. У різних галузях науки і техніки для того, щоб пізнати закономірності функціонування деякого об'єкта чи природного явища проводяться експерименти самого різного характеру [1;2]. Мета експерименту –

виявлення головних закономірностей явища або процесу і формування на їх основі деякої математичної моделі. Часто на практиці зустрічаються ситуації, коли об'єкт дослідження або принципово недоступний, або ж проведення експериментів значно коштовне. Тоді основним методом досліджень виступає математичне моделювання. З точки зору співвідношення «причини – наслідки» всі задачі математичного моделювання можна розбити на два великих класи – прямі (відомі причини, знайти наслідки) і обернені (відомі наслідки, виявити причини).

Робота присвячена розробці і обґрунтуванню методів і алгоритмів розв'язання деяких обернених задач для звичайних диференціальних рівнянь, що у своїй математичній постановці можуть бути зведені до динамічних задач оптимального розбиття множин. Такі задачі виникають при реконструкції нелінійних систем. Наприклад [3;4], багато енергоємних об'єктів, такі як електричні печі, функціонують в широкому інтервалі температур і не можуть бути представлені лінійними диференціальними рівняннями з постійними коефіцієнтами. Часто це обумовлено зміною теплофізичних властивостей матеріалів від температури, наявністю запізнення, зміненням об'єму завантаження та іншими факторами. Для такого роду об'єктів досить точно динамічні режими можна описати з послідовним використанням декількох лінійних диференціальних рівнянь. Для цього весь температурний режим (діапазон) розбивається на декілька стадій або зон таким чином, щоб у межах однієї стадії динаміка із достатньою точністю описувалася б одним (векторним) лінійним диференціальним рівнянням. Задача ідентифікації таких систем полягає у пошуку не тільки коефіцієнтів функцій правих частин диференціальних рівнянь, а й границь між зонами функціонування. Далі розроблений алгоритм реконструкції систем звичайних диференціальних рівнянь з правими частинами, що перемикаються [5], на основі поєднання елементів теорії обернених задач і методів розв'язання неперервних задач оптимального розбиття множин (ОРМ) [6].

Математична модель задачі ідентифікації «багатозонної» динамічної моделі. Нехай стан деякого об'єкта характеризується двома фазовими координатами (x_1, x_2) , зміна яких з часом описується системою диференціальних рівнянь з розривною правою частиною

$$\begin{aligned} \dot{x}_1 &= x_2; \\ \dot{x}_2 &= a_1^i x_1 + a_2^i x_2 + f_i(b^i, t), \quad x_1 \in \Omega_i, \quad i = \overline{1, N}. \end{aligned} \quad (1)$$

Тут $(\Omega_1, \dots, \Omega_N)$ – множини значень першої фазової координати системи, які визначають можливі стани (зони) функціонування системи, $a_j^i, b^i, i = 1, N$ – задані параметри, $f_i(b^i, t)$ – відомі неперервні за своїми аргументами функції.

Пряма задача для рівняння (1) формулюється так: за відомими границями між зонами $(\Omega_1, \Omega_2, \dots, \Omega_N)$, знайти розв'язок системи (1), що задовольняє умови:

$$x_1(0) = x_{10}, \quad x_2(0) = x_{20}. \quad (2)$$

Обернена задача для системи (1) формулюється за припущенням, що границі переключення зон невідомі, а значення змінної x_1 спостерігаються на деякому відрізку часу $\bar{T} = \{t : 0 \leq t \leq T\}$: потрібно за відомими значеннями першої фазової координати $x_1 = \tilde{x}_1$ на $\bar{T}$ і за умов (2) визначити в системі (1) границі між зонами $(\Omega_1, \Omega_2, \dots, \Omega_N)$ а також параметри $a_j^i, b^i, i = 1, N$, якщо вони невідомі. Таким чином, в оберненій задачі у системі (1) визначенню підлягає права частина рівняння, про яку відомо, що вона належить до класу кусково неперервних на $R^2 \times \bar{T}$ функцій.

Зведення задачі ідентифікації до динамічної задачі оптимального розбиття множини. Задача ідентифікації «багатозонної» моделі формулюється в термінах теорії неперервних динамічних задач оптимального розбиття множин у такий спосіб: за результатами вимірювань вхідних і вихідних змінних об'єкта, а також заданій кількості N зон потрібно визначити розбиття $\bar{\omega} = (\Omega_1, \dots, \Omega_N) \in P_N(\Omega)$ множини Ω всіх можливих значень першої фазової координати і невідомі параметри правої частини системи диференціальних рівнянь так, щоб різниця між експериментальними і розрахунковими даними досягала мінімального значення, тобто досягав мінімального значення деякий критерій якості.

Якщо через $P_N(\Omega)$ позначити клас усіх можливих розбиттів множини Ω на N підмножин:

$$P_N(\Omega) = \{\bar{\omega} = (\Omega_1, \dots, \Omega_N) : \bigcup_{i=1}^N \Omega_i = \Omega, \text{mes}(\Omega_i \cap \Omega_j) = 0, i \neq j, i, j = \overline{1, N}\}$$

то обернена задача для системи (1) має вигляд

$$I(\Omega_1, \dots, \Omega_N) \rightarrow \min_{(\Omega_1, \dots, \Omega_N) \in P_N(\Omega)}. \quad (3)$$

В якості функціонала задачі (3) можна розглядати такі:

$$1) \quad I^1(\Omega_1, \dots, \Omega_N) = \int_0^T (\dot{x}_2(t) - \dot{\tilde{x}}_2(t))^2 dt, \quad (4)$$

де $\dot{\tilde{x}}_2(t)$ – наближені значення похідної фазової координати $x_2(\cdot)$, обчислені за результатами спостережень змінної $x_1(\cdot)$ за допомогою операції чисельного диференціювання, а $x_2(\cdot)$ задовольняє систему (1);

$$2) \quad I^2(\Omega_1, \dots, \Omega_N) = \int_0^T \beta_0 (x_1(t) - \tilde{x}_1(t))^2 + \beta_1 (x_2(t) - \tilde{x}_2(t))^2 dt, \quad (5)$$

де $x_1(\cdot)$, $x_2(\cdot)$ – розв'язок задачі Коші для відновлюваної системи, $\tilde{x}_1(t)$ – значення фазової змінної, що спостерігається, $\tilde{x}_2(t)$ – наближені значення фазової координати $x_2(\cdot)$, обчислені, при необхідності (тобто якщо $\beta_1 \neq 0$) за результатами спостережень змінної $x_1(\cdot)$ за допомогою операції чисельного диференціювання.

Наведемо алгоритм розв'язання задачі (1) – (3) з критерієм якості (4).

Метод розв'язання задачі. Розв'язання задачі визначення правої частини рівняння (1) за відомими результатами спостережень зводиться до застосування операції чисельного диференціювання для знаходження за спостережуваними значеннями фазової змінної $\tilde{x}_1(t)$ наближеного значення її другої похідної $\dot{\tilde{x}}_2(t)$ і апроксимації останньої кусково-неперервною функцією із класу

$$\Phi \equiv \{\varphi(t, x; \bar{\omega}) : \varphi(t, x; \bar{\omega}) = a_1^i x_1 + a_2^i x_2 + \bar{f}_i(b^i, t) \text{ м.в. для } x \in \Omega_i, i = \overline{1, N}, \bar{\omega} \in P_N(\Omega)\}$$

Основна проблема, що виникає при реалізації цього методу в тому, що задача диференціювання функції, заданої наближено, в рівномірній метриці є некоректною [1;7].

Оптимальний розв'язок задачі мінімізації цільового функціонала

$$I^1(\bar{\omega}) = \int_0^T (\phi(x, t; \bar{\omega}) - \dot{\tilde{x}}_2(t)) dt \rightarrow \min_{\bar{\omega} \in P_N(\Omega)}$$

отримується шляхом зведення цієї задачі до задачі ОРМ. Останню далі представимо в термінах характеристичних функцій підмножин:

$$\bar{I}(\lambda(\cdot), \tau) = \int_0^T \left(\sum_{i=1}^N f_i(t, \tilde{x}(t); \tau_i) \lambda_i(\tilde{x}_1(t)) - \dot{\tilde{x}}_2(t) \right)^2 dt \rightarrow \min_{(\lambda(\cdot), \tau) \in \Lambda_0 \times R^{3N}}, \quad (6)$$

де $\lambda(\cdot) \in \Lambda$;

$$\Lambda = \left\{ \lambda(\cdot) = \left(\lambda_1(\cdot), \dots, \lambda_N(\cdot) : \lambda_i(x) = 0 \vee 1; i = \overline{1, N}; \sum_{i=1}^N \lambda_i(x) = 1 \text{ м.в. } x \in \Omega \right) \right\};$$

$\tau_i = (a_1^i, a_2^i, b^i)$, $i = \overline{1, N}$ – заздалегідь вважаються невідомими;

$$f_i(t, x; \tau_i) = a_1^i x_1 + a_2^i x_2 + \bar{f}_i(b^i, t), i = \overline{1, N}.$$

Оптимальний розв'язок задачі (6) визначається наступним чином для кожного $\tilde{x}_1(t) \in \Omega_i^*, i = \overline{1, N}$:

$$\lambda_i^*(\tilde{x}_1(t)) = \begin{cases} 1, & \text{якщо } |f_i(t, \tilde{x}; \tau_i^*) - \dot{\tilde{x}}_2(t)| = \min_{k=1, N} |f_k(t, \tilde{x}; \tau_k^*) - \dot{\tilde{x}}_2(t)|; \\ 0, & \text{в іншому випадку.} \end{cases} \quad (7)$$

де в якості параметрів $\tau_i^* = (a_1^i, a_2^i, b^i)$, $i = \overline{1, N}$ вибирається оптимальний розв'язок задачі

$$G(\tau_1, \dots, \tau_N) = \int_0^T \min_{i=1, N} (f_i(t, \tilde{x}(t); \tau_i) - \dot{\tilde{x}}_2(t)) dt \rightarrow \min_{\tau \in R^{3N}}. \quad (8)$$

Задача (8) належить до класу задач скінченно-вимірної недиференційованої оптимізації, до розв'язання якої застосовується метод узагальненого градієнтного спуску з розтягненням простору в напрямку різниці двох послідовних субградієнтів (Г-алгоритм Н.З. Шора).

Слід зазначити, що одночасно із знаходженням границь перемикання зон динамічної системи, можуть бути отримані моменти часу переходу

стану системи з i -ої в j -ту зону. Ці моменти можуть бути визначені в процесі розв'язання нелінійного рівняння $|f_i(t, \tilde{x}; \tau_i^*) - \dot{\tilde{x}}_2(t)| = |f_j(t, \tilde{x}; \tau_j^*) - \dot{\tilde{x}}_2(t)|$.

Описаний вище підхід виявився конструктивним у тому розумінні, що чисельний алгоритм розв'язання вихідної задачі ідентифікації «багатостадійної» динамічної системи за даними спостережень відтворює наведену схему.

Зауваження. При розв'язанні задачі пошуку лише границь перемикання зон динамічної системи застосовується модифікація підходу. В цьому випадку немає необхідності вирішувати задачу скінченновимірної оптимізації.

Результати ідентифікації «багатозонної» моделі. Визначення границь між зонами. Дослідження роботи запропонованого алгоритму спочатку проводилися на прикладі задачі ідентифікації кусково неперервної функції правої частини одного звичайного диференціального рівняння, тобто коли динамічна система описувалась у такий спосіб:

$$\dot{x} = f_k(t, x), x \in \Omega_k, (\Omega_1, \dots, \Omega_N) \in P_N(\Omega) \quad (9)$$

$$x(t_0) = x_0. \quad (10)$$

Результат розбиття області значень фазової координати на 10 зон, у випадку коли всі зони ідентифікуються, графічно представлений на рис. 1. На рис.2 наведене розбиття фазової області на зони, отримане за умов, коли динамічна система (1) має ковзний режим за даними, що спостерігаються.

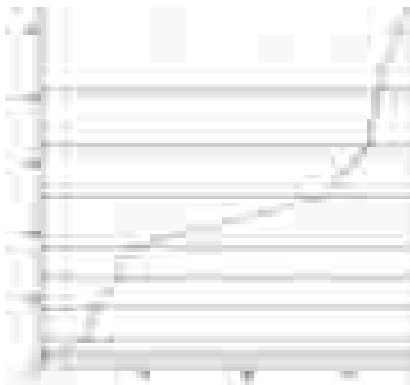


Рис. 1 Розбиття множини значень x при $f_k(t, x) = a_k, k = 1..10$

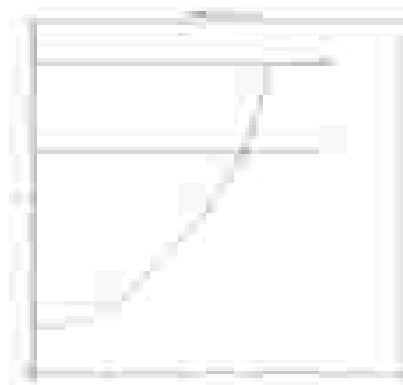


Рис. 1 Визначення зон функціонування динамічної системи (9)-(10).
Наявність ковзного режиму

На рис. 3,4 представлено графічне зображення границі між зонами у просторі фазових змінних (x_1, x_2) системи

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \begin{cases} -1, & x_2 > kx_1 \\ 1, & x_2 < kx_1 \end{cases} \end{cases} \quad (11)$$

відновленої за низкою послідовностей спостережень.

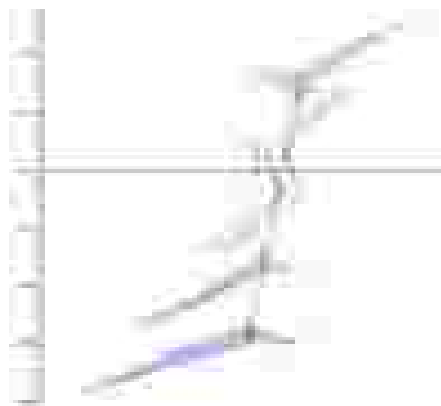


Рис. 3 Лінія перемикання зон в системі (11), $k=0.8$

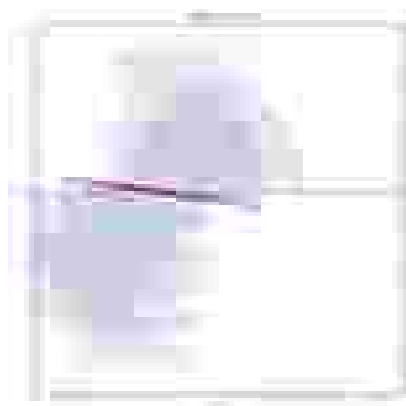


Рис. 42 Лінія перемикання зон в системі (11), $k=-0.1$

На рис. 5 наведений результат відновлення розривної за фазовою змінною правої частини диференціального рівняння 2-го порядку, коли в якості функцій правих частин обирались функції вигляду

$f_i = a_{0i}x_1 + a_{1i}x_2 + b_i \times t, \quad i = 1, N$. Тут можна прослідити за видом функцій, що складають праву частину системи диференціальних рівнянь, початкові і відновлені параметри, фазові портрети побудовані за результатами чисельного інтегрування відновлених систем, і відповідних їм спостережень, крім того на рисунках вказано час роботи алгоритмів і значення середньоквадратичного відхилення реконструйованої фазової траєкторії і заданої.

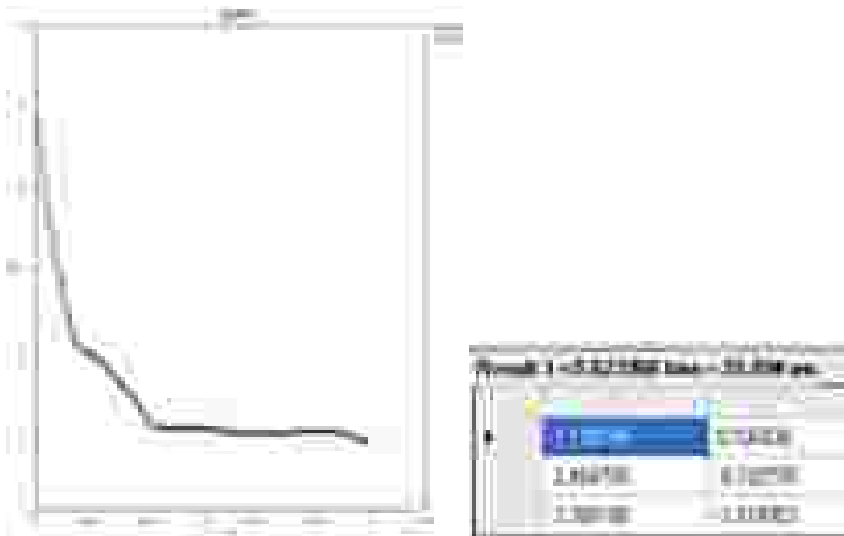


Рис. 5 Ідентифікація системи при наявності ковзного режиму

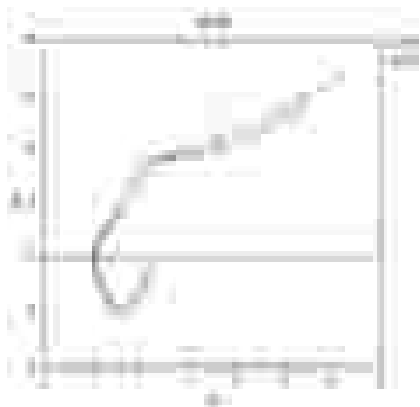
Цікавими виявилися приклади динамічних систем, які характеризуються наявністю ковзного режиму. На рис. 2 наведено приклад системи, в якій як тільки фазова точка потрапляє на границю перемикання зон, на цій границі і залишається. Іноді фазова точка може вийти з ковзного режиму (рис. 5), це залежить від вигляду функцій правої частини диференціального рівняння. Аналізуючи результати обчислювальних експериментів, можна простежити неєдиність розв'язків задач ідентифікації динамічної системи, тобто значення параметрів можуть суттєво відрізнятися але при цьому апроксимація правої частини здійснюється із задовільною точністю. Цей факт не суперечить властивостям обернених задач.

Було виявлено, що збільшення кількості спостережень у декілька разів, навіть на порядок, не суттєво впливає на якість реконструкції (рис.6).

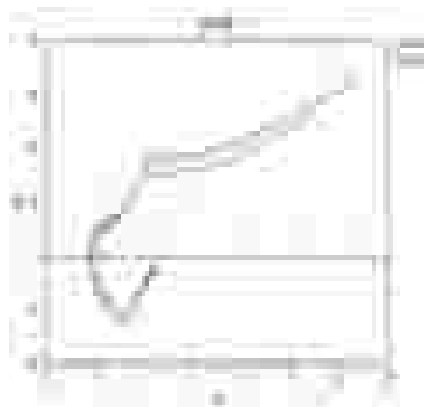
Зазначимо ще декілька особливостей процесу розв'язання задач відновлення кусково_неперервної за фазовими змінними функції правої частини системи (1), виявлених з обчислювальних експериментів:

- якщо в наборі спостережень присутні точки перемикання, то вибірка спостережень повинна містити точки з різних зон функціонування, кількість їх залежить від виду функції правих частин;

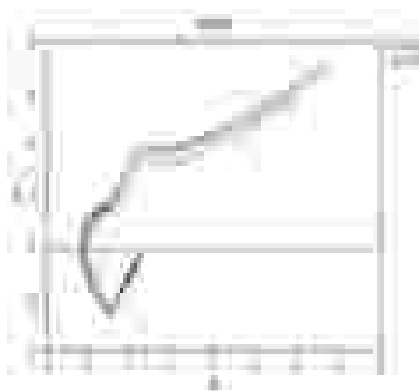
- якість реконструкції динамічної системи за допомогою наведеного алгоритму залежить від початкових наближень невідомих параметрів.



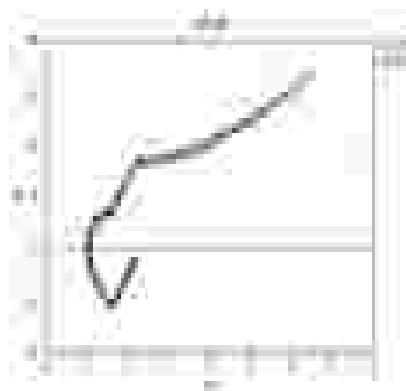
а) $N = 25$



б) $N = 50$



в) $N = 100$



г) $N = 300$

Рис.6. Результат ідентифікації «багатозонної динамічної системи за результатами N спостережень

Висновки. Робота є послідовним розвитком теорії неперервних задач оптимального розбиття множин у напрямі розробки і обґрунтування методів розв’язування динамічних задач ОРМ. Сформульована обернена задача для системи диференціальних рівнянь з розривною правою частиною. Продемонстрована можливість формулювання задачі ідентифікації «багатозонної» моделі динаміки в термінах теорії неперервних задач оптимального розбиття множин і використання методів ОРМ при розробці алгоритмів реконструкції динамічних систем. Запропонований математичний апарат може бути надалі застосований при розв’язуванні задач оптимального керування на множині станів функціонування, коли у процесі керування можуть змінюватися параметри моделі динаміки, значення фазових координат тощо.

Бібліографічні посилання

1. **Тихонов А.Н.** Математические задачи компьютерной томографии / А.Н. Тихонов, В.Я. Арсенин, А.А. Тимонов – М., 1987. – 160 с.
2. **Денисов А.М.** Введение в теорию обратных задач: Учеб. пособие. / А.М. Денисов – М., 1994. – 208 с.
3. **Матвейкин В.Г.** Теоретические основы энергосберегающего управления динамическими режимами установок производственно-технического назначения: монография / В.Г. Матвейкин, Д.Ю. Муромцев. – М., 2007. – 128 с.
4. **Егоров А.И.** Обыкновенные дифференциальные уравнения с приложениями – 2-е изд., испр. / А.И. Егоров – М.: ФИЗМАТЛИТ, 2005. – 384 с.
5. **Филиппов А.Ф.** Дифференциальные уравнения с разрывной правой частью / А.Ф. Филиппов – М., 1985. – 225 с.
6. **Киселева Е.М.** Непрерывные задачи оптимального разбиения множеств: теория, алгоритмы, приложения: Монография / Е.М. Киселева, Н.З. Шор – К., 2005. – 564.
7. **Тихонов А.Н.** Методы решения некорректных задач / А.Н. Тихонов, В.Я. Арсенин – М., 1979.

Надійшла до редколегії 15.06.2011

УДК 519.85

¹⁴О. М. Кісельова*, В. О. Стросьва**

*Дніпропетровський національний університет імені Олеся Гончара

**Дніпродзержинський державний технічний університет

РОЗВ'ЯЗАННЯ НЕЛІНІЙНИХ НЕПЕРЕРВНИХ БАГАТОПРОДУКТОВИХ ЗАДАЧ ОПТИМАЛЬНОГО РОЗБИТТЯ МНОЖИН З ФІКСОВАНИМИ ЦЕНТРАМИ ПІДМНОЖИН

Приведено алгоритм розв'язання нелінійної неперервної багатопродуктової задачі оптимального розбиття множини Ω з n -вимірного евклідового простору на її підмножини з фіксованими центрами цих підмножин при обмеженнях у формі рівностей та нерівностей, з функціями попиту та вартості транспортування одиниці продукції, які можуть бути різними для різних видів продукції.

Приведен алгоритм решения нелинейной непрерывной многопродуктовой задачи оптимального разбиения множества Ω из n -мерного евклидова пространства на его непересекающиеся подмножества с фиксированными центрами этих подмножеств при ограничениях в форме равенств и неравенств, с функциями спроса и стоимости транспортировки единицы продукции, которые могут быть разными для разных видов продукции.

The algorithm of the decision of a non-linear continuous multigrocery problem of a set Ω from n -dimensional Euclidean space on its not crossed subsets with fixed the centres is given at restrictions in the form of equalities and inequalities, with functions of demand and costs of transportation of a unit of production which can be different for different kinds of production.

Ключові слова: нескінченновимірне математичне програмування, оптимальне розбиття множин, багатопродуктова задача.

Вступ. Багато практично важливих задач оптимізації та задач з різних розділів прикладних наук зводиться до задач розбиття заданої множини визначеної структури на її неперетинні підмножини з метою мінімізації деякого критерію якості розбиття.

В [1] викладається математична теорія неперервних задач оптимального розбиття множин (ОПМ) n -вимірного евклідового простору, які є некла-

сичними задачами нескінченновимірною математичного програмування з булевими змінними. На основі розроблених і теоретично обґрунтованих у монографії [1] методів розв'язання задач оптимального розбиття множин сформульовані алгоритми, складовою частиною яких є r -алгоритм Н. З. Шора або його модифікації.

В [2]–[5] задачі з [1] узагальнюються на нелінійний випадок. Дослідження цієї роботи продовжує розгляд нелінійних неперервних багатопродуктових задач ОРМ з [3]–[5]. Метою роботи є приведення алгоритму розв'язання нелінійної неперервної багатопродуктової задачі оптимального розбиття множини Ω на її підмножини з фіксованими центрами при обмеженнях у формі рівностей та нерівностей, його програмна реалізація на випадок, коли функції попиту та вартості транспортування одиниці продукції можуть бути різними для різних видів продукції та ілюстрація цього алгоритму на модельних задачах.

Постановка задачі. Розглядається нелінійна неперервна багатопродуктова задача оптимального розбиття множини $\Omega \in E^n$ на її неперетинні вимірні за Лебегом підмножини $\Omega_1^1, \dots, \Omega_N^1; \Omega_1^2, \dots, \Omega_N^2; \dots; \Omega_1^M, \dots, \Omega_N^M$ (серед яких можуть бути й пусті) так, щоб

$$\text{mes}(\Omega_{ik}^j \setminus \Omega) = 0, \quad i \neq k, \quad i, k = 1, \dots, N, \quad j = 1, \dots, M,$$

де $\text{mes}()$ – міра Лебега,

$$\bigcup_{i=1}^N \Omega_i^j = \Omega, \quad j = 1, 2, \dots, M,$$

з фіксованими координатами центрів підмножин при обмеженнях у вигляді рівностей та нерівностей.

Позначимо сукупність усіх можливих розбиттів множини Ω на N підмножин по M продуктів через Σ_{Ω}^{NM} , тобто

$$\Sigma_{\Omega}^{NM} = \left\{ (\Omega_1^1, \dots, \Omega_N^1; \dots; \Omega_1^M, \dots, \Omega_N^M) : \begin{aligned} & \bigcup_{i=1}^N \Omega_i^j = \Omega, \\ & \text{mes}(\Omega_{ik}^j \setminus \Omega) = 0, \quad i \neq k, \quad i, k = 1, \dots, N, \quad j = 1, \dots, M \end{aligned} \right\}.$$



$$\text{Задача А.} \quad \min_{\{\Omega_1^1, \dots, \Omega_{NM}^M, \Omega\}} \sum_{j=1}^{MN} \sum_{i=1}^M \left[\Phi_{ij}^j \int_{\Omega} \rho^j(x) dx + c^j(x, \tau) \rho^j(x) \right] \quad (1)$$

при умовах

$$\{\Omega_1^1, \dots, \Omega_{NM}^M, \Omega\} \in \Sigma_{\Omega}^{NM} \quad (2)$$

$$\sum_{j=1}^M \int_{\Omega_i^j} \rho^j(x) dx = b_i, \quad i=1, \dots, N, \quad (3)$$

$$\sum_{j=1}^M \int_{\Omega_i^j} \rho^j(x) dx \leq b_i, \quad i=1, \dots, N.$$

де $x = (x^{(1)}, \dots, x^{(n)}) \in \Omega$; $\tau = (\tau_1, \dots, \tau_{NM}) \in \Omega$; $b_1, \dots, b_N$ – задані невід’ємні числа, причому виконуються умови розв’язності задачі

$$S = M \cdot \int_{\Omega} \rho^j(x) dx \leq \sum_{i=1}^N b_i, \quad 0 < b \leq S, \quad i=1, \dots, N.$$

Функції $c^j(x, \tau_i)$ – дійсні, обмежені, визначені на $\Omega \times \Omega$, вимірні по аргументу x при будь-якому фіксованому $\tau_i = (\tau_i^{(1)}, \dots, \tau_i^{(n)}) \in \Omega$ для всіх $i=1, \dots, N$; функції $\rho^j(x)$ – дійсні, обмежені, вимірні і невід’ємні на Ω для всіх $j=1, \dots, NM$; $\tau_i = (\tau_i^{(1)}, \dots, \tau_i^{(n)})$ – задана точка підмножини Ω , одна й та ж для всіх $j=1, \dots, NM$, яку ми називаємо спільним центром підмножин $\Omega_1^1, \dots, \Omega_{NM}^M$, не обов’язково належить кожній $\Omega_1^1, \dots, \Omega_{NM}^M$; Φ_{ij}^j , $i=1, \dots, N$, $j=1, \dots, NM$ – дійсні, обмежені, опуклі, двічі неперервно диференційовні функції свого аргументу.

Вважаємо, що міра Лебега множин і граничних точок підмножин Ω_i^j , $i=1, \dots, N$, $j=1, \dots, NM$, дорівнює нулю.

Задача А записується у термінах характеристичних функцій, а потім, від отриманої нескінченновимірної задачі математичного програмування з булевими значеннями змінних $\lambda_i^j(x)$, $i=1, \dots, N$, $j=1, \dots, NM$, здійснено перехід до відповідної задачі зі значеннями $\lambda_i^j(x)$ на відрізку $[0, 1]$:

Задача С.
$$\min_{(\lambda_0) \in \Gamma_1} \sum_{j=1}^{MN} [\Phi_i^j(\int_{\Omega} \rho^j(x) \lambda_i^j(x) dx) + c^j(x, \tau_i) \rho^j(x) \lambda_i^j(x) dx],$$

де $\Gamma_1 = \{ \lambda(x) : \lambda(x) \in \Gamma \text{ майже скрізь (м. с.) для } x \in \Omega;$

$$\int_{\Omega} \sum_{j=1}^M \rho \lambda_i(x) dx = b, i = 1, \dots, p,$$

$$\int_{\Omega} \sum_{j=1}^M \rho \lambda_i(x) dx \leq b, i = p+1, \dots, N \};$$

$$\Gamma = \{ \lambda(x) = (\lambda_1^1(x), \dots, \lambda_N^1(x); \dots; \lambda_1^M(x), \dots, \lambda_N^M(x)) : 0 \leq \lambda_i^j(x) \leq 1, x \in \Omega,$$

$$i = 1, \dots, N, j = 1, \dots, M, \sum_{i=1}^N \lambda_i(x) = 1 \text{ м. с. для } x \in \Omega \}.$$

По аналогії з [2], [3] можна привести обґрунтування методу розв'язання задачі С, який засновано на переході від вихідної нелінійної нескінченновимірної задачі через функціонал Лагранжа виду

$$\begin{aligned} h(\lambda, \psi) = & - \sum_{i=1}^N \psi_{ii} + \sum_{j=1}^{MN} \left[\Phi_{ii}^j \left(\int_{\Omega} \rho^j(x) \lambda_i^j(x) dx \right) + \right. \\ & \left. - \int_{\Omega} (c^j(x, \tau_i) - \psi \rho^j(x) \lambda_i^j(x) dx) \right] = \\ & = - \sum_{i=1}^N \psi_{ii} b_i + \sum_{j=1}^{MN} \Phi_{ii}^j(\lambda_i(x), \psi_i), \end{aligned} \quad (4)$$

де $\psi = (\psi_1, \dots, \psi_N) \in \Lambda$ – N -вимірний вектор з дійсними компонентами, причому $\psi_1, \dots, \psi_p$ – довільного знаку, а $\psi_{p+1}, \dots, \psi_N$ – невід'ємні; $\lambda(x) \in \Gamma$ для $x \in \Omega$, $\tau_i = (\tau_i^{(1)}, \dots, \tau_i^{(n)}) \in \Omega$, до скінченновимірної задачі з негладким цільовим функціоналом.

Як випливає з [2], [3], [6], у випадку виконання умов сильної регулярності:

$$\text{mes} \left\{ x \in \Omega : \int_{\Omega} \rho^j(x, \tau_i) dx + \psi + \phi_{ij}^j \left(\int_{\Omega} \rho^j(x) \lambda_i^*(x) dx \right) - \rho^j(x) = 0 \right\} = 0, i = 1, \dots, N, j = 1, \dots, M,$$

вектор-функція $\lambda_i^*(x)$, яка доставляє мінімум лінійному функціоналу

$\Phi_i^{jj}(\lambda \Psi, \cdot)_i$ з правої частини (4), визначається при кожному фіксованому $\Psi \in \Lambda$ із наступного операторного рівняння:

$$\lambda_i^{*j}(x) = \begin{cases} 1, & c^j(x, \tau_i) + \Psi_i + \Phi_{Y_i^j}^{ij} \left(\int_{\Omega} \rho^j(x) \lambda_i^{*j}(x) dx \right) \leq c^j(x, \tau_k) + \Psi_k + \Phi_{K_k^j}^{ij} \left(\int_{\Omega} \rho^j(x) \lambda_k^{*j}(x) dx \right), \\ i \neq k \text{ м.с. для } x \in \Omega \text{ (іншими словами, } i = k \text{ тільки на множині міри нуль,} \\ \text{тобто у точках границі між підмножинами } \Omega_k^j \text{ і } \Omega \text{)}, \quad i, k = 1, \dots, N, \quad j = 1, \dots, M, \\ 0, & \text{в інших випадках,} \end{cases}$$

$$\text{де } Y_i^j = \int_{\Omega} \rho \lambda_i^j(x) dx.$$

Отже, слідую [2]–[3], можна привести перехід від нескінченновимірної задачі С до пошуку сідлової точки функціонала (4) у вигляді наступної теореми, на основі якої буде сформульовано алгоритм розв'язання поставленої задачі.

Теорема 1. Якщо $\Phi_i^j(\cdot)$, $i = 1, \dots, N$, $j = 1, \dots, M$ – опуклі, двічі неперервно-диференційовні функції свого аргументу та при кожному фіксованому $\Psi \in \Lambda$ має місце умова сильної регулярності

$$\text{mes} \left\{ x \in \Omega : \left(c^j(x, \tau_i) + \Psi + \Phi_{Y_i^j}^{ij} \left(\int_{\Omega} \rho^j(x) \lambda_i^{*j}(x) dx \right) - \rho^j(x) = 0, i = 1, \dots, N, j = 1, \dots, M, \right. \right.$$

тоді сідлова точка $(\lambda \Psi, \cdot)^*$ (де перша компонента є оптимальним розв'язком задачі С) функціонала (4) на множині $\{\Gamma \times \Omega^N\} \times \Lambda$ визначається для $i = 1, \dots, N$, $j = 1, \dots, M$ та майже всіх $x \in \Omega$ наступним чином:

$$\lambda_i^{*j}(x) = \begin{cases} 1 & \text{при } x \in \Omega_{\tau_i}^{jj} \text{ та } q \leq i, \\ 0 & \text{при } x \notin \Omega_{\tau_i}^{jj}, \end{cases}$$

$$\text{де } \Omega_{\tau_i}^{jj} = \left\{ x \in \Omega : c^j(x, \tau_i) + \Psi_k^{**} + \Phi_{Y_i^j}^{ij}(Y_i^j) = \min_{k \in \bar{N}_1} (c^j(x, \tau_k) + \Psi_k^{**} + \Phi_{K_k^j}^{ij}(Y_k^j)), i \neq k \text{ м.с. для } x \in \Omega \right\},$$

в якості $Y_1^{**}, \dots, Y_N^{**}$, $\Psi_1^{**}, \dots, \Psi_N^{**}$ обирається оптимальний розв'язок слідуючої двоїстої задачі:

$$G\Psi = \max_{Y \in U} \left[\sum_{i=1}^N \psi_i + \sum_{j=1}^{MN} \left(\phi(Y_i^j) - \phi'(Y_i^j) \cdot Y_i^j + \int_{\Omega} \min_{k \in M_i} \left(\phi^j(x, \tau_k) + \psi_k + \phi'_{kY_k^j}(Y_k^j) \cdot \rho^j(x) \right) dx \right) \right] \rightarrow \max, \quad (5)$$

$$\Psi \in \Lambda, Y \in U = \left\{ Y = (Y_1^{11}, \dots, Y_N; \dots; Y_1^M, \dots, Y_N^M) \in E_{NM} : 0 \leq \sum_{j=1}^M Y_i^j \leq b_i, \quad 1, \dots, N \right\}.$$

при умовах $\psi_i \geq 0, \quad i = 1, \dots, N.$ (6)

Алгоритм розв'язання задачі. Сформулюємо алгоритм розв'язання задачі (5), (6), що базується на теоремі 1, в якому для розв'язання допоміжної скінченновимірної негладкої оптимізаційної задачі (5), (6), буде застосовано r -алгоритм Н.З. Шора, [7].

Для цього від задачі (5), (6) перейдемо до задачі безумовної оптимізації по Ψ за допомогою введення в цільову функцію задачі (5), (6) негладких штрафних функцій множин:

$$\{\psi_i \geq 0, i = 1, \dots, N\}, \{Y_i^j \geq 0, i = 1, \dots, N; j = 1, \dots, M\}, \left\{ \sum_{j=1}^M Y_i^j \leq b_i, \quad 1, \dots, N \right\}.$$

А саме, знайти

$$\max_{\Psi \in \Lambda} \max_{Y \in U} P(\Psi, Y), \quad (7)$$

де

$$P(\Psi, Y) = G(\Psi, Y) -$$

$$- S_1 \cdot \sum_{i=1}^N \max\{0, -\psi_i\} - S_2 \cdot \sum_{j=1}^M \sum_{i=1}^N \max\{0, -Y_i^j\} - S_3 \cdot \sum_{i=1}^N \max\left\{0, \sum_{j=1}^M Y_i^j - b_i\right\}. \quad (8)$$

Тут S_1, S_2, S_3 – достатньо великі додатні числа.

Визначаємо вектор узагальненого градієнту функції (8) у точці $(Y, \Psi) = (Y_1^{11}, \dots, Y_N; \dots; Y_1^M, \dots, Y_N^M; \psi_1, \dots, \psi_N)$ наступним чином:

$$g_Y(\Psi, Y) = (g_{Y_1^{11}}(Y, \Psi), g_{Y_1^2}(Y, \Psi), \dots, g_{Y_N^M}(Y, \Psi)) =$$

$$= (g_p^{Y_1^j}(Y, \Psi), \dots, g_p^{Y_N^j}(Y, \Psi); \dots; g_p^{Y_1^M}(Y, \Psi), \dots, g_p^{Y_N^M}(Y, \Psi); g_p^{\Psi_1}(Y, \tau\Psi), \dots, g_p^{\Psi_N}(Y, \tau\Psi))$$

при цьому його компоненти мають вигляд:

$$g_p^{Y_i^j}(Y, \Psi) = -\phi''_{iY_i^j}(Y_i^j) \cdot Y_i^j + \phi''_{iY_i^j}(Y_i^j) \cdot \int_{\Omega} \rho^j(x) \lambda_i^j(x) dx + \quad (9)$$

$$+ S_2 \max[0, \text{sign}(-Y_i^j) - S \max[0, \text{sign}(\sum_{j=1}^M Y_i^j - b_i)] ; i=1, \dots, N, j=1, \dots, M,$$

$$g_p^{\Psi_i}(Y, \Psi) = \begin{cases} \int_{\Omega} \sum_{j=1}^M \rho^j(x) \lambda_i^j(x) dx - b_i, & i=1, \dots, p \\ \int_{\Omega} \sum_{j=1}^M \rho^j(x) \lambda_i^j(x) dx - b_i + S_1 \max[0, \text{sign}(-\Psi_i)] & i=p+1, \dots, N, \end{cases} \quad (10)$$

де $\lambda_i^j(x)$ визначаються за формулами:

$$\lambda_i^j(x) = \begin{cases} 1, & \text{якщо } c^j(x, \tau_i) + \Psi_i + \phi''_{iY_i^j}(Y_i^j) = \min_{k \in I_1} [c^j(x, \tau_k) + \Psi_k + \phi''_{kY_k^j}(Y_k^j)], \\ & i \neq k \text{ м. с. для } x \in \Omega, \\ 0, & \text{в інших випадках,} \end{cases} \quad (11)$$

$$i=1, \dots, N, j=1, \dots, M.$$

Опишемо послідовні етапи алгоритму.

Алгоритм. Область Ω заключаємо в n -вимірний паралелепіпед Π , сторони якого паралельні вісям декартової системи координат, покладемо $\rho^j(x) = 0$ при $x \in \Pi \setminus \Omega$, $j=1, \dots, M$. Паралелепіпед Π покриваємо прямокутною сіткою і задаємо початкове наближення $(Y, \Psi) = (Y^{(0)}, \Psi^{(0)})$. Обчислюємо значення $\lambda_i^{(0)}(x)$ у вузлах сітки за формулами (11) при $Y = Y^{(0)}$, $\Psi = \Psi^{(0)}$. Значення $g_p^{Y_i^j}(Y^{(0)}, \Psi^{(0)})$, $g_p^{\Psi_i}(Y^{(0)}, \Psi^{(0)})$ у вузлах сітки визначаємо за формулами (9)-(10) при $Y = Y^{(0)}$, $\Psi = \Psi^{(0)}$, $\lambda(x) = \lambda^{(0)}(x)$. Вибираємо початковий крок $h_0 > 0$ r -алгоритму.

Перший крок алгоритму проводимо за формулами:

$$Y^{(1)} = Y^{(0)} - h_0 \nabla_Y (Y^{(0)}, \Psi^{(0)}),$$

$$\Psi^{(1)} = \Psi^{(0)} - h g_p^\Psi(Y^{(0)}, \Psi^{(0)}),$$

де P_Π – оператор проектування на Π .

Переходимо до другого кроку.

Нехай у результаті обчислень після $k, k = 1, 2, \dots$ кроків алгоритму отримано значення $Y^{(k)}, \Psi^{(k)}, (k-1)$ у вузлах сітки.

Опишемо $(k+1)$ -й крок алгоритму.

1. Обчислюємо значення $\lambda^{(k)}(x)$ у вузлах сітки за формулами (11) при $Y^{(k)}, \Psi^{(k)}$
2. Обчислюємо значення $g_p^Y(x, \Psi)$ за формулами (9), (10) при $\lambda(x) = \lambda^{(k)}(x), Y^{(k)}, \Psi^{(k)}$.
3. Проводимо $(k+1)$ -й крок Γ -алгоритму узагальнених градієнтів із розтягненням простору, близького до Γ -алгоритму у H -формі [7], коротка схема якого має вигляд

$$Y^{(k+1)} = Y^{(k)} - h_k \frac{H_{k+1} g_p^Y(Y^{(k)}, \Psi^{(k)})}{\sqrt{(H_{k+1} g_p^Y(Y^{(k)}, \Psi^{(k)}), g_p^Y(Y^{(k)}, \Psi^{(k)}))}},$$

$$\Psi^{(k+1)} = \Psi^{(k)} - h_k \frac{H_{k+1} g_p^\Psi(Y^{(k)}, \Psi^{(k)})}{\sqrt{(H_{k+1} g_p^\Psi(Y^{(k)}, \Psi^{(k)}), g_p^\Psi(Y^{(k)}, \Psi^{(k)}))}},$$

де H_{k+1} – матриця розтягнення простору з коефіцієнтом α у напрямку різниці двох послідовних узагальнених градієнтів, яка визначається наступним чином:

$$H_{k+1} = H_k + \left(\frac{1}{\alpha^2} - 1 \right) \frac{H_k \Delta_k^T}{(H_k \Delta_k, \Delta_k)},$$

$$\Delta_k = g_p^Y(Y^{(k)}, \Psi^{(k)}) - g_p^Y(Y^{(k-1)}, \Psi^{(k-1)}),$$

(тут T – одна зі змінних Y або Ψ).

4. Якщо умова

$$\left\| \left(Y^{(k)}, \Psi^{(k)} \right) - \left(Y^{(k+1)}, \Psi^{(k+1)} \right) \right\| \leq \varepsilon, \quad \varepsilon > 0 \quad (12)$$

не виконується, переходимо до $(k+2)$ -го кроку алгоритму, у противному випадку – до п. 5.

5. Покладемо $Y = 0$, $\Psi = 0$, $\lambda_{ij}(x) = 0$ (), де l – номер ітерації, на якій виконалася умова (12).

6. Обчислимо оптимальне значення цільового функціоналу за формулою (8) при $Y = Y^*$, $\Psi = \Psi^*$ і, для контролю правильності обчислень, за формулою

$$I(\lambda_*) = \sum_{j \in J} \sum_{i \in I} [\Phi_i^j \left(\int_{\Omega} \rho^j(x) \lambda_{*i}^j(x) dx \right) + c^j(x, \tau_i) \rho^j(x) \lambda_{*i}^j(x) dx].$$

Алгоритм описано.

Розв’язання модельних задач. Описаний алгоритм реалізовано для модельних нескінченновимірних транспортних задач. У заданій області розміщені дев’ять підприємств. Вони виробляють продукцію трьох видів для розміщеного в цій області з заданою щільністю споживача, з обмеженнями на потужність підприємств у вигляді рівностей та нерівностей. Вартість транспортування одиниці продукції від підприємства до споживача задається однією з метрик: евклідової, Чебишева або манхеттенської, в залежності від виду продукції. Функція попиту на продукцію у *модельній задачі 1* задається аналітично, а в *модельній задачі 2* дорівнює одиниці.

Модельна задача 1. Задано множину Ω споживачів трьох видів продукції, яка може виготовлятися дев’ятьма підприємствами. Границя області визначена наступним чином $\Omega = \{(x, y) : 0 \leq x \leq 10, 0 \leq y \leq 10\}$. Вартість транспортування одиниці продукції з i -го підприємства $i = \overline{1, 9}$ до споживача (x, y) задається відповідно до виду продукції:

$$c^j(x, y, \tau_i) = \begin{cases} \sqrt{(x - \tau_i^{(1)})^2 + (y - \tau_i^{(2)})^2}, & \text{якщо } j = 1; \\ \max \{ |x - \tau_i^{(1)}|, |y - \tau_i^{(2)}| \}, & \text{якщо } j = 2; \\ |x - \tau_i^{(1)}| + |y - \tau_i^{(2)}|, & \text{якщо } j = 3. \end{cases}$$

Попит $\rho^j(xy)$ на продукцію кожного виду розподілений в області Ω з відповідними щільностями:

$$\rho^1(xy) = \frac{1}{\ln(xy)^2 - 110.003}, \quad \rho^2(xy) = \frac{0.4}{|\lg xy - 110.003|}, \quad \rho^3(xy) = \frac{50}{|xy + 110.003|}.$$

Графічне зображення функцій попиту для трьох видів продукції представлено на рис. 1.

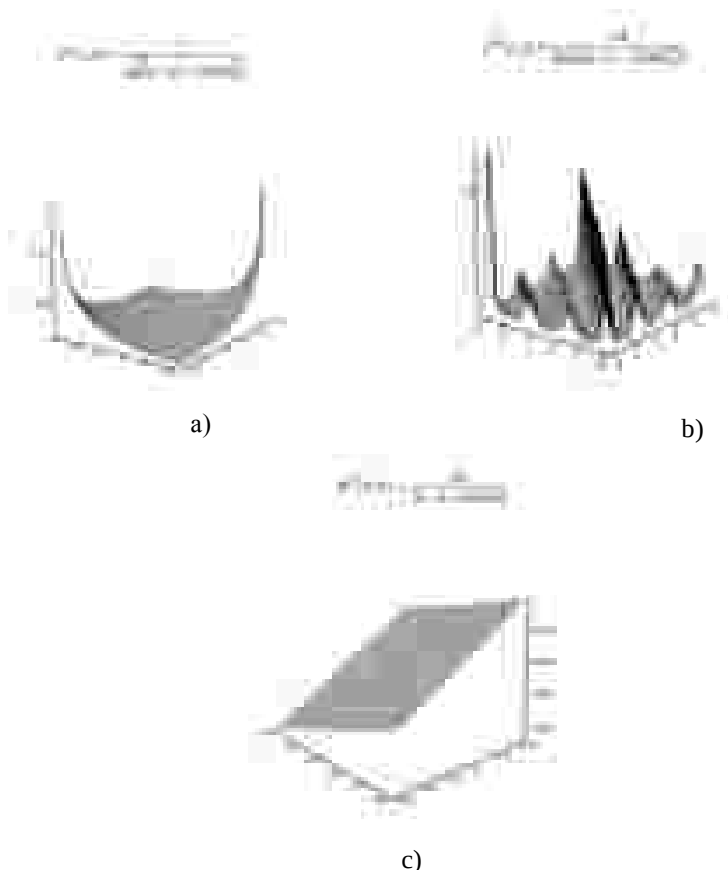


Рис. 1. Графічне зображення функцій попиту $\rho^j(xy)$ для кожного з трьох видів продукції: а) для I-го виду продукції; б) для II-го виду продукції;

с) для III-го виду продукції

Функції $\varphi_{ii}^{jj}(Y)$, які описують залежність вартості виробництва j -го виду продукції на i -ому підприємстві від його потужності, мають вигляд

$$\varphi_i^j(Y_i^j) = (Y_i^j)^2, \quad i = \overline{1, 9}, \quad j = \overline{1, 2, 3},$$

де потужність Y_i^j i -го підприємства по виробництву j -го виду продукції визначається за формулою:

$$Y_i^j = \iint_{\Omega_i^j} \rho^j(xy) dx dy.$$

Потужність i -го пункту виробництва по всім видам продукції визначається сумарним попитом споживачів, які належать Ω_i^j , $i = \overline{1, 9}$, та для пунктів виробництва $i = 1, 2, 4, 5, 7, 9$ не повинна перевищувати задані об'єми, тобто на потужності підприємств накладені наступні обмеження:

$$0 \leq \sum_{j=1}^3 \iint_{\Omega_i^j} \rho^j(xy) dx dy \leq b_i, \quad i = 1, 2, 4, 5, 7, 9,$$

$$b_1 = 100, \quad b_2 = 60, \quad b_4 = 80, \quad b_5 = 17, \quad b_7 = 100, \quad b_9 = 12,$$

а для пунктів виробництва з номерами $i = 3, 6, 8$ повинна дорівнювати

заданим об'ємам:

$$\sum_{j=1}^3 \iint_{\Omega_i^j} \rho^j(xy) dx dy = b_i, \quad i = 3, 6, 8,$$

$$b_3 = 36, \quad b_6 = 5, \quad b_8 = 15.$$

Необхідно розбити множину споживачів Ω на їх зони обслуговування дев'ятьма підприємствами по кожному виду продукції, тобто на підмножини Ω_i^j , $i = \overline{1, 9}$, $j = \overline{1, 3}$, так, щоб мінімізувати функціонал сумарних витрат на виробництво продукції та доставку її до споживача:

$$F(\{\{\Omega_1^1, \dots, \Omega_9^1, \Omega_1^2, \dots, \Omega_9^2, \Omega_1^3, \dots, \Omega_9^3\}\}) =$$

$$= \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} \left[\Phi_{ii}^j \left(\iint_{\Omega_{ij}^0} \rho^j(xy) dx dy \right) + \iint_{\Omega_{ij}^0} c^j(x, \tau) \rho^j(xy) dx dy \right. \\ \left. \left[\Phi_{ii}^j \left(\iint_{\Omega_{ij}^0} \rho^j(xy) dx dy \right) + \iint_{\Omega_{ij}^0} c^j(x, \tau) \rho^j(xy) dx dy \right] \right. .$$

Не є виключенням той випадок, коли деякі з підмножин $\overline{\Omega_i^j}$, $i = 1, 9$, $j = 1, 3$, опиняться порожніми.

Множина Ω покривалася сіткою з вузлами (ij) , $i = 1, \dots, 21$, $j = 1, \dots, 21$.

В якості початкових значень двоїстих змінних задано $\Psi_i^{(0)} = 0$, $i = \overline{1, 9}$; початкові значення потужностей: $Y_1^{j(0)} = 10$, $Y_2^{j(0)} = 100$, $Y_3^{j(0)} = 10$, $Y_4^{j(0)} = 10$, $Y_5^{j(0)} = 100$, $Y_6^{j(0)} = 10$, $Y_7^{j(0)} = 10$, $Y_8^{j(0)} = 100$, $Y_9^{j(0)} = 10$, а також задані координати розміщення підприємств:

$$\tau^0 = \begin{pmatrix} 0.7; 1.4; 4.8; 8.7; 5.1; 8.9; 2.5; 5.8; 8.5 \\ 2.4; 5.6; 4.2; 2.5; 7.5; 8.3; 9.1; 1.5; 5.7 \end{pmatrix} .$$

Умовою завершення обчислень є виконання нерівності:

$$\left\| (Y^{(k)}, \Psi^{(k)}) - (Y^{(k+1)}, \Psi^{(k+1)}) \right\| \leq \varepsilon, \quad \varepsilon = 10^{-3}.$$

За 146 ітерацій отримано:

- максимальне значення функціоналу двоїстої задачі $G^* \approx 4646.87$;
- мінімальне значення функціоналу прямої задачі $F_* \approx 4668.40$;
- оптимальні потужності кожного з дев'яти підприємств:

$$Y_1^* = 24.82; Y_2^* = 20.81; Y_3^* = 35.93; Y_4^* = 24.41; Y_5^* = 16.98;$$

$$Y_6^* = 4.95; Y_7^* = 24.23; Y_8^* = 14.65; Y_9^* = 11.89.$$

Оптимальне розбиття множини споживачів Ω на дев'ять зон обслуговування кожним з дев'яти підприємств по трьом видам продукції представлено на рис. 2.

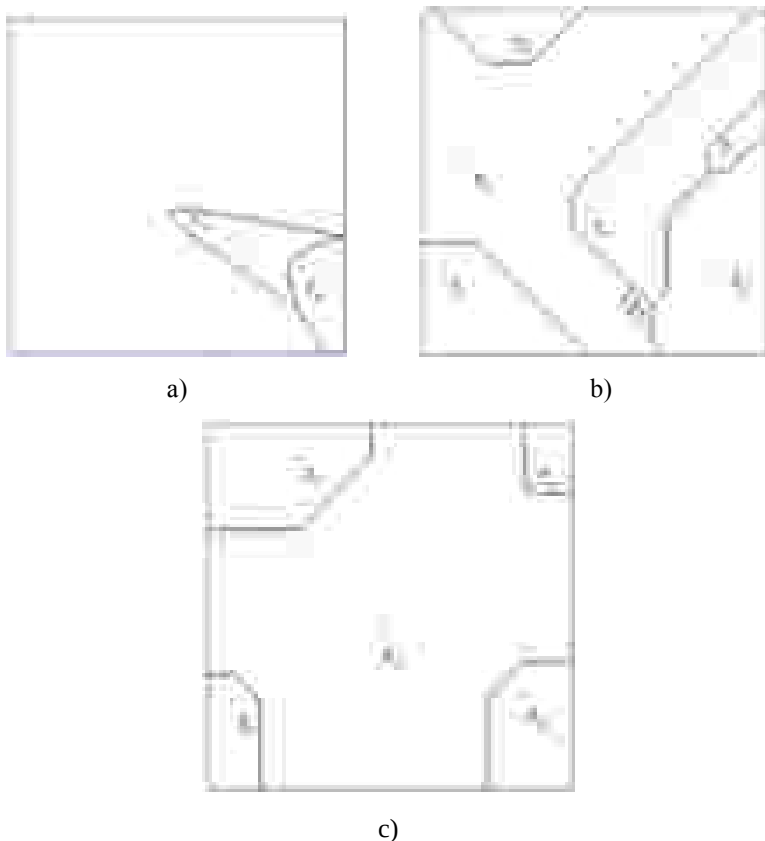


Рис. 2. Оптимальне розбиття множини Ω на зони обслуговування кожним з дев'яти підприємств з фіксованими центрами по трьом видам продукції для модельної задачі 1: а) для I-го виду продукції; б) для II-го виду продукції; в) для III-го виду продукції.

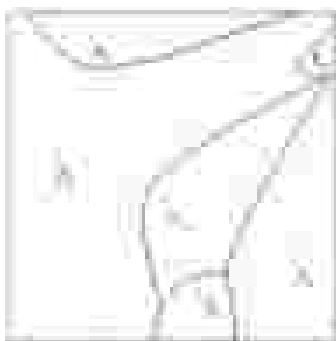
Модельна задача 2. У постановці першої модельної задачі задамо функцію попиту $p^j(x_j) = 1, j = 1, 2, 3$. За таких умов, після 204 ітерації, отримали наступні результати:

- максимальне значення функціоналу двоїстої задачі $G^* \approx 14254.00$;
- мінімальне значення функціоналу прямої задачі $F_* \approx 14242.49$;
- оптимальні потужності кожного з дев'яти підприємств:

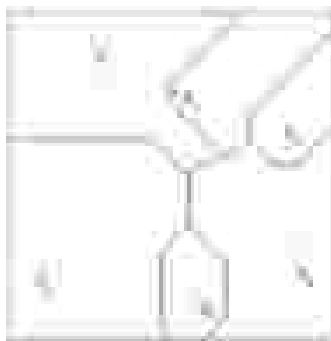
$$Y_1^* = 53.65; Y_2^* = 52.89; Y_3^* = 35.93; Y_4^* = 54.46; Y_5^* = 16.94;$$

$$Y_6^* = 4.95; Y_7^* = 54.02; Y_8^* = 14.91; Y_9^* = 11.98.$$

Оптимальне розбиття множини споживачів Ω на дев'ять зон обслуговування кожним з дев'яти підприємств по трьом видам продукції представлено на рис. 3.



a)



b)



c)

Рис. 3. Оптимальне розбиття множини Ω на зони обслуговування кожним з дев'яти підприємств з фіксованими центрами по трьом видам продукції для модельної задачі 2:

а) для I-го виду продукції; б) для II-го виду продукції; в) для III-го виду продукції.

За результатами розглянутих модельних задач можна зробити наступні висновки:

1) для кожної із задач виконуються умови розв'язності задачі А :

$$S = M \cdot \int_{\Omega} \rho^j(x) dx \leq \sum_{i=1}^9 b_{ib} \quad 0 < b \leq S, \quad i=1,\dots,9, \quad j=1,2,3,$$

тобто загальна оптимальна потужність дев'яти підприємств, яка отримана за алгоритмом розв'язку (для задачі 1 це 178.67, а для задачі 2 це 299.73) не перевищує $S = 425$ – суми заданих об'ємів потужностей підприємств;

2) отримані оптимальні потужності 3_го, 6_го та 8_го підприємства у кожній із задач відповідають обмеженням у вигляді рівностей, тобто дорівнюють заданим значенням, а саме $Y_3^* \approx 36$, $Y_6^* \approx 5$, $Y_8^* \approx 15$;

3) у задачі 1 функція попиту $\rho^j(x)$, $j=1, 2, 3$, тому деякі з підмножин, за рахунок вигляду функції ρ^j , опинились порожніми, що не суперечить постановці вихідної задачі. Так, наприклад, на рисунку 2 а) порожніми є підмножини з номерами 1, 2, 5, 6, 7, 8, 9 і т.п.

Описаний алгоритм реалізовано мовою Visual Fortran 6.5 у середовищі Microsoft Visual Studio.

Висновки. Приведено алгоритм розв'язання нелінійної неперервної багатопродуктової задачі оптимального розбиття множини Ω з n – вимірного евклідового простору на її неперетинні підмножини з фіксованими центрами цих підмножин при обмеженнях у формі рівностей та нерівностей, з функціями попиту та вартості транспортування одиниці продукції, які можуть бути різними для різних видів продукції. Програмна реалізація алгоритму проілюстрована на модельних задачах.

Бібліографічні посилання

1. **Киселёва Е.М.** Непрерывные задачи оптимального разбиения множеств: теория, алгоритмы, приложения / Е.М. Киселёва, Н.З. Шор // Монография. – К., 2005.
2. **Киселёва Е.М.** Решение непрерывной нелинейной задачи оптимального разбиения множеств с размещением центров подмножеств для случая выпуклого целевого функционала / Е.М. Киселёва, М.С. Дунайчук // Кибернетика и системный анализ. — К., 2008. — № 2. — С. 134—152.
3. **Кисельова О.М.** Розв'язання нелінійної неперервної багатопродуктової задачі оптимального розбиття множин з розташуванням центрів підмножин / О.М. Кисельова, В.О. Строева // Питання прикладної математики і математичного моделювання. – Д.: ДНУ, 2010. – С. 155–164.
4. **Кисельова О.М.** Про алгоритм розв'язку нелінійної багатопродуктової задачі оптимального розбиття множин з фіксованими центрами / О.М. Кисельова, В.О. Строева // Математичне та програмне забезпечення інтелектуальних систем: VIII Міжнар. наук.-практ. конф.: тезис докл. – Д. ДНУ, 2010. – С. 104.
5. **Кисельова О.М.** Застосування операторних рівнянь при розв'язку нелінійної багатопродуктової задачі оптимального розбиття множин / О.М. Кисельова, В.О. Строева // Проблеми математичного моделювання: Міжнар. наук.-практ. конф.: тезис докл. – Д. ДДТУ, 2011. – С. 9–11.
6. **Трухаев Р.Н.** Теория неклассических вариационных задач / Р.Н. Трухаев, В.В. Хоменюк – Л., 1971.
7. **Шор Н.З.** Методы минимизации недифференцируемых функций и их приложение. // – К., 1979.

Дата надходження до редколегії: 21.03.11