Київський національний університет імені Тараса Шевченка

Журнал обчислювальної та прикладної

2012 **МАТЕМАТИКИ** 1(107)

Заснований в 1965 році

Адреса редколегії:

Кафедра обчислювальної математики, факультет кібернетики,
Київський національний університет імені Тараса Шевченка,
вул. Володимирська 64, Київ, МСП 01601, Україна.

E-mail: opmjournal@gmail.com html://www.opmj.univ.kiev.ua

Затверджено Вченою радою факультету кібернетики
28 травня 2012 року, протокол №9

Свідоцтво про державну реєстрацію КВ 4246 від 26.05.2000

Підписано до друку 28.05.2012

Для публікації в журналі приймаються не опубліковані раніше наукові праці з провідних методів дослідження та розв'язування задач оптимізації, обчислювальної математики, математичної фізики, теорії фільтрації, теорії пружності, математичної кібернетики, теорії управління та інших напрямків.

У журналі публікуються статті українською, російською й англійською мовами.

Стаття у друкованому вигляді з підписами авторів подається особисто або електронною поштою з подальшим дублюванням на папері в одному примірнику.

До статті додається анкета автора (-ів): ПІБ автора(-ів) українською, російською й англійською мовами, учені ступені та звання, місце роботи та посада, адреса проживання й організації, телефон, факс, електронна пошта тощо та додаткова інформація: назва статті, резюме та ключові слова українською, російською й англійською мовами на електронному носії особисто або електронною поштою (див.“Анкета автора” на сайті журналу www.opmj.univ.kiev.ua) та рецензія сторонньої організації із зазначенням вченого ступеня рецензента.

Обсяг статті не має перевищувати 12 сторінок, включаючи резюме (обсягом до 500 друкованих знаків, яке містить коротку характеристику досліджуваної проблеми, мету роботи й конкретну інформацію про отримані результати), рисунки, таблиці, графіки, список літератури та індекс УДК. Текст статті повинен відповідати структурній схемі: короткий літературний огляд проблеми, що розглядається; мета дослідження; постановка задачі та її актуальність; методи її вирішення й отримані результати; висновки про наукову новизну та практичне значення результатів. Використана література наводиться загальним списком наприкінці статті за порядком її згадування в тексті (посилання у тексті — арабськими цифрами у квадратних дужках) мовою оригіналу, відповідно до ГОСТу 7.1:2006 (ознайомитися можна на сайті). Стаття має бути підготовлена за допомогою редактора LaTeX.

Редактор LaTeX може бути наданий або надісланий на електронному носії поштою на прохання автора. Додаткову інформацію, шаблони статті та анкету автора можна знайти на сайті журналу.

Місцезнаходження редакції: Україна, 01601, м. Київ, просп. Глушкова 4д, Київський національний університет імені Тараса Шевченка, факультет кібернетики, к. 218, редакція журналу “Журнал обчислювальної та прикладної математики”, e-mail: opmjournal@gmail.com.

Журнал обчислювальної та прикладної математики | 2012, №1(107) | Journal of Computational & Applied Mathematics

УДК 517.9

ІТЕРАЦІЙНІ АЛГОРИТМИ ДЛЯ МОНОТОННИХ ДВОРІВНЕВИХ ВАРІАЦІЙНИХ НЕРІВНОСТЕЙ

Р. Я. Апостол, А. А. Гриненко, В. В. Семенов

Резюме. У роботі досліджуються дворівневі монотонні варіаційні нерівності у нескінченновимірному гільбертовому просторі. Запропоновано два алгоритми проекційного типу. Доведено теореми сильної збіжності у випадку сильно монотонних операторів нерівності другого рівня.

Вступ

Варіаційні нерівності — один з центральних об'єктів прикладного нелінійного аналізу [1, 2, 3, 4, 5, 6]. Дворівневі варіаційні нерівності виникають як природнє узагальнення задач лексикографічної (послідовної) оптимізації з двома критеріями [7, 8], а також при аналізі звичайних оптимізаційних задач з обмеженнями у формі варіаційної нерівності [9]. Як самостійний математичний об'єкт, дворівнева варіаційна нерівність у скінченновимірному випадку розглядалась у [10]. Розв'язності більш загальних n-рівневих варіаційних нерівностей та побудові одноетапних алгоритмів їх розв'язання присвячено роботи [11, 12].

Дана стаття пов'язана з роботами [13]–[23]. Для розв'язання дворівневих монотонних варіаційних нерівностей у нескінченновимірному гільбертовому просторі запропонованно два алгоритми проекційного типу. Один з них є своєрідною регуляризацією відомого методу Корпелевич [24].

Основні результати — теореми про сильну збіжність алгоритмів у випадку сильно монотонного оператора нерівності другого рівня.

Опишемо коротко структуру статті. У першому пункті сформульовано дворівневу варіаційну нерівність та основні припушення. Другий пункт присвячено доведенню збіжності запропонованого варіанту алгоритму Корпелевич. У третьому пункті розглянуто алгоритм для дворівневої нерівності з обернено сильно монотонним оператором нерівності першого рівня. В останній, четвертий, пункт винесено доведення двох допоміжних нерівностей, що використовувались у другому пункті.

1. Постановка задачі

Нехай H — дійсний гільбертовий простір з скалярним добутком $(\cdot,\cdot)$ та нормою $\|\cdot\|$. Для оператора $A : H \to H$ та множини $M \subseteq H$ позначимо

$$VI(A, M) = \{x \in M : \ (Ax, y - x) \geq 0 \ \ \forall\, y \in M\}.$$

Через P_C позначимо оператор метричного проектування гільбертового простору H на замкнену опуклу множину $C \subseteq H$.

Нехай:

A1) $C \subseteq H$ — замкнена опукла множина;
A2) $A_1 : H \to H$ — монотонний та L_1-ліпшицевий оператор;
A3) $VI(A_1, C) \neq \emptyset$;
A4) $A_2 : H \to H$ — l_2-сильно монотонний та L_2-ліпшицевий оператор.

Зауваження 1. Множина $VI(A_1, C)$ — замкнена та опукла [1].

Розглянемо задачу:

$$\text{знайти} \;\; x \in VI(A_2, VI(A_1, C)). \tag{1}$$

Зауваження 2. Розв'язок задачі (1) існує та єдиний [1].

2. Аналог алгоритму Корпелевич

Для апроксимації розв'язку задачі (1) використовуємо

Алгоритм 1. Обираємо $x_1 \in H$ та генеруємо послідовність елементів (x_n) за допомогою ітераційної схеми

$$\begin{cases} y_n = P_C(x_n - \lambda_n A_1 x_n), \\ z_n = P_C(x_n - \lambda_n A_1 y_n), \\ x_{n+1} = z_n - \alpha_n A_2 z_n, \end{cases}$$

де $\alpha_n \in [0, 1)$, $\lim_{n\to\infty} \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = +\infty$, $0 < \lambda^* \le \lambda_n \le \lambda^{**} < \frac{1}{L_1}$.

Зауваження 3. Для задачі обчислення проекції $P_{VI(A_1,C)}a$ $(a \in H)$ алгоритм 1 має вигляд

$$\begin{cases} x_1 \in H, \\ y_n = P_C(x_n - \lambda_n A_1 x_n), \\ z_n = P_C(x_n - \lambda_n A_1 y_n), \\ x_{n+1} = \alpha_n a + (1 - \alpha_n) z_n, \end{cases}$$

де $\alpha_n \in [0, 1)$, $\lim_{n\to\infty} \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = +\infty$, $0 < \lambda^* \le \lambda_n \le \lambda^{**} < \frac{1}{L_1}$.

Мають місце наступні тведження (доведення винесені у останній пункт).

Лема 1. *Для $z \in VI(A_1, C)$ та породжених алгоритмом 1 послідовностей (x_n), (y_n) та (z_n) виконується нерівність*

$$\|z_n - z\|^2 \le \|x_n - z\|^2 - (1 - L_1\lambda_n)\,\|x_n - y_n\|^2 - \\ - (1 - L_1\lambda_n)\,\|y_n - z_n\|^2 \quad \forall n \in \mathbb{N}. \tag{2}$$

Лема 2. *Для $z \in VI(A_1, C)$ та породжених алгоритмом 1 послідовностей (x_n), (y_n) та (z_n) виконується нерівність*

$$\|x_{n+1} - z\|^2 - \|x_n - z\|^2 + \|x_{n+1} - z_n\|^2 + (1 - L_1\lambda_n)\,\|x_n - y_n\|^2 + \\ + (1 - L_1\lambda_n)\,\|y_n - z_n\|^2 \le -2\alpha_n (A_2 z_n, x_{n+1} - z) \;\; \forall n \in \mathbb{N}. \tag{3}$$

Перейдемо до доведення збіжності алгоритму 1.

Лема 3. *Породжені алгоритмом 1 послідовності* (x_n), (y_n) *та* (z_n) *— обмежені.*

Доведення. Для $z \in VI(A_1, C)$ маємо

$$\|x_{n+1} - z\| = \|z_n - \alpha_n A_2 z_n - z\| \leq \\ \leq \|(z_n - \alpha_n A_2 z_n) - (z - \alpha_n A_2 z)\| + \alpha_n \|A_2 z\| .$$

Для $\mu > 0$ маємо

$$\|(z_n - \alpha_n A_2 z_n) - (z - \alpha_n A_2 z)\| \leq \left(1 - \frac{\alpha_n}{\mu}\right) \|z_n - z\| + \\ + \frac{\alpha_n}{\mu} \|(z_n - \mu A_2 z_n) - (z - \mu A_2 z)\| .$$

З ліпшицевості та сильної монотонності A_2 випливає

$$\|(x - \mu A_2 x) - (y - \mu A_2 y)\|^2 = \mu^2 \|A_2 x - A_2 y\|^2 + \|x - y\|^2 - \\ - 2\mu(A_2 x - A_2 y, x - y) \leq L_2^2 \mu^2 \|x - y\|^2 + \|x - y\|^2 - 2 l_2 \mu \|x - y\|^2 = \\ = \left(1 - 2 l_2 \mu + L_2^2 \mu^2\right) \|x - y\|^2 \quad \forall x \in H \ \forall y \in H \ \forall \mu > 0. \quad (4)$$

Тому для $\mu \in \left(0, \frac{2l_2}{L_2^2}\right)$ та $\alpha_n \in [0, \mu)$ отримуємо

$$\|(z_n - \alpha_n A_2 z_n) - (z - \alpha_n A_2 z)\| \leq \left(1 - \frac{\alpha_n}{\mu} \beta\right) \|z_n - z\| ,$$

де $\beta = 1 - \sqrt{1 - 2 l_2 \mu + L_2^2 \mu^2} \in (0, 1)$. З леми 1 випливає

$$\|z_n - z\| \leq \|x_n - z\| . \quad (5)$$

Отже, отимали оцінку

$$\|x_{n+1} - z\| \leq \left(1 - \frac{\beta}{\mu} \alpha_n\right) \|x_n - z\| + \frac{\beta}{\mu} \alpha_n \left(\frac{\mu}{\beta} \|A_2 z\|\right) \leq \\ \leq \max \left\{\|x_n - z\| , \frac{\mu}{\beta} \|A_2 z\|\right\} .$$

Оскільки $\lim_{n \to \infty} \alpha_n = 0$, то не зменшуючи загальності, можна вважати $\alpha_n \in [0, \mu)$ для всіх $n \in \mathbb{N}$. Тому,

$$\|x_{n+1} - z\| \leq \max \left\{\|x_1 - z\| , \frac{\mu}{\beta} \|A_2 z\|\right\} . \quad (6)$$

З (6) випливає обмеженість послідовності (x_n). А з нерівності (5) випливає обмеженість послідовності (z_n). З нерівності (2) випливає оцінка

$$\|y_n\|^2 \leq 2 \|x_n\|^2 + 2 \|x_n - y_n\|^2 \leq 2 \|x_n\|^2 + \\ + \frac{2}{(1 - L_1 \lambda_n)} \left\{\|x_n - z\|^2 - \|z_n - z\|^2\right\} ,$$

з якої випливає обмеженість послідовності (y_n). $\square$

Тепер сформулюємо основний результат.

Теорема 1. *Нехай виконуються умови A1) – A4). Тоді породжені алгоритмом 1 послідовності* (x_n), (y_n) *та* (z_n) *сильно збігаються до єдиного розв'язку задачі (1).*

Доведення. Нехай $\bar{x} \in H$ — єдиний розв'язок задачі (1). З леми 3 випливає існування такого $M > 0$, що $|(A_2 z_n, x_{n+1} - \bar{x})| \le M$ для всіх $n \in \mathbb{N}$. Тоді з леми 2, одержуємо оцінку

$$\|x_{n+1} - \bar{x}\|^2 - \|x_n - \bar{x}\|^2 + \|x_{n+1} - z_n\|^2 + (1 - L_1\lambda_n)\|x_n - y_n\|^2 + \\ + (1 - L_1\lambda_n)\|y_n - z_n\|^2 \le 2\alpha_n M. \quad (7)$$

Розглянемо числову послідовність $(\|x_n - \bar{x}\|)$. Можливо два варіанти:

(a) існує номер $\bar{n} \in \mathbb{N}$ такий, що

$$\|x_{n+1} - \bar{x}\| \le \|x_n - \bar{x}\| \quad \forall\, n \ge \bar{n};$$

(b) існує зростаюча послідовність номерів (n_k) така, що

$$\|x_{n_k+1} - \bar{x}\| > \|x_{n_k} - \bar{x}\| \quad \forall\, k \in \mathbb{N}.$$

Розглянемо варіант (a). У цьому випадку існує $\lim_{n\to\infty}\|x_n - \bar{x}\| = c \in \mathbb{R}$. Припустимо, що $c > 0$. Оскільки $\|\bar{x} - x_{n+1}\|^2 - \|x_n - \bar{x}\|^2 \to 0$ та $\alpha_n \to 0$, маємо

$$\|x_{n+1} - z_n\| \to 0, \ \ \|x_n - y_n\| \to 0, \ \ \|y_n - z_n\| \to 0. \quad (8)$$

З сильної монотонності оператора A_2 випливає

$$(A_2 z_n, x_{n+1} - \bar{x}) \ge l_2\|z_n - \bar{x}\|^2 + (A_2\bar{x}, z_n - \bar{x}) + (A_2 z_n, x_{n+1} - z_n). \quad (9)$$

З обмеженості (z_n) та (8) випливає

$$\lim_{n\to\infty}(A_2 z_n, x_{n+1} - z_n) = 0. \quad (10)$$

Числова послідовність $((A_2\bar{x}, z_n - \bar{x}))$ — обмежена. Існує така підпослідовність (z_{n_k}), що

$$\liminf_{n\to\infty}(A_2\bar{x}, z_n - \bar{x}) = \lim_{k\to\infty}(A_2\bar{x}, z_{n_k} - \bar{x}), \quad (11)$$

$$z_{n_k} \rightharpoonup \bar{z} \in C. \quad (12)$$

Покажемо, що $\bar{z} \in VI(A_1, C)$. Маємо

$$\lambda_{n_k}(A_1 y_{n_k}, \eta - z_{n_k}) + (z_{n_k} - x_{n_k}, \eta - z_{n_k}) \ge 0 \ \ \forall\eta \in C.$$

Звідки

$$(A_1\eta, \eta - y_{n_k}) \ge (A_1 y_{n_k}, \eta - y_{n_k}) \ge \\ \ge (A_1 y_{n_k}, z_{n_k} - y_{n_k}) + \frac{(z_{n_k} - x_{n_k}, z_{n_k} - \eta)}{\lambda_{n_k}} \quad \forall\eta \in C. \quad (13)$$

З (8) та (12) випливає $y_{n_k} \rightharpoonup \bar{z}$, $\|z_n - x_n\| \to 0$. Після граничного переходу в (13), одержимо

$$(A_1\eta, \eta - \bar{z}) \geq \limsup_{k\to\infty}(A_1 y_{n_k}, \eta - y_{n_k}) \geq$$
$$\geq \lim_{k\to\infty}(A_1 y_{n_k}, z_{n_k} - y_{n_k}) \geq 0 \quad \forall \eta \in C,$$

тобто, $\bar{z} \in VI(A_1, C)$. Таким чином, в (11) одержимо

$$\liminf_{n\to\infty}(A_2\bar{x}, z_n - \bar{x}) = \lim_{k\to\infty}(A_2\bar{x}, z_{n_k} - \bar{x}) = (A_2\bar{x}, \bar{z} - \bar{x}) \geq 0. \tag{14}$$

З нерівностей

$$\|x_n - \bar{x}\| - \|z_n - x_n\| \leq \|z_n - \bar{x}\| \leq \|x_n - \bar{x}\| + \|z_n - x_n\|$$

випливає

$$\lim_{n\to\infty}\|z_n - \bar{x}\| = c. \tag{15}$$

Урахувавши (10), (14), та (15) в (9), одержимо

$$\liminf_{n\to\infty}(A_2 z_n, x_{n+1} - \bar{x}) \geq l_2 c^2 > 0. \tag{16}$$

Оберемо $\delta \in (0, l_2 c^2)$. Існує таке $n_* \in \mathbb{N}$, що для всіх $n \geq n_*$

$$(A_2 z_n, x_{n+1} - \bar{x}) \geq \delta.$$

Тоді з нерівності (3) випливає, що для всіх $n \geq n_*$ маємо $\|x_{n+1} - \bar{x}\|^2 - \|x_n - \bar{x}\|^2 \leq -2\delta\alpha_n$. Звідки

$$\|x_{n+1} - \bar{x}\|^2 \leq \|x_{n_*} - \bar{x}\|^2 - 2\delta\sum_{i=n_*}^{n}\alpha_i.$$

З умови $\sum_{n=1}^{\infty}\alpha_n = +\infty$ випливає $\|x_n - \bar{x}\| \to -\infty$, що неможливо. Отже, $c = 0$, тобто, $\lim_{n\to\infty}\|x_n - \bar{x}\| = 0$. Ясно, що

$$\lim_{n\to\infty}\|y_n - \bar{x}\| = \lim_{n\to\infty}\|z_n - \bar{x}\| = 0.$$

Розглянемо варіант (b). Використаємо прийом з роботи [25]. У цьому випадку можна розглянути послідовність номерів

$$\pi_n = \max\left\{n_1 \leq k \leq n : \ \|x_{k+1} - \bar{x}\| > \|x_k - \bar{x}\|\right\}.$$

Послідовність (π_n) має властивості:

(i) $\pi_n \nearrow +\infty$;

(ii) $\|x_{\pi_n+1} - \bar{x}\| > \|x_{\pi_n} - \bar{x}\|$ для всіх $n \geq n_1$;

(iii) $\|x_{\pi_n+1} - \bar{x}\| \geq \|x_n - \bar{x}\|$ для всіх $n \geq n_1$.

З (7) та (ii) випливає

$$\|x_{\pi_n+1} - z_{\pi_n}\|^2 + (1 - L_1\lambda_{\pi_n})\|x_{\pi_n} - y_{\pi_n}\|^2 +$$
$$+ (1 - L_1\lambda_{\pi_n})\|y_{\pi_n} - z_{\pi_n}\|^2 \leq 2\alpha_{\pi_n} M.$$

Звідси

$$\lim_{n\to\infty}\|x_{\pi_n+1} - z_{\pi_n}\| = 0, \quad \lim_{n\to\infty}\|x_{\pi_n} - y_{\pi_n}\| = 0, \quad \lim_{n\to\infty}\|y_{\pi_n} - z_{\pi_n}\| = 0.$$

Доведемо сильну збіжність послідовності (z_{π_n}) до точки $\bar{x}$. З обмеженості (z_{π_n}) випливає існування підпослідовності $(z_{\pi_{n_k}})$, слабко збіжної до деякої точки $\bar{z} \in C$. Як і попередньому випадку показуємо, що $\bar{z} \in VI(A_1, C)$. З (ii) та нерівності (2) випливає

$$(A_2 z_{\pi_n}, x_{\pi_n+1} - \bar{x}) < 0 \quad \forall n \geq n_1. \tag{17}$$

Використовуючи сильну монотонність оператора A_2 для $n \geq n_1$, отримуємо

$$(A_2 z_{\pi_n} - A_2\bar{x}, z_{\pi_n} - \bar{x}) = (A_2 z_{\pi_n}, x_{\pi_n+1} - \bar{x}) + (A_2 z_{\pi_n}, z_{\pi_n} - x_{\pi_n+1}) - \\ - (A_2\bar{x}, z_{\pi_n} - \bar{x}) \geq l_2 \left\| z_{\pi_n} - \bar{x} \right\|^2.$$

Урахувавши (17), маємо

$$\left\| z_{\pi_n} - \bar{x} \right\|^2 \leq \left\{ (A_2 z_{\pi_n}, z_{\pi_n} - x_{\pi_n+1}) - (A_2\bar{x}, z_{\pi_n} - \bar{x}) \right\} / l_2. \tag{18}$$

Оскільки

$$\lim_{n\to\infty} (A_2 z_{\pi_n}, z_{\pi_n} - x_{\pi_n+1}) = 0, \quad \lim_{k\to\infty} (A_2\bar{x}, z_{\pi_{n_k}} - \bar{x}) = (A_2\bar{x}, \bar{z} - \bar{x}),$$

то з (18) випливає $\limsup_{k\to\infty} \|z_{\pi_{n_k}} - \bar{x}\|^2 \leq -(A_2\bar{x}, \bar{z} - \bar{x})/l_2 \leq 0$. Отже, $\lim_{k\to\infty} \left\| z_{\pi_{n_k}} - \bar{x} \right\| = 0$. З єдиності $\bar{x}$ та $\bar{z} = \bar{x}$ випливає $\lim_{n\to\infty} \|z_{\pi_n} - \bar{x}\| = 0$. З нерівності $\|x_{\pi_n+1} - \bar{x}\| \leq \|x_{\pi_n+1} - z_{\pi_n}\| + \|z_{\pi_n} - \bar{x}\|$ випливає

$$\lim_{n\to\infty} \|x_{\pi_n+1} - \bar{x}\| = 0.$$

Ураховуючи (iii), отримуємо

$$\lim_{n\to\infty} \|x_n - \bar{x}\| = 0. \tag{19}$$

З (19) та (7) випливає $\lim_{n\to\infty} \|x_n - y_n\| = \lim_{n\to\infty} \|z_n - y_n\| = 0$, а тому $\lim_{n\to\infty} \|y_n - \bar{x}\| = \lim_{n\to\infty} \|z_n - \bar{x}\| = 0$. □

3. Алгоритм для задачі з обернено сильно монотонним оператором

У випадку більш підсиленої монотонності операторів нерівності першого рівня для розв'язання задачі (1) можна використовувати методи з одним проектуванням на кроці.

Нехай:

A5) $A_1 : H \to H$ — L_1-обернено сильно монотонний оператор[1].

[1] Оператор $A : D(A) \to H$ називають обернено сильно монотонним (ко-коерцитивним), якщо існує додатня константа α така, що

$$(Ax - Ay, x - y) \geq \alpha \|Ax - Ay\|^2 \quad \forall x, y \in D(A).$$

У цьому випадку кажуть, що A — α-обернено сильно монотонний. Ясно, що α-обернено сильно монотонний оператор є $1/\alpha$-ліпшицевим. Якщо оператор A є μ-сильно монотонним та L-ліпшицевим, то він буде μ/L^2-обернено сильно монотонним. Якщо g — заданий на замкненій опуклій множині C опуклий диференційовний функціонал з похідною, що задовольняє умову Ліпшиця зі сталою $L > 0$, то похідна ∇g — $1/L$-обернено сильно монотонний на C оператор [26].

Для L_1-обернено сильно монотонного оператора $A_1 : H \to H$ та $\gamma > 0$ має місце

$$\begin{aligned}
&\|P_C(x-\gamma A_1 x) - P_C(y-\gamma A_1 y)\|^2 \le \\
&\qquad \le \|(x-\gamma A_1 x) - (y-\gamma A_1 y)\|^2 = \\
&= \|x-y\|^2 - 2\gamma(x-y, A_1x - A_1y) + \gamma^2 \|A_1 x - A_1 y\|^2 \le \\
&\le \|x-y\|^2 - 2\gamma L_1 \|A_1 x - A_1 y\|^2 + \gamma^2 \|A_1 x - A_1 y\|^2 = \\
&\qquad = \|x-y\|^2 - \gamma(2L_1 - \gamma)\|A_1 x - A_1 y\|^2 . \qquad (20)
\end{aligned}$$

Таким чином, при $\gamma \in (0, 2L_1]$ оператор $P_C(I - \gamma A_1)$ — нерозтягуючий.

Розглянемо наступний

Алгоритм 2. Обираємо $x_1 \in H$ та генеруємо послідовність елементів (x_n) за допомогою ітераційної схеми

$$\begin{cases} y_n = P_C(x_n - \lambda_n A_1 x_n), \\ x_{n+1} = y_n - \alpha_n A_2 y_n, \end{cases}$$

де $\alpha_n \in [0,1)$, $\lim_{n\to\infty} \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = +\infty$, $\sum_{n=1}^{\infty} |\alpha_{n+1} - \alpha_n| < +\infty$ або $\lim_{n\to\infty}(\alpha_{n+1} - \alpha_n)/\alpha_{n+1} = 0$, $0 < \lambda^* \le \lambda_n \le \lambda^{**} < 2L_1$.

Зауваження 4. Даний алгоритм запропоновано у [13]. Чисельні експеременти проведено у бакалаврській роботі А.А. Гриненка.

Лема 4. *Породжені алгоритмом 2 послідовності* (x_n), (y_n) *— обмежені.*

Доведення. Для $z \in VI(A_1, C)$ маємо

$$\begin{aligned}
&\|x_{n+1} - z\| = \|y_n - \alpha_n A_2 y_n - z\| \le \\
&\qquad \le \|(y_n - \alpha_n A_2 y_n) - (z - \alpha_n A_2 z)\| + \alpha_n \|A_2 z\| .
\end{aligned}$$

Для $\mu > 0$ маємо

$$\begin{aligned}
&\|(y_n - \alpha_n A_2 y_n) - (z - \alpha_n A_2 z)\| \le \left(1 - \frac{\alpha_n}{\mu}\right)\|y_n - z\| + \\
&\qquad + \frac{\alpha_n}{\mu}\|(y_n - \mu A_2 z_n) - (z - \mu A_2 z)\| .
\end{aligned}$$

Використовуючи (4), для $\mu \in \left(0, \frac{2l_2}{L_2^2}\right)$ та $\alpha_n \in [0,\mu)$, отримуємо

$$\|(y_n - \alpha_n A_2 y_n) - (z - \alpha_n A_2 z)\| \le \left(1 - \frac{\alpha_n}{\mu}\beta\right)\|y_n - z\| ,$$

де $\beta = 1 - \sqrt{1 - 2l_2\mu + L_2^2\mu^2} \in (0,1)$. З (20) та рівності $z = P_C(z - \lambda_n A_1 z)$ випливає

$$\|y_n - z\| \le \|x_n - z\| . \qquad (21)$$

Отже, отимали оцінку

$$\|x_{n+1} - z\| \le \left(1 - \frac{\beta}{\mu}\alpha_n\right)\|x_n - z\| + \frac{\beta}{\mu}\alpha_n\left(\frac{\mu}{\beta}\|A_2 z\|\right) \le$$
$$\le \max\left\{\|x_n - z\|, \frac{\mu}{\beta}\|A_2 z\|\right\}.$$

Оскільки $\lim_{n\to\infty}\alpha_n = 0$, то не зменшуючи загальності, можна вважати $\alpha_n \in [0,\mu)$ для всіх $n \in \mathbb{N}$. Тому,

$$\|x_{n+1} - z\| \le \max\left\{\|x_1 - z\|, \frac{\mu}{\beta}\|A_2 z\|\right\}. \tag{22}$$

З (22) випливає обмеженість послідовності (x_n). А з нерівності (21) випливає обмеженість послідовності (y_n). □

Для доведення збіжності алгоритму 2 використаємо такий факт.

Лема 5 ([27])**.** *Нехай (ξ_n) — послідовність невід'ємних чисел, що задовольняє рекурентну нерівність $\xi_{n+1} \le (1-\alpha_n)\xi_n + \alpha_n\beta_n + \gamma_n$ $\forall n \in \mathbb{N}$, де послідовності (α_n), (β_n) і (γ_n) мають властивості: 1) $\alpha_n \in [0,1)$ та $\sum_{n=1}^{\infty}\alpha_n = +\infty$; 2) $\limsup_{n\to\infty}\beta_n \le 0$; 3) $\gamma_n \in [0,+\infty)$ та $\sum_{n=1}^{\infty}\gamma_n < +\infty$. Тоді*

$$\lim_{n\to\infty}\xi_n = 0.$$

Покажемо, що

$$\lim_{n\to\infty}\|x_{n+1} - y_n\| = 0 \tag{23}$$

та

$$\lim_{n\to\infty}\|x_{n+1} - x_n\| = 0. \tag{24}$$

Маємо

$$\|x_{n+1} - y_n\| = \alpha_n\|A_2 y_n\|.$$

З обмеженості (y_n) та $\lim_{n\to\infty}\alpha_n = 0$ випливає (23).

Оцінимо $\|x_{n+1} - x_n\|$. Маємо

$$\|x_{n+1} - x_n\| = \|(y_n - \alpha_n A_2 y_n) - (y_{n-1} - \alpha_{n-1}A_2 y_{n-1})\| \le$$
$$\le \|(y_n - \alpha_n A_2 y_n) - (y_{n-1} - \alpha_n A_2 y_{n-1})\| + |\alpha_n - \alpha_{n-1}|\,\|A_2 y_{n-1}\| \le$$
$$\le \left(1 - \frac{\beta}{\mu}\alpha_n\right)\|y_n - y_{n-1}\| + |\alpha_n - \alpha_{n-1}|\,\|A_2 y_{n-1}\|.$$

Звідки

$$\|x_{n+1} - x_n\| \le \left(1 - \frac{\beta}{\mu}\alpha_n\right)\|x_n - x_{n-1}\| + |\alpha_n - \alpha_{n-1}|\,\|A_2 y_{n-1}\|.$$

З леми 5 випливає (24).

З (23) та (24) випливає

$$\|x_n - y_n\| \le \|x_n - x_{n+1}\| + \|x_{n+1} - y_n\| \to 0. \tag{25}$$

Нехай $\bar{x} \in H$ — єдиний розв'язок задачі (1). Доведемо, що

$$\limsup_{n\to\infty}(-A_2\bar{x}, x_n - \bar{x}) \le 0. \tag{26}$$

Виділимо з (x_n) таку підпослідовність (x_{n_k}), що

$$\limsup_{n\to\infty}(-A_2\bar{x}, x_n - \bar{x}) = \lim_{k\to\infty}(-A_2\bar{x}, x_{n_k} - \bar{x}).$$

Можна вважати, що $x_{n_k} \rightharpoonup \tilde{x}$. З (25) випливає, що $y_{n_k} \rightharpoonup \tilde{x}$ та $\tilde{x} \in C$. Покажемо, що $\tilde{x} \in VI(A_1, C)$. Маємо

$$\lambda_{n_k}(A_1 x_{n_k}, \eta - y_{n_k}) + (y_{n_k} - x_{n_k}, \eta - y_{n_k}) \geq 0 \ \ \forall \eta \in C.$$

Звідки

$$(A_1\eta, \eta - x_{n_k}) \geq (A_1 x_{n_k}, \eta - x_{n_k}) \geq$$
$$\geq (A_1 x_{n_k}, y_{n_k} - x_{n_k}) + \frac{(y_{n_k} - x_{n_k}, y_{n_k} - \eta)}{\lambda_{n_k}} \quad \forall \eta \in C. \quad (27)$$

Після граничного переходу в (27), одержимо

$$(A_1\eta, \eta - \tilde{x}) \geq \limsup_{k\to\infty}(A_1 x_{n_k}, \eta - x_{n_k}) \geq$$
$$\geq \lim_{k\to\infty}\left\{(A_1 x_{n_k}, y_{n_k} - x_{n_k}) + \frac{(y_{n_k} - x_{n_k}, y_{n_k} - \eta)}{\lambda_{n_k}}\right\} \geq 0 \quad \forall \eta \in C$$

тобто, $\tilde{x} \in VI(A_1, C)$. Тому

$$\lim_{k\to\infty}(-A_2\bar{x}, x_{n_k} - \bar{x}) = -(A_2\bar{x}, \tilde{x} - \bar{x}) \leq 0,$$

чим і доводимо (26).

Покажемо тепер, що $x_n \to \bar{x}$. Маємо

$$\|x_{n+1} - \bar{x}\|^2 = \|y_n - \alpha_n A_2 y_n - \bar{x}\|^2 =$$
$$= \|(y_n - \alpha_n A_2 y_n) - (\bar{x} - \alpha_n A_2\bar{x}) + \alpha_n A_2\bar{x}\|^2 \leq$$
$$\leq \|(y_n - \alpha_n A_2 y_n) - (\bar{x} - \alpha_n A_2\bar{x})\|^2 + 2\left(-\alpha_n A_2\bar{x}, x_{n+1} - \bar{x}\right) \leq$$
$$\leq \left(1 - \frac{\beta}{\mu}\alpha_n\right)\|x_n - \bar{x}\|^2 + 2\alpha_n\left(-A_2\bar{x}, x_{n+1} - \bar{x}\right). \quad (28)$$

Застосувавши до (28) лему 5, робимо висновок, що $\|x_n - \bar{x}\| \to 0$.

Отже, має місце

Теорема 2. *Нехай виконуються умови A1), A5), A3) та A4). Тоді породжені алгоритмом 2 послідовності (x_n) та (y_n) сильно збігаються до єдиного розв'язку задачі (1).*

4. Доведення допоміжних нерівностей (1), (2)

У даному пункті наведені доведення лем 1, 2. Нерівності з цих лем — основні складові доведення сильної збіжності алгоритму 1.

Доведення леми 1. Оскільки $z_n = P_C(x_n - \lambda_n A_1 y_n)$, то

$$\lambda_n(A_1 y_n, \eta - z_n) + (z_n - x_n, \eta - z_n) \geq 0 \ \ \forall \eta \in C. \quad (29)$$

Поклавши в (29) $\eta = z \in VI(A_1, C)$ та врахувавши нерівність

$$(A_1 y_n, z - y_n) \leq 0,$$

одержимо

$$(z_n - x_n, z - z_n) \geq \lambda_n (A_1 y_n, z_n - y_n). \tag{30}$$

Оскільки

$$(A_1 y_n, z_n - y_n) \geq (A_1 x_n, z_n - y_n) - \frac{L_1}{2} \|x_n - y_n\|^2 - \frac{L_1}{2} \|y_n - z_n\|^2 ,$$

то з (30) випливає нерівність

$$(z_n - x_n, z - z_n) \geq \lambda_n (A_1 x_n, z_n - y_n) - \\ - \frac{L_1}{2} \lambda_n \|x_n - y_n\|^2 - \frac{L_1}{2} \lambda_n \|y_n - z_n\|^2 . \tag{31}$$

Оскільки $y_n = P_C(x_n - \lambda_n A_1 x_n)$, то

$$\lambda_n (A_1 x_n, z_n - y_n) + (y_n - x_n, z_n - y_n) \geq 0. \tag{32}$$

Урахувавши (32) в (31), прийдемо до нерівності

$$(z_n - x_n, z - z_n) \geq (y_n - x_n, y_n - z_n) - \\ - \frac{L_1}{2} \lambda_n \|x_n - y_n\|^2 - \frac{L_1}{2} \lambda_n \|y_n - z_n\|^2 . \tag{33}$$

Має місце рівність

$$2(z_n - x_n, z - z_n) = \|x_n - z\|^2 - \|z_n - x_n\|^2 - \|z_n - z\|^2 .$$

Тому (33) можна переписати у вигляді

$$\|x_n - z\|^2 - \|z_n - x_n\|^2 - \|z_n - z\|^2 \geq \\ \geq 2(y_n - x_n, y_n - z_n) - L_1 \lambda_n \|x_n - y_n\|^2 - L_1 \lambda_n \|y_n - z_n\|^2 .$$

Звідки

$$\|z_n - z\|^2 \leq \|x_n - z\|^2 - \|z_n - x_n\|^2 - 2(y_n - x_n, y_n - z_n) + \\ + L_1 \lambda_n \|x_n - y_n\|^2 + L_1 \lambda_n \|y_n - z_n\|^2 .$$

Оскільки

$$\|z_n - x_n\|^2 = \|z_n - y_n\|^2 + \|y_n - x_n\|^2 + 2(y_n - x_n, z_n - y_n),$$

то

$$\|z_n - z\|^2 \leq \|x_n - z\|^2 - \|z_n - y_n\|^2 - \|y_n - x_n\|^2 + \\ + L_1 \lambda_n \|x_n - y_n\|^2 + L_1 \lambda_n \|y_n - z_n\|^2 = \|x_n - z\|^2 - \\ - (1 - L_1 \lambda_n) \|x_n - y_n\|^2 - (1 - L_1 \lambda_n) \|y_n - z_n\|^2 ,$$

що і треба було довести. □

Доведення леми 2. Маємо

$$\|x_{n+1} - z\|^2 = \|z_n - \alpha_n A_2 z_n - z\|^2 = \\ = \|z_n - z\|^2 - 2\alpha_n (A_2 z_n, z_n - z) + \alpha_n^2 \|A_2 z_n\|^2 = \\ = \|z_n - z\|^2 - 2\alpha_n (A_2 z_n, x_{n+1} - z) - \|x_{n+1} - z_n\|^2 .$$

З леми 1 випливає

$$\|x_{n+1} - z\|^2 \leq \|x_n - z\|^2 - \\ - (1 - L_1\lambda_n) \|x_n - y_n\|^2 - (1 - L_1\lambda_n) \|y_n - z_n\|^2 - \\ - \|x_{n+1} - z_n\|^2 - 2\alpha_n(A_2 z_n, x_{n+1} - z).$$

Звідки випливає нерівність (3). □

Література

1. Лионс Ж.-Л. Некоторые методы решения нелинейных краевых задач / Ж.-Л. Лионс. — Москва: Мир, 1972. — 587 с.
2. Гловински Р. Численное исследование вариационных неравенств / Р. Гловински, Ж.-Л. Лионс, Р. Тремольер. — Москва: Мир, 1979. — 576 с.
3. Дюво Г. Неревенства в механике и физике / Г. Дюво, Ж.-Л. Лионс. — Москва: Наука, 1980. — 382 с.
4. Байокки К. Вариационные и квазивариационные неравенства / К. Байокки, А. Капело. — Москва: Наука, 1988. — 448 с.
5. Обен Ж.-П. Прикладной нелинейный анализ / Ж.-П. Обен, И. Экланд. — Москва: Мир, 1988. — 510 с.
6. Nagurney A. Network economics: A variational inequality approach / A. Nagurney. — Dordrecht: Kluwer Academic Publishers, 1999. — xxi + 325 p.
7. Еремин И. И. О задачах последовательного программирования / И. И. Еремин // Сиб. мат. журн. — 1973. — 14, №1. — С. 53–63.
8. Подиновский В. В. Оптимизация по последовательно применяемым критериям / В. В. Подиновский, В. Н. Гаврилов. — Москва: Советское радио, 1975. — 192 с.
9. Ye J. J. Necessary Optimality Conditions for Optimization Problems with Variational Inequality Constraints / J. J. Ye, X. Y. Ye // Mathematics of Operations Research. — 1997. — V. 22, №4. — P. 977–997.
10. Калашников В. В. Решение двухуровневого вариационного неравенства / В. В. Калашников, Н. И. Калашникова // Кибернетика и системный анализ. — 1994. — №4. — С. 178–180.
11. Коннов И. В. О системах вариационных неравенств / И. В. Коннов // Изв. вузов. Матем. — 1997. — №12. — С. 79–88.
12. Попов Л. Д. Лексикографические вариационные неравенства и некоторые приложения / Л. Д. Попов // Математическое программирование. Регуляризация и аппроксимация, Сборник статей. — Тр. ИММ. — 8, №1. — 2002. — С. 103–115.
13. Войтова Т. А. Ітераційні алгоритми для опуклих дворівневих задач оптимізації / Т. А. Войтова, А. А. Гриненко // III Міжнародна конференція "Обчислювальна та прикладна математика"(присвячена пам'яті академіка НАНУ І. І. Ляшка), Україна, Київ, 11–12 вересня 2009 року, Матеріали конференції, Київ, 2009. — С. 27.
14. Войтова Т. А. Метод решения двухуровневого операторного включения / Т. А. Войтова, В. В. Семенов // XVI Int. Conf."PDMU-2010". Abstracts. — Yalta, Ukraine. — 2010. — P. 44.

15. Семенов В. В. О сходимости методов решения двухуровневых вариационных неравенств с монотонными операторами / В. В. Семенов // Журн. обчисл. та прикл. матем. — 2010. — №1 (100). — С. 121–129.
16. Войтова Т. А. Метод решения двухэтапных операторных включений / Т. А. Войтова, В. В. Семенов // Журн. обчисл. та прикл. матем. — 2010. — №3 (102). — С. 34–39.
17. Войтова Т. А. Экономичная схема для задач с ограничением в виде вариационного неравенства / Т. А. Войтова, В. В. Семенов// Тези XVII Int. Conf. "PDMU–2011". Abstracts. — Skhidnytsia, Ukraine. — 2011. — P. 51.
18. Ляшко С. И. Экономичная модификация метода Корпелевич для монотонных задач о равновесии / С. И. Ляшко, Т. А. Войтова, В. В. Семенов // Кибернетика и системный анализ. — 2011. — №4. — С. 146–154.
19. Апостол Р. Я. Сильно збіжний варіант методу Корпелевич для задач про рівновагу / Р. Я. Апостол, С. В. Денисов, В. В. Семенов // IV Міжнародна конференція "Обчислювальна та прикладна математика". Матеріали конференції. — Київ, 2011. — С. 40.
20. Войтова Т. А. Збіжність альтернуючого проксимального алгоритму для задачі дворівневої опуклої мінімізації / Т. А. Войтова, В. В. Семенов // IV Міжнародна конференція "Обчислювальна та прикладна математика". Матеріали конференції. — Київ, 2011. — С. 58.
21. Войтова Т. А. Методи регуляризації та декомпозиції варіаційних задач / Т. А. Войтова, Ю. В. Маліцький, В. В. Семенов // Праці міжнародної молодіжної математичної школи "Питання оптимізації обчислень (ПОО–XXXVII)— Київ: Інститут кібернетики імені В.М. Глушкова НАН України, 2011. — С. 30–31.
22. Войтова Т. А. Сильно збіжний модифікований варіант методу Корпелевич для задач рівноважного програмування / Т. А. Войтова, С. В. Денисов, В. В. Семенов // Журн. обчисл. та прикл. матем. — 2011. — №1 (104). — С. 10–23.
23. Денисов С. В. Проксимальний алгоритм для дворівневих варіаційних нерівностей: сильна збіжність / С. В. Денисов, В. В. Семенов // Журн. обчисл. та прикл. матем. — 2011. — №3 (106). — С. 27–32.
24. Корпелевич Г. М. Экстраградиентный метод для отыскания седловых точек и других задач / Г. М. Корпелевич // Экономика и математические методы. — 1976. — Т. 12, №4. — С. 747–756.
25. Mainge P.–E. Strong Convergence of Projected Subgradient Methods for Nonsmooth and Nonstrictly Convex Minimization / P.–E. Mainge // Set–Valued Analysis. — 2008. — V. 16. — P. 899–912.
26. Гольштейн Е. Г. Модифицированные функции Лагранжа. Теория и методы оптимизации / Е. Г. Гольштейн, Н. В. Третьяков. — Москва: Наука, 1989. — 400 с.
27. Xu H. K. Another control condition in an iterative method for nonexpansive mappings / H. K. Xu // Bull. Austral. Math. Soc.— 2002. — 65 — P. 109–113.

Факультет кібернетики, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64, Київ, 01601, Україна.

Надійшла 05.09.2011

Журнал обчислювальної та прикладної математики 2012, №1(107) Journal of Computational & Applied Mathematics

UDC 517.9

OPTIMAL CONTROL PROBLEMS IN COEFFICIENTS FOR DEGENERATE VARIATIONAL INEQUALITIES OF MONOTONE TYPE. II. ATTAINABILITY PROBLEM

O. P. Kupenko

Abstract. In this paper we study an optimal control problem for a nonlinear monotone variational inequality with degenerate weight function and with the coefficients which we adopt as controls in $L^\infty(\Omega)$. Since these types of variational inequalities can exhibit the Lavrentieff phenomenon, we consider the optimal control problem in coefficients in the so-called class of H-admissible solutions. We prove attainability of H-optimal pairs via optimal solutions of some non-degenerate perturbed optimal control problems.x

Introduction

The aim of this paper is to study optimal control problems (OCPs) associated to nonlinear degenerate elliptic variational inequalities. The control is a matrix of coefficients in the main part of nonlinear elliptic operator. It is well known, that degenerate control problems of this type may admit non-uniqueness of admissible solution classes, which implies non-uniqueness of optimal solutions of particular kind. Here we consider the optimal control problem in coefficients in the so-called class of H-admissible solutions. Mainly, we are interested about attainability of H-optimal solutions to degenerate problems via optimal solutions of non-degenerate problems.

More precisely, we consider the following OCP

$$I(\mathcal{U}, y) = \int_\Omega |y(x) - z_\partial(x)|^p \, dx \to \inf, \tag{1}$$

$$\mathcal{U} \in M_p^{\alpha,\beta}(\Omega), \ y \in K, \tag{2}$$

$$\langle -\mathrm{div}\left(\mathcal{U}(x)\rho(x)\left[(\nabla y)^{p-2}\right]\nabla y\right) + |y|^{p-2}y, v - y\rangle_W \geq \langle f, v - y\rangle_W, \ \forall v \in K \tag{3}$$

where $[\eta^{p-2}] = \mathrm{diag}\{|\eta_1|^{p-2}, |\eta_2|^{p-2}, \dots, |\eta_N|^{p-2}\} \ \forall \eta \in \mathbb{R}^N$.

Here, Ω is a bounded open subset of $\mathbb{R}^N$ $(N \geq 1)$ with Lipschitz boundary, $\rho > 0$ is a weight function, $z_\partial \in L^p(\Omega)$ and $f \in L^q(\Omega)$ — fixed elements, $M_p^{\alpha,\beta}(\Omega) \subset L^\infty(\Omega; \mathbb{R}^{N\times N})$ is a class of admissible controls, $K \subset W$ — is a closed convex subset, and $W = W(\Omega, \rho\, dx)$ is a set of functions $y \in W_0^{1,1}(\Omega)$

for which the norm

$$\|y\|_\rho = \left(\int_\Omega \left(|y|^p + \rho \sum_{i=1}^N \left|\frac{\partial y}{\partial x_i}\right|^p\right) dx\right)^{1/p} \tag{4}$$

is finite.

Let p be a real number such that $2 \le p < \infty$ and let q be its conjugate, namely $p^{-1} + q^{-1} = 1$. We say that a weight function $\rho = \rho(x)$ is degenerate in $\mathbb{R}^N$ if

$$\rho(x) > 0 \quad \text{a.e. in } \mathbb{R}^N \quad \text{and} \quad \rho + \rho^{-1/(p-1)} \in L^1_{loc}(\mathbb{R}^N), \tag{5}$$

and the sum $\rho + \rho^{-1/(p-1)}$ does not belong to $L^\infty(\Omega)$, in general.

Dealing with degenerate problems leads us to the concept of weighted Sobolev spaces such as $W(\Omega, \rho\, dx)$. In general, these spaces are not new in the literature (see [7], [6]). They allow to enlarge the class of boundary value problems which are solvable by functional-analytical methods. In fact, we consider variational inequality (3) with degenerate weight ρ which is not bounded away from zero and infinity but only satisfying local integrability conditions (5). Under these assumptions the nonlinear differential operator in (3) is not coercive in the classical sense. Here we encounter non-uniqueness of a particular kind: the smooth functions are, in general, not dense in the weighted Sobolev space $W(\Omega, \rho\, dx)$; that is, if $H(\Omega, \rho\, dx)$ is the closure of $C_0^\infty(\Omega)$ with respect to the norm (4) then $H(\Omega, \rho\, dx) \ne W(\Omega, \rho\, dx)$. In literature this fact is called the Lavrentieff phenomenon and it leads to surprising consequences like non-uniqueness of solutions to problem (2)–(3) (see [15, 17, 19]) and, hence, to several possible alternative settings of OCPs, depending on the choice of solution space.

OCPs in coefficients have been widely studied by many authors since this topic includes optimal shape design problems, optimization of certain evolution systems, some problems originating in mechanics and others. We could mention Buttazzo & Dal Maso [1], Lions [12], Murat [14] and others.

As François Murat showed for OCPs in coefficients for elliptic equations, even if the weight function ρ is non-degenerate, such problems have no solution, in general. So we have to restrict problem (1)–(3) by introducing some additional control constraints (see, for instance, [9]). In [5] the existence of H-optimal solutions for problem (1)–(3) in the class of so-called generalized solenoidal controls was proved.

The paper is organized as follows. Section 1 contains some notation and preliminaries. In Section 2 we impose additional control constrains like **div**-conditions of a certain type. After that we discuss the classification of admissible solutions to problem (1)–(3). In particular, we define the class of W-admissible solutions and the class of so-called H-admissible solutions. However, we restrict our analysis with the later one. The aim of Section 3 is to give the collection of preliminary results.

In Section 4, we deal with attainability of H-optimal solutions via the optimal solutions to the special perturbed problems for non-degenerate variational inequalities. In applications a degenerate weight ρ occurs as the limit of a sequence of non-degenerate weights ρ_ε for which the corresponding 'approximate' OCP is solvable (see [10]). The results of this section answer the following question: if limit points of the family of admissible solutions $(\mathcal{U}_\varepsilon, y_\varepsilon)$ to the perturbed problems appear to be H-admissible solutions to the original problem (1)–(3), whether all H-optimal solutions are attainable in this sense? Note that for the above OCP the attainability and approximability questions remain in the focus of attention. In particular, similar questions were raised by Zhikov and Pastukhova in [19], [15] for the degenerate boundary value problems without controls.

1. Notation and Preliminaries

Weighted Sobolev spaces. For any subset $E \subset \Omega$ we denote by $|E|$ its N-dimensional Lebesgue measure $\mathcal{L}^N(E)$. The space $W_0^{1,1}(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in the classical Sobolev space $W^{1,1}(\Omega)$. Let ρ be a degenerate weight in the sense of (5). For a given $\Omega \subset \mathbb{R}^N$ we associate to this function two weighted Sobolev spaces $W = W(\Omega, \rho\, dx)$ and $H = H(\Omega, \rho\, dx)$, where H is the closure of $C_0^\infty(\Omega)$ in W.

Note that the spaces W and H are reflexive Banach spaces with respect to the norm $\|\cdot\|_\rho$ due to the estimate

$$\int_\Omega |\nabla y|\, dx \leq \left(\int_\Omega \rho |\nabla y|_p^p\, dx\right)^{1/p} \left(\int_\Omega \rho^{-1/(p-1)}\, dx\right)^{p/p-1} \leq C\|y\|_\rho,$$

where $|\eta|_p = \left(\sum\limits_{k=1}^N |\eta_k|^p\right)^{1/p}$ is a Hölder norm of order p in $\mathbb{R}^N$. It is clear that $H \subseteq W$.

Since the smooth functions are in general not dense in the weight Sobolev space W, it follows that $H \neq W$; that is, for a "typical" degenerate weight ρ the identity $W = H$ is not always valid (for the corresponding examples we refer to [2, 16, 17]). However, if ρ is a non-degenerate weight function, that is, ρ is bounded between two positive constants, then it is easy to verify that $W = H = W_0^{1,p}(\Omega)$. We recall that the dual space of H is $H^* = W^{-1,-p/(p-1)}(\Omega, \rho^{-1/(p-1)}\, dx)$ (for more details see [6]).

Remark 1. *Assume that there exists a value* $\nu \in \left(\frac{N}{p}, +\infty\right) \cap \left[\frac{1}{p-1}, +\infty\right)$ *such that* $\rho^{-\nu} \in L^1(\Omega)$. *Then the following result takes place (see* [6, pp.46]*): relation* $(5)_2$ *implies that*

$$|||y|||_{\rho,\Omega} = \left[\int_\Omega \sum_{i=1}^N \left|\frac{\partial y}{\partial x_i}\right|^p \rho\, dx\right]^{1/p}$$

is a norm of the space $H(\Omega, \rho\, dx)$ *equivalent to* (4) *and the embedding*

$$H(\Omega, \rho\, dx) \hookrightarrow L^p(\Omega)$$

is compact and dense.

Monotone operators. Let α and β be constants such that $0 < \alpha \le \beta < +\infty$. We define $M_p^{\alpha,\beta}(\Omega)$ as a set of all symmetric matrices $\mathcal{U}(x) = \{a_{i\,j}(x)\}_{1\le i,j\le N}$ in $L^\infty(\Omega;\mathbb{R}^{N\times N})$ such that the following conditions of growth, monotonicity, and strong coercivity are fulfilled:

$$|a_{i\,j}(x)| \le \beta \quad \text{a.e. in } \ \Omega\ \forall\, i,j \in \{1,\dots,N\}, \tag{6}$$

$$\big(\mathcal{U}(x)([\zeta^{p-2}]\zeta - [\eta^{p-2}]\eta), \zeta-\eta\big)_{\mathbb{R}^N} \ge 0 \quad \text{a.e. in } \ \Omega\ \forall\, \zeta,\eta \in \mathbb{R}^N, \tag{7}$$

$$\big(\mathcal{U}(x)[\zeta^{p-2}]\zeta,\zeta\big)_{\mathbb{R}^N} = \sum_{i,j=1}^N a_{ij}(x)|\zeta_j|^{p-2}\,\zeta_j\,\zeta_i \ge \alpha\,|\zeta|_p^p \quad \text{a.e. in } \ \Omega. \tag{8}$$

Lemma 1. [5] *For every fixed control* $\mathcal{U} \in M_p^{\alpha,\beta}(\Omega)$*, the operator* $A_{\mathcal{U}} : H \to H^*$ *defined as*

$$\langle A_{\mathcal{U}}(y), v\rangle_H = \sum_{i,j=1}^N \int_\Omega \left(a_{ij}(x)\left|\frac{\partial y}{\partial x_j}\right|^{p-2}\frac{\partial y}{\partial x_j}\right)\frac{\partial v}{\partial x_i}\rho\,dx + \int_\Omega |y|^{p-2}y\,v\,dx,$$

is strictly monotone, coercive and semicontinuous (here by the semicontinuity property we mean that the scalar function $t \to \langle A_{\mathcal{U}}(y+tv), w\rangle_H$ *is continuous for all* y, v, $w \in H$*).*

Elliptic Variational Inequalities. Let V be a Banach space and $K \subset V$ be a closed convex subset. Suppose also that $A : K \to V^*$ is a nonlinear operator and $f \in V^*$ is a given element of the dual space.

Let us consider the following variational problem: to find an element $y \in K$ such that

$$\langle Ay, v-y\rangle_V \ge \langle f, v-y\rangle_V, \quad \forall v \in K. \tag{9}$$

Referring to Lions [13], we make use of the following assumptions.

Hypothesis 1. There exists a reflexive Banach space X such that $X \subset V^*$, the imbedding $X \hookrightarrow V^*$ is continuous, and X is dense in V^*.

Hypothesis 2. There can be found a duality mapping $J : X \to X^*$ such that $\forall\, y \in K$, $\forall\, \varepsilon > 0$ there exists an $y_\varepsilon \in K$ such that $A(y_\varepsilon) \in X$ and

$$y_\varepsilon + \varepsilon J(A(y_\varepsilon)) = y.$$

Theorem 1. [13, Theorem 8.7] *Assume that Hypotheses 1 and 2 hold true*[1]*. Let operator* $A : V \to V^*$ *be monotone, semicontinuous, bounded and satisfy the following assumption: there exists an element* $v_0 \in K$ *such that*

$$\frac{\langle Ay, y-v_0\rangle_V}{\|y\|_V} \to +\infty \ \textit{as}\ \|y\|_V \to \infty,\ y \in K.$$

Then for any solution y *of variational inequality* (9) *the inclusion* $Ay \in X$ *takes place provided* $f \in X$*.*

Smoothing. Throughout the paper ε denotes a small parameter which varies within a strictly decreasing sequence of positive numbers converging to 0. When we write $\varepsilon > 0$, we consider only the elements of this sequence, while writing $\varepsilon \ge 0$, we also consider its limit $\varepsilon = 0$.

[1](see the example for $V = H_0^1(\Omega)$ and $X = L^2(\Omega)$ in [13, Theorem 8.8.])

Definition 1. We say that a weight function ρ with properties (5) is approximated by non-degenerate weight functions $\{\rho^\varepsilon\}_{\varepsilon>0}$ on Ω if:

$$\rho^\varepsilon(x) > 0 \quad \text{a.e. in } \Omega, \quad \rho^\varepsilon + (\rho^\varepsilon)^{-1} \in L^\infty(\Omega), \quad \forall\, \varepsilon > 0, \tag{10}$$

$$\rho^\varepsilon \to \rho, \quad (\rho^\varepsilon)^{-1/(p-1)} \to \rho^{-1/(p-1)} \;\text{ in }\; L^1(\Omega) \;\text{ as }\; \varepsilon \to 0. \tag{11}$$

Remark 2. *The family $\{\rho^\varepsilon\}_{\varepsilon>0}$ satisfying properties* (10)–(11) *is called the non-degenerate perturbation of the weight function ρ.*

Examples of such perturbations can be constructed using the classical smoothing. For instance, let Q be some positive compactly supported function such that $Q \in L^\infty(\mathbb{R}^N)$, $\int_{\mathbb{R}^N} Q(x)\,dx = 1$, and $Q(x) = Q(-x)$. Then, for a given weight function $\rho \in L^1_{loc}(\mathbb{R}^N)$, we can take $\rho^\varepsilon = (\rho)_\varepsilon$, where

$$(\rho)_\varepsilon(x) = \frac{1}{\varepsilon^N}\int_{\mathbb{R}^N} Q\left(\frac{x-z}{\varepsilon}\right)\rho(z)\,dz = \int_{\mathbb{R}^N} Q(z)\rho(x+\varepsilon z)\,dz. \tag{12}$$

In this case, we say that the perturbation $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ of the original degenerate weight function ρ is constructed by the “direct ”smoothing scheme.

Lemma 2 ([15])**.** *If ρ, $\rho^{-1/(p-1)} \in L^1_{loc}(\mathbb{R}^N)$ then the “direct” smoothing $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ possesses properties* (10)–(11).

Weak Compactness Criterion in $L^1(\Omega)$. Throughout the paper we will often use the concepts of the weak and strong convergence in $L^1(\Omega)$. Let $\{a_\varepsilon\}_{\varepsilon>0}$ be a bounded sequence in $L^1(\Omega)$. We recall that $\{a_\varepsilon\}_{\varepsilon>0}$ is called equi-integrable if for any $\delta > 0$ there exists $\tau = \tau(\delta)$ such that $\int_S |a_\varepsilon|\,dx < \delta$ for every $\varepsilon > 0$ and every measurable subset $S \subset \Omega$ of Lebesgue measure $|S| < \tau$. Then the following assertions are equivalent: (i) A sequence $\{a_\varepsilon\}_{\varepsilon>0}$ is weakly compact in $L^1(\Omega)$. (ii) The sequence $\{a_\varepsilon\}_{\varepsilon>0}$ is equi-integrable. (iii) Given $\delta > 0$ there exists $\lambda = \lambda(\delta)$ such that $\sup_{\varepsilon>0}\int_{\{|a_\varepsilon|>\lambda\}} |a_\varepsilon|\,dx < \delta$.

Theorem 2 (Lebesgue's Theorem)**.** *If a bounded sequence $\{a_\varepsilon\}_{\varepsilon>0} \subset L^1(\Omega)$ is equi-integrable and $a_\varepsilon \to a$ almost everywhere on Ω, then $a_\varepsilon \to a$ in $L^1(\Omega)$.*

Radon measures and convergence in variable spaces. By a nonnegative Radon measure on Ω we mean a nonnegative Borel measure which is finite on every compact subset of Ω. The space of all nonnegative Radon measures on Ω will be denoted by $\mathcal{M}_+(\Omega)$. If μ is a nonnegative Radone measure on Ω, we will use $L^r(\Omega, d\mu)$, $1 \le r \le \infty$, to denote the usual Lebesgue space with respect to the measure μ with the corresponding norm

$$\|f\|_{L^r(\Omega,d\mu)} = \left(\int_\Omega |f(x)|^r\,d\mu\right)^{1/r}.$$

Let $\{\mu_\varepsilon\}_{\varepsilon>0}$, μ be Radon measures such that $\mu_\varepsilon \stackrel{*}{\rightharpoonup} \mu$ in $\mathcal{M}_+(\Omega)$; that is,

$$\lim_{\varepsilon\to 0}\int_\Omega \varphi\,d\mu_\varepsilon = \int_\Omega \varphi\,d\mu \quad \forall\,\varphi \in C_0(\mathbb{R}^N), \tag{13}$$

where $C_0(\mathbb{R}^N)$ is the space of all compactly supported continuous functions. A typical example of such measures is $d\mu_\varepsilon = \rho^\varepsilon(x)\,dx$, $d\mu = \rho(x)\,dx$, where

$0 \le \rho^\varepsilon \rightharpoonup \rho$ in $L^1(\Omega)$. Let us recall the definition and main properties of convergence in the variable L^p-space [17].

1. A sequence $\{v_\varepsilon \in L^p(\Omega, d\mu_\varepsilon)\}$ is called bounded if

$$\limsup_{\varepsilon\to 0} \int_\Omega |v_\varepsilon|^p \, d\mu_\varepsilon < +\infty.$$

2. A bounded sequence $\{v_\varepsilon \in L^p(\Omega, d\mu_\varepsilon)\}$ converges weakly to $v \in L^p(\Omega, d\mu)$ if

$$\lim_{\varepsilon\to 0} \int_\Omega v_\varepsilon \varphi \, d\mu_\varepsilon = \int_\Omega v\varphi \, d\mu$$

for any $\varphi \in C_0^\infty(\Omega)$ and we write $v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, d\mu_\varepsilon)$.

3. The strong convergence $v_\varepsilon \to v$ in $L^p(\Omega, d\mu_\varepsilon)$ means that $v \in L^p(\Omega, d\mu)$ and

$$\lim_{\varepsilon\to 0} \int_\Omega v_\varepsilon z_\varepsilon \, d\mu_\varepsilon = \int_\Omega vz \, d\mu \quad \text{as} \quad z_\varepsilon \rightharpoonup z \quad \text{in} \quad L^q(\Omega, d\mu_\varepsilon). \tag{14}$$

The following convergence properties in variable spaces hold:

(a) *Compactness criterium:* if a sequence is bounded in $L^p(\Omega, d\mu_\varepsilon)$, then this sequence is compact with respect to the weak convergence.

(b) *Property of lower semicontinuity:* if $v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, d\mu_\varepsilon)$, then

$$\liminf_{\varepsilon\to 0} \int_\Omega |v_\varepsilon|^p \, d\mu_\varepsilon \ge \int_\Omega v^p \, d\mu. \tag{15}$$

(c) *Criterium of strong convergence:* $v_\varepsilon \to v$ if and only if $v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, d\mu_\varepsilon)$ and

$$\lim_{\varepsilon\to 0} \int_\Omega |v_\varepsilon|^p \, d\mu_\varepsilon = \int_\Omega v^p \, d\mu. \tag{16}$$

Concluding this section, we recall some well-known results concerning the convergence in the variable space $L^p(\Omega, \rho^\varepsilon dx)$.

Lemma 3 ([15, 19, 17])**.** *If $\{\rho^\varepsilon\}_{\varepsilon>0}$ is a non-degenerate perturbation of the weight function $\rho(x) \ge 0$, then: (A_1) $(\rho^\varepsilon)^{-1} \to \rho^{-1}$ in $L^q(\Omega, \rho^\varepsilon dx)$. (A_2) $[\, v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, \rho^\varepsilon dx)\,] \Longrightarrow [\, v_\varepsilon \rightharpoonup v$ in $L^1(\Omega)\,]$. (A_3) If a sequence $\{v_\varepsilon \in L^p(\Omega, \rho^\varepsilon dx)\}_{\varepsilon>0}$ is bounded, then the weak convergence $v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, \rho^\varepsilon dx)$ is equivalent to the weak convergence $\rho^\varepsilon v_\varepsilon \rightharpoonup \rho v$ in $L^1(\Omega)$. (A_4) If $a \in L^\infty(\Omega)$ and $v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, \rho^\varepsilon dx)$, then $av_\varepsilon \rightharpoonup av$ in $L^p(\Omega, \rho^\varepsilon dx)$.*

Variable Sobolev spaces. Let $\rho(x)$ be a degenerate weight function and let $\{\rho^\varepsilon\}_{\varepsilon>0}$ be a non-degenerate perturbation of the function ρ in the sense of Definition 1. We denote by $H(\Omega, \rho^\varepsilon \, dx)$ the closure of $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{\rho^\varepsilon}$. Since for every ε the function ρ^ε is non-degenerate, the space $H(\Omega, \rho^\varepsilon \, dx)$ coincides with the classical Sobolev space $W_0^{1,p}(\Omega)$.

Definition 2. We say that a sequence $\{y_\varepsilon \in H(\Omega, \rho^\varepsilon dx)\}_{\varepsilon>0}$ converges weakly to an element $y \in W$ as $\varepsilon \to 0$, if the following hold: (i) This sequence is bounded. (ii) $y_\varepsilon \rightharpoonup y$ in $L^p(\Omega)$. (iii) $\nabla y_\varepsilon \rightharpoonup \nabla y$ in $L^p(\Omega, \rho^\varepsilon dx)^N$.

Compensated Compactness Lemma in variable Lebesgue and Sobolev spaces. Let $\{\rho^\varepsilon\}_{\varepsilon>0}$ be a non-degenerate perturbation of a weight function ρ. We associate to every ρ^ε the space

$$X(\Omega,\rho^\varepsilon dx)=\left\{\vec{f}\in L^q(\Omega,\rho^\varepsilon dx)^N \mid \operatorname{div}\left(\rho^\varepsilon\vec{f}\right)\in L^q(\Omega)\right\}\quad \forall\,\varepsilon>0 \qquad (17)$$

with the norm

$$\|\vec{f}\|_{X(\Omega,\rho^\varepsilon dx)}=\left(\|\vec{f}\|^q_{L^q(\Omega,\rho^\varepsilon dx)^N}+\|\operatorname{div}\left(\rho^\varepsilon\vec{f}\right)\|^q_{L^q(\Omega)}\right)^{1/q}.$$

We say that a sequence $\left\{\vec{f}_\varepsilon\in X(\Omega,\rho^\varepsilon dx)\right\}_{\varepsilon>0}$ is bounded if

$$\limsup_{\varepsilon\to 0}\|\vec{f}_\varepsilon\|_{X(\Omega,\rho^\varepsilon dx)}<+\infty.$$

In order to discuss the attainability of H-optimal solutions to the problem (1)-(3), (19), we use the following result (for comparison, we refer the reader to the Compensated Compactness Lemma in [14]).

Lemma 4. [5] *Let $\{\rho^\varepsilon\}_{\varepsilon>0}$ be a non-degenerate perturbation of a weight function $\rho(x)>0$. Let sequences $\left\{\vec{f}_\varepsilon\in L^q(\Omega,\rho^\varepsilon dx)^N\right\}_{\varepsilon>0}$ and $\{g_\varepsilon\in H(\Omega,\rho^\varepsilon dx)\}_{\varepsilon>0}$ be such that $\left\{\vec{f}_\varepsilon\right\}_{\varepsilon>0}$ is bounded in the variable space $X(\Omega,\rho^\varepsilon dx)$, $\vec{f}_\varepsilon\rightharpoonup\vec{f}$ in $L^q(\Omega,\rho^\varepsilon dx)^N$, $\{g_\varepsilon\}_{\varepsilon>0}$ is bounded in the variable space $H(\Omega,\rho^\varepsilon dx)$, $g_\varepsilon\rightharpoonup g$ in $L^p(\Omega)$, and $\nabla g_\varepsilon\rightharpoonup\nabla g$ in $L^p(\Omega,\rho^\varepsilon dx)^N$. Then*

$$\lim_{\varepsilon\to 0}\int_\Omega\varphi\left(\vec{f}_\varepsilon,\nabla g_\varepsilon\right)_{\mathbb{R}^N}\rho^\varepsilon dx=\int_\Omega\varphi\left(\vec{f},\nabla g\right)_{\mathbb{R}^N}\rho\,dx,\quad \forall\,\varphi\in C_0^\infty(\Omega). \qquad (18)$$

2. Setting of the optimal control problem

The OCP, we consider in this paper, is to minimize the discrepancy between a given distribution $z_\partial\in L^p(\Omega)$ and the solution $y=y_{\mathcal{U},f}$ of the degenerate variational inequality by choosing an appropriate matrix $\mathcal{U}\in L^\infty(\Omega;\mathbb{R}^{N\times N})$. In fact, we deal with the minimization problem in the form (1)–(3).

Definition 3. We say that a matrix $\mathcal{U}=[a_{i\,j}]$ is an admissible control to degenerate problem (2)–(3) if $\mathcal{U}\in U_{ad}$, where the set U_{ad} is defined as follows

$$U_{ad}=\Big\{\mathcal{U}=[\vec{a}_1,\dots,\vec{a}_N]\in M_p^{\alpha,\beta}(\Omega)\Big| \\ |\operatorname{div}\,(\rho\,\vec{a}_i)|\le\gamma_i,\ \text{a.e. in}\,\Omega,\ \forall\,i=1,\dots,N\Big\}. \qquad (19)$$

Here, $\gamma=(\gamma_1,\dots,\gamma_N)\in\mathbb{R}^N$ is a strictly positive vector.

In what follows, depending on the choice of solution space, we introduce the main types of solutions to the above elliptic variational inequality.

Definition 4. We say that a function $y=y(\mathcal{U},f)\in K$ is a W-solution to degenerate variational inequality (2)–(3) if

$$\langle -\operatorname{div}\left(\mathcal{U}(x)\rho(x)\,[(\nabla y)^{p-2}]\nabla y\right)+|y|^{p-2}y,v-y\rangle_W\ge\langle f,v-y\rangle_W, \qquad (20)$$

holds for any $v\in K$.

Definition 5. Let $\widetilde{K}$ be a closure in the space $C_0^\infty(\Omega)$ of the set $K \bigcap C_0^\infty(\Omega)$. We say that a function $y = y(\mathcal{U}, f) \in \widetilde{K}$ is an H-solution to variational inequality (2)–(3) if

$$\langle -\mathrm{div}\left(\mathcal{U}(x)\rho(x)\,[(\nabla y)^{p-2}]\nabla y\right) + |y|^{p-2}y, v - y\rangle_H \geq \langle f, v - y\rangle_H, \tag{21}$$

holds for any $v \in \widetilde{K}$.

Remark 3. *It is easy to see that the set $\widetilde{K} \subset H$ is closed and convex.*

Proposition 1. [5] *For every control $\mathcal{U} \in M_p^{\alpha,\beta}(\Omega)$ and every $f \in L^q(\Omega)$ there exists a unique H-solution to degenerate elliptic variational inequality* (2)–(3).

Remark 4. *Similar result with Proposition 1 concerning existence and uniqueness of W-solution to problem* (2)–(3) *can be easily obtained using similar argumentation.*

Taking this fact into account we can introduce two sets of admissible pairs to the optimal control problem (1)–(3), (19):

$$\Xi_W = \{(\mathcal{U}, y) \in U_{ad} \times W \;\mid y \in K, (\mathcal{U}, y) \text{ are related by (20)}\}, \tag{22}$$

$$\Xi_H = \left\{(\mathcal{U}, y) \in U_{ad} \times H \;\middle|\; y \in \widetilde{K}, (\mathcal{U}, y) \text{ are related by (21)}\right\}. \tag{23}$$

Hence for the given control object described by relations (2)–(3) with both fixed control constraints ($\mathcal{U} \in U_{ad}$) and fixed cost functional (1), we have two different statements of the original optimal control problem, namely

$$\left\langle \inf_{(\mathcal{U},y)\in\Xi_W} I(\mathcal{U}, y)\right\rangle \text{ and } \left\langle \inf_{(\mathcal{U},y)\in\Xi_H} I(\mathcal{U}, y)\right\rangle.$$

As a matter of fact, there is no comparison between these problems, in general. Indeed, having assumed that $W \neq H$ for a given degenerate weight function $\rho \geq 0$, we can come to the effect which is usually called the Lavrentieff phenomenon. It means that for some $\mathcal{U} \in U_{ad}$ and $f \in L^q(\Omega)$ an H-solution $y_H(\mathcal{U}, f)$ to problem (2)–(3) does not coincide with its W-solution $y_W(\mathcal{U}, f)$ [17]. In this paper we deal with H-solutions to problem (2)–(3).

Remark 5. *In view of proposition 1, the set Ξ_H is always nonempty.*

Taking this observation into account, we adopt the following concept.

Definition 6. We say that a pair $(\mathcal{U}^0, y^0) \in L^\infty(\Omega; \mathbb{R}^{N\times N}) \times H$ is an H-optimal solution to problem (1)–(3), (19) if $(\mathcal{U}^0, y^0) \in \Xi_H$ and $I(\mathcal{U}^0, y^0) = \inf_{(\mathcal{U},y)\in\,\Xi_H} I(\mathcal{U}, y)$.

3. Auxiliary results

The aim of this section is to give a collection of auxiliary results which will be useful in the sequel. To begin with, we cite the well-known result of the classical smoothing theory (see [18]).

Lemma 5. *Let $a \in L^q(\Omega)$ be a given element. Then $(a)_\varepsilon \to a$ in $L^q(\Omega)$ as $\varepsilon \to 0$, where $(a)_\varepsilon \in L^q(\Omega)$ is a classical smoothing of a.*

Further, we consider a special "lifting" operator

$$T_\varepsilon : L^p(\Omega, \rho dx) \to L^p(\Omega, \rho^\varepsilon dx)$$

defined as follows

$$\int_\Omega T_\varepsilon y\, \varphi \rho^\varepsilon dx = \int_\Omega y(\varphi)_\varepsilon \rho\, dx \quad \forall\, \varphi \in C_0^\infty(\Omega),\ \forall\, \varepsilon > 0. \tag{24}$$

Firstly this operator was constructed in [18] for the case of an arbitrary measure. The following result is well-known (for the details see [15, Lemma 7.2]).

Proposition 2. *Let $\rho \in L^1_{loc}(\mathbb{R}^N)$ be a degenerate weight and let $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ be a "direct "smoothing of ρ. Then for every element $y \in L^p(\Omega, \rho\, dx)$ there exists a sequence $\{T_\varepsilon y \in L^p(\Omega, \rho^\varepsilon dx)\}_{\varepsilon>0}$ such that $T_\varepsilon y \to y$ in $L^p(\Omega, \rho^\varepsilon dx)$.*

We need the following specification of Proposition 2.

Lemma 6. *Let $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ be a "direct "smoothing of a degenerate weight function $\rho \in L^1_{loc}(\mathbb{R}^N)$. Also let $v \in L^p(\Omega, \rho\, dx)$ and $\{v_\varepsilon \in L^p(\Omega, \rho\, dx)\}_{\varepsilon>0}$ be such that $v_\varepsilon \rightharpoonup v$ in $L^p(\Omega, \rho\, dx)$ as $\varepsilon \to 0$, and $\|v_\varepsilon\|_{L^p(\Omega,\rho\, dx)} \le \gamma$, $\forall\, \varepsilon > 0$ (here, $\gamma > 0$ is a given constant independent of ε). Then*

$$T_\varepsilon v_\varepsilon \rightharpoonup v \quad \textit{in variable space } \ L^p(\Omega, \rho^\varepsilon dx) \ \textit{as} \ \ \varepsilon \to 0, \tag{25}$$

$$\textit{and} \quad \|T_\varepsilon v_\varepsilon\|_{L^p(\Omega,\rho^\varepsilon dx)} \le \gamma, \tag{26}$$

where the "lifting" operator $T_\varepsilon : L^p(\Omega, \rho\, dx) \to L^p(\Omega, \rho^\varepsilon dx)$ is defined by (24) $\forall\, \varepsilon > 0$.

Proof. By Proposition 2 for every $\varepsilon > 0$ there exists a function

$$T_\varepsilon v_\varepsilon \in L^p(\Omega, \rho^\varepsilon dx)$$

such that (24) holds true. Moreover, by the estimate

$$\begin{aligned}
\int_\Omega v_\varepsilon (\varphi)_\varepsilon \rho\, dx \le C_\varepsilon \left(\int_\Omega |(\varphi)_\varepsilon|^q \rho\, dx \right)^{1/q} &\overset{\text{by Jensen's inequality}}{\le} \\
\le C_\varepsilon \left(\int_\Omega (|\varphi|^q)_\varepsilon \rho\, dx \right)^{1/q} &= \\
= C_\varepsilon \left(\frac{1}{\varepsilon^N} \int_\Omega \int_{\mathbb{R}^N} Q\left(\frac{x-z}{\varepsilon} \right) |\varphi|^q(z) \rho(x)\, dz\, dx \right)^{1/q} &\le \\
\le C_\varepsilon \left(\int_\Omega |\varphi|^q (\rho)_\varepsilon\, dx \right)^{1/q} = C_\varepsilon \left(\int_\Omega |\varphi|^q \rho^\varepsilon dx \right)^{1/q}&,
\end{aligned}$$

where $C_\varepsilon^p = \int_\Omega v_\varepsilon^p \rho\, dx$ and $\varphi \in C_0^\infty(\Omega)$, we obtain, that

$$\langle T_\varepsilon v_\varepsilon, \varphi \rangle_{L^q(\Omega,\rho^\varepsilon dx)} \le \|v_\varepsilon\|_{L^p(\Omega,\rho dx)} \|\varphi\|_{L^q(\Omega,\rho^\varepsilon dx)}$$

for all $\varphi \in C_0^\infty(\Omega)$. Then, obviously

$$\|T_\varepsilon v_\varepsilon\|_{L^p(\Omega,\rho^\varepsilon dx)} \le \|v_\varepsilon\|_{L^p(\Omega,\rho dx)} \le \gamma.$$

Hence, the sequence $\{T_\varepsilon v_\varepsilon\}_{\varepsilon>0}$ is bounded in $L^p(\Omega,\rho^\varepsilon dx)$. Let us show that $T_\varepsilon v_\varepsilon \rightharpoonup v$ in $L^p(\Omega,\rho^\varepsilon dx)$ as $\varepsilon\to 0$. Indeed, by the initial assumption, we have the estimate

$$\left|\int_\Omega v_\varepsilon(\varphi)_\varepsilon\rho\,dx-\int_\Omega v\varphi\rho\,dx\right|\le\int_\Omega |v_\varepsilon|\,|(\varphi)_\varepsilon-\varphi|\,\rho\,dx+ \\ +\left|\int_\Omega v_\varepsilon\varphi\rho\,dx-\int_\Omega v\varphi\rho\,dx\right|=I_1^\varepsilon+I_2^\varepsilon.$$

Since $v_\varepsilon\rightharpoonup v$ in $L^p(\Omega,\rho\,dx)$, it follows that $\lim_{\varepsilon\to0}I_2^\varepsilon=0$. Further, using the classical properties of smoothing, we conclude that $(\varphi)_\varepsilon\to\varphi$ uniformly on Ω. Since

$$I_1^\varepsilon\le\sup_{x\in\Omega}|(\varphi)_\varepsilon-\varphi|\left(\int_\Omega\rho(x)\,dx\right)^{1/q}\left(\int_\Omega|v|_\varepsilon^p\rho(x)\,dx\right)^{1/p}\le C_1\sup_{x\in\Omega}|(\varphi)_\varepsilon-\varphi|\,,$$

it follows that $\lim_{\varepsilon\to0}I_1^\varepsilon=0$. Thus $v_\varepsilon(\varphi)_\varepsilon\rightharpoonup v\varphi$ in $L^p(\Omega,\rho\,dx)$. Hence, the weak convergence (25) is a direct consequence of the relation $\int_\Omega T_\varepsilon v_\varepsilon\varphi\rho^\varepsilon dx=$ $=\int_\Omega v_\varepsilon(\varphi)_\varepsilon\rho\,dx\ \forall\,\varphi\in C_0^\infty(\Omega)$. The proof is complete. □

Using the same arguments and Lemma 5, we come to the following well-known result.

Lemma 7. *Let $v\in L^q(\Omega)$ and $\{v_\varepsilon\in L^q(\Omega)\}_{\varepsilon>0}$ be such that $v_\varepsilon\rightharpoonup v$ in $L^q(\Omega)$ as $\varepsilon\to0$, and $\|v_\varepsilon\|_{L^q(\Omega)}\le\gamma$, $\forall\,\varepsilon>0$ (here, $\gamma>0$ is a given constant independent of ε). Then*

$$(v_\varepsilon)_\varepsilon\rightharpoonup v\quad\text{in}\quad L^q(\Omega)\quad\text{as}\quad\varepsilon\to0,\quad\text{and}\ \|(v_\varepsilon)_\varepsilon\|_{L^q(\Omega)}\le\gamma\quad\forall\,\varepsilon>0.\tag{27}$$

Let us recall, that an element $a\in L^q(\Omega)$ and a vector $b\in L^q(\Omega,\rho\,dx)^N$ are related by the equality $\operatorname{div}(\rho\,b)=a$ if

$$\int_\Omega(b,\nabla\varphi)_{\mathbb{R}^N}\rho\,dx=-\int_\Omega a\varphi\,dx\ \text{for all}\ \varphi\in C_0^\infty(\Omega).\tag{28}$$

In a similar way, for $a^\varepsilon\in L^q(\Omega)$ and $b^\varepsilon\in L^q(\Omega,\rho^\varepsilon\,dx)^N$, we have $\operatorname{div}\,(\rho^\varepsilon b^\varepsilon)=a^\varepsilon$ if

$$\int_\Omega(b^\varepsilon,\nabla\varphi)_{\mathbb{R}^N}\rho^\varepsilon\,dx=-\int_\Omega a^\varepsilon\varphi\,dx\ \text{for all}\ \varphi\in C_0^\infty(\Omega).\tag{29}$$

Note that by arguments of completion, the above identities can be extended to test functions from H and $H(\Omega,\rho^\varepsilon\,dx)$, respectively.

As an obvious consequence of Lemmas 6 and 7, we have the following results.

Lemma 8. *Let $\{\rho^\varepsilon=(\rho)_\varepsilon\}_{\varepsilon>0}$ be a "direct" smoothing of a degenerate weight function $\rho\in L^1_{loc}(\mathbb{R}^N)$, and let $\left\{\vec{f}_\varepsilon\right\}_{\varepsilon>0}$ be a sequence of vector-functions in $L^q(\Omega,\rho\,dx)^N$ such that $\vec{f}_\varepsilon\rightharpoonup\vec{f}$ in $L^q(\Omega,\rho\,dx)^N$, $\operatorname{div}\left(\rho\,\vec{f}_\varepsilon\right)\rightharpoonup\xi$ in $L^q(\Omega)$ as*

$\varepsilon \to 0$. *Then*

$$\xi = \operatorname{div}\left(\rho \vec{f}\right); \text{ that is, } \quad \int_\Omega \left(\vec{f}, \nabla\varphi\right)_{\mathbb{R}^N} \rho\, dx = -\int_\Omega \xi\varphi\, dx, \tag{30}$$

$$\operatorname{div}\left(\rho^\varepsilon T_\varepsilon \vec{f_\varepsilon}\right) = \left(\operatorname{div}(\rho \vec{f_\varepsilon})\right)_\varepsilon \in L^q(\Omega) \quad \forall\, \varepsilon > 0, \tag{31}$$

$$\left(\operatorname{div}(\rho \vec{f_\varepsilon})\right)_\varepsilon \rightharpoonup \operatorname{div}(\rho \vec{f}) \quad \text{in } L^q(\Omega) \text{ as } \varepsilon \to 0. \tag{32}$$

Lemma 9. *If $a \in L^q(\Omega)$ and $b \in L^q(\Omega, \rho\, dx)^N$ are related by (28), then $a^\varepsilon = (a)_\varepsilon$ and $b^\varepsilon = T_\varepsilon b$ are related by the equality (29).*

Following [15], we can give a dual description of the weighted Sobolev space H. To this end, let us consider two spaces: The first is X^p_ρ as the closure of the set $\{(y, \nabla y),\ y \in C^\infty_0(\Omega)\}$ in $L^p(\Omega) \times L^p(\Omega, \rho\, dx)^N$, hence, the elements of this space are pairs (y, v), where y is a function in H and $v = \nabla y$ is its gradient. The second space $\widetilde{X}^p_\rho$ consists of pairs (y, v), where $y \in L^p(\Omega)$ and $v \in L^p(\Omega, \rho\, dx)^N$ are such that

$$\int_\Omega ya\, dx = -\int_\Omega (v, b)_{\mathbb{R}^N} \rho\, dx \tag{33}$$

for any (a, b) satisfying the conditions

$$a \in L^q(\Omega), \quad b \in L^q(\Omega, \rho\, dx)^N, \quad a = \operatorname{div}(\rho\, b). \tag{34}$$

It is easy to see, that X^p_ρ and $\widetilde{X}^p_\rho$ are closed in $L^p(\Omega) \times L^p(\Omega, \rho\, dx)^N$ and $X^p_\rho \subseteq \widetilde{X}^p_\rho$.

Lemma 10. $X^p_\rho = \widetilde{X}^p_\rho$.

Proof. Assume that $X^p_\rho \neq \widetilde{X}^p_\rho$, that is there exists at least one pair $(y_0, v_0) \in \widetilde{X}^p_\rho \setminus X^p_\rho$. Then there exists a couple of elements $(y^0, v^0) \in L^q(\Omega)\times$ $\times L^q(\Omega, \rho\, dx)^N$ such that

$$\int_\Omega y^0\varphi\, dx + \int_\Omega (v^0, \nabla\varphi)_{\mathbb{R}^N} \rho\, dx = 0 \text{ for all } \varphi \in C^\infty_0(\Omega),$$
$$\int_\Omega y^0 y_0\, dx + \int_\Omega (v^0, v_0)_{\mathbb{R}^N} \rho\, dx = 1. \tag{35}$$

However, the first equality means that the pair (y^0, v^0) satisfies the conditions (34), and, therefore, it can be taken as a test pair in relation (33) for (y_0, v_0). As a result, we obtain $\int_\Omega y^0 y_0\, dx = -\int_\Omega (v^0, v_0)_{\mathbb{R}^N} \rho\, dx$ and come to the contradiction with (35). $\square$

Now we can prove the main result of this section.

Theorem 3. *Let $\rho^\varepsilon = (\rho)_\varepsilon$ be a direct smoothing of a degenerate weight $\rho \in L^1_{loc}(\mathbb{R}^N)$ and let $y^\varepsilon \in H(\Omega, \rho^\varepsilon\, dx)$, $y^\varepsilon \rightharpoonup y$ in $L^p(\Omega)$, $\nabla y^\varepsilon \rightharpoonup v$ in $L^p(\Omega, \rho^\varepsilon\, dx)^N$. Then $y \in H$ and $v = \nabla y$.*

Proof. Let $a \in L^q(\Omega)$ and $b \in L^q(\Omega, \rho\, dx)^N$ be such that $a = \operatorname{div}(\rho\, b)$. By Lemma 9 we can construct the elements $a^\varepsilon \in L^q(\Omega)$ and $b^\varepsilon \in L^q(\Omega, \rho^\varepsilon\, dx)^N$ such that $a^\varepsilon = \operatorname{div}(\rho^\varepsilon b^\varepsilon)$; that is, $\int_\Omega a^\varepsilon y^\varepsilon\, dx = -\int_\Omega (b^\varepsilon, \nabla y^\varepsilon)_{\mathbb{R}^N} \rho^\varepsilon\, dx$. Passing to the limit in this relation as $\varepsilon \to 0$ and using the strong convergence $a^\varepsilon = (a)_\varepsilon \to a$ in $L^q(\Omega)$ and $b^\varepsilon = T_\varepsilon b \to b$ in $L^q(\Omega, \rho^\varepsilon\, dx)$ (see Proposition 2 and Lemma 5), we arrive to $\int_\Omega ay\, dx = -\int_\Omega (b, \nabla v)_{\mathbb{R}^N} \rho\, dx$. Hence, $(y, v) \in \widetilde{X}^p_\rho = X^p_\rho$, and, therefore, $y \in H$ and $v = \nabla y$. The proof is complete. $\square$

4. Attainability of H-optimal solutions

The aim of this section is to propose an appropriate non-degenerate perturbation for the original degenerate OCP (1)–(3), (19) and to show that H-optimal solutions of (1)–(3), (19) can be attained by optimal solutions of perturbed problems. Hereinafter in this section we assume that the set of H-optimal solutions to the problem (1)–(3), (19) is non-empty.

Let ρ be a degenerate weight function with properties (5), and let $\{\rho^\varepsilon\}_{\varepsilon>0}$ be a non-degenerate perturbation of ρ in the sense of Definition 1.

Definition 7. We say that a bounded sequence

$$\left\{(\mathcal{U}_\varepsilon, y_\varepsilon) \in \mathbb{Y}(\Omega, \rho^\varepsilon dx) = L^\infty(\Omega; \mathbb{R}^{N\times N}) \times H(\Omega, \rho^\varepsilon dx)\right\}_{\varepsilon>0}$$

w-converges to $(\mathcal{U}, y) \in L^\infty(\Omega; \mathbb{R}^{N\times N}) \times W$ in the variable space $\mathbb{Y}(\Omega, \rho^\varepsilon dx)$ as $\varepsilon \to 0$ (in symbols, $(\mathcal{U}_\varepsilon, y_\varepsilon) \overset{w}{\rightharpoonup} (\mathcal{U}, y)$), if $\mathcal{U}_\varepsilon \overset{*}{\rightharpoonup} \mathcal{U}$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$, $y_\varepsilon \rightharpoonup y$ in $L^p(\Omega)$, and $\nabla y_\varepsilon \rightharpoonup \nabla y$ in $L^p(\Omega, \rho^\varepsilon dx)^N$.

Definition 8. We say that a minimization problem

$$\left\langle \inf_{(\mathcal{U},y)\in \Xi_H} I(\mathcal{U}, y) \right\rangle \tag{36}$$

is a weak variational limit (or variational w-limit) of the sequence

$$\left\{ \left\langle \inf_{(\mathcal{U}_\varepsilon,y_\varepsilon)\,\in \Xi_\varepsilon} I_\varepsilon(\mathcal{U}_\varepsilon, y_\varepsilon) \right\rangle; \quad \Xi_\varepsilon \subset \mathbb{Y}(\Omega, \rho^\varepsilon dx), \quad \varepsilon > 0 \right\}, \tag{37}$$

with respect to w-convergence in the variable space $\mathbb{Y}(\Omega, \rho^\varepsilon dx)$, if the following conditions are satisfied:

(1) if $\{\varepsilon_k\}$ is a subsequence of $\{\varepsilon\}$ such that $\varepsilon_k \to 0$ as $k \to \infty$, and a sequence $\{(\mathcal{U}_k, y_k) \in \Xi_{\varepsilon_k}\}_{\varepsilon>0}$ w-converges to a pair $(\mathcal{U}, y)$, then

$$(\mathcal{U}, y) \in \Xi_H; \quad I(\mathcal{U}, y) \le \liminf_{k\to\infty} I_{\varepsilon_k}(\mathcal{U}_k, y_k); \tag{38}$$

(2) for every pair $(\mathcal{U}, y) \in \Xi_H$ and any value $\delta > 0$ there exists a realizing sequence $\left\{(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon) \in \mathbb{Y}(\Omega, \rho^\varepsilon dx)\right\}_{\varepsilon>0}$ such that

$$(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon) \in \Xi_\varepsilon \ \forall\, \varepsilon > 0, \quad (\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon) \overset{w}{\to} (\widehat{\mathcal{U}}, \widehat{y}), \tag{39}$$

$$\|\mathcal{U} - \widehat{\mathcal{U}}\|_{L^\infty(\Omega;\mathbb{R}^{N\times N})} + \|y - \widehat{y}\|_\rho \le \delta, \qquad I(\mathcal{U}, y) \ge \limsup_{\varepsilon\to 0} I_\varepsilon(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon) - \delta. \tag{40}$$

Definition 8 is motivated by the following property of variational w-limits (for the details we refer to [4]).

Theorem 4. *Assume that* (36) *is a weak variational limit of the sequence* (37)*, and the constrained minimization problem* (36) *has a solution. Suppose* $\{(\mathcal{U}^0_\varepsilon, y^0_\varepsilon) \in \Xi_\varepsilon\}_{\varepsilon>0}$ *is a sequence of optimal pairs to* (37)*. Then there exists a pair* $(\mathcal{U}^0, y^0) \in \Xi_H$ *such that* $(\mathcal{U}^0_\varepsilon, y^0_\varepsilon) \xrightarrow{w} (\mathcal{U}^0, y^0)$*, and*

$$\inf_{(\mathcal{U},y)\in\,\Xi_H} I(A,y) = I\left(\mathcal{U}^0, y^0\right) = \lim_{\varepsilon\to 0} \inf_{(\mathcal{U}_\varepsilon,y_\varepsilon)\in\,\Xi_\varepsilon} I_\varepsilon(\mathcal{U}_\varepsilon, y_\varepsilon).$$

Let us consider the sequence $\{K_\varepsilon\}_{\varepsilon>0}$ of non-empty closed and convex subsets, which sequentially converges to the set $\widetilde{K}$ in the sense of Kuratovski as $\varepsilon \to 0$ with respect to weak topology of the space $H(\Omega, \rho^\varepsilon dx)$ and let Hypothesis 2 hold true for $X = L^q(\Omega)$ and $V = H(\Omega, \rho^\varepsilon dx)$ $\forall\, \varepsilon > 0$. Taking into account theorem 4, we consider the following collection of perturbed OCPs in coefficients for non-degenerate elliptic variational inequalities:

$$\text{Minimize} \quad \left\{ I_\varepsilon(\mathcal{U}, y) = \int_\Omega |y(x) - z_\partial(x)|^p \, dx \right\}, \tag{41}$$

$$\mathcal{U} \in U^\varepsilon_{ad}, \ y \in K_\varepsilon, \tag{42}$$

$$\langle -\mathrm{div}\left(\rho^\varepsilon \mathcal{U}[(\nabla\, y)^{p-2}]\, \nabla\, y\right) + |y|^{p-2} y, v - y\rangle_{H(\Omega,\rho^\varepsilon dx)} \geq \\ \geq \langle f, v - y\rangle_{H(\Omega,\rho^\varepsilon dx)} \ \forall v \in K_\varepsilon, \tag{43}$$

$$U^\varepsilon_{ad} = \Big\{ \mathcal{U} = [\vec{a}_1, \dots, \vec{a}_N] \in M^{\alpha,\beta}_p(\Omega) \Big| \\ |\mathrm{div}\ (\rho^\varepsilon\, \vec{a}_i)| \leq \gamma_i, \ \text{a.e. in} \ \ \Omega, \ \forall\, i = 1, \dots, N \Big\}, \tag{44}$$

where the elements $z_\partial \in L^p(\Omega)$, $f \in L^q(\Omega)$ and $\gamma = (\gamma_1, \dots, \gamma_N) \in \mathbb{R}^N$ are the same as for the original problem (1)–(3), (19). For every $\varepsilon > 0$ we define Ξ_ε as a set of all admissible pairs to the problem (41)–(44), namely $(\mathcal{U}, y) \in \Xi_\varepsilon$ if and only if the pair $(\mathcal{U}, y)$ satisfies (42)–(44).
Note that each of perturbed OCPs (41)–(44) is solvable provided $\{\rho^\varepsilon\}_{\varepsilon>0}$ is a non-degenerate perturbation of $\rho \geq 0$ (see [10]).

Remark 6. *Let us recall that sequential* K*-upper and* K*-lower limits of a sequence of sets* $\{E_k\}_{k\in\mathbb{N}}$ *are defined as follows, respectively:*

$$K_s\text{–}\overline{\lim}\ E_k = \{y \in X : \exists \sigma(k) \to \infty, \ \exists y_k \to y, \ \forall k \in \mathbb{N} : \ y_k \in E_{\sigma(k)}\},$$

$$K_s\text{–}\underline{\lim}\ E_k = \{y \in X : \exists y_k \to y, \ \exists k \geq k_0 \in \mathbb{N} : \ y_k \in E_k\}.$$

The sequence $\{E_k\}_{k\in\mathbb{N}}$ *sequentially converges in the sense of Kuratovski to the set* E *(shortly,* K_s*-converges), if* $E = K_s\text{–}\underline{\lim}\ E_k = K_s\text{–}\overline{\lim}\ E_k$.

Lemma 11. *Let* $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ *be a "direct" smoothing of a degenerate weight function* $\rho(x) \geq 0$*. Let* $\{(\mathcal{U}_\varepsilon, y_\varepsilon) \in \Xi_\varepsilon\}_{\varepsilon>0}$ *be a sequence of admissible pairs to the problem* (41)–(44)*. Then there exist a pair* $(\mathcal{U}^*, y^*)$ *and a subsequence* $\{(\mathcal{U}_{\varepsilon_k}, y_{\varepsilon_k})\}_{k\in\mathbb{N}}$ *of* $\{(\mathcal{U}_\varepsilon, y_\varepsilon) \in \Xi_\varepsilon\}_{\varepsilon>0}$ *such that* $(\mathcal{U}_{\varepsilon_k}, y_{\varepsilon_k}) \xrightarrow{w} (\mathcal{U}^*, y^*)$ *as* $k \to \infty$ *and* $(\mathcal{U}^*, y^*) \in \Xi_H$.

Proof. As follows from (44) and the variational inequality

$$\langle -\mathrm{div}\left(\rho^\varepsilon \mathcal{U}_\varepsilon[(\nabla\, y_\varepsilon)^{p-2}]\, \nabla y_\varepsilon\right) + \\ + |y_\varepsilon|^{p-2} y_\varepsilon, v_\varepsilon - y_\varepsilon\rangle_{H(\Omega,\rho^\varepsilon dx)} \geq \langle f, v_\varepsilon - y_\varepsilon\rangle_{H(\Omega,\rho^\varepsilon dx)}, \ \forall\, v_\varepsilon \in K_\varepsilon, \tag{45}$$

the sequences $\{\mathcal{U}_\varepsilon\}_{\varepsilon>0}$, $\{y_\varepsilon\}_{\varepsilon>0}$, and $\{\nabla y_\varepsilon\}_{\varepsilon>0}$ are bounded in $L^\infty(\Omega;\mathbb{R}^{N\times N})$, $L^p(\Omega)$, and $L^p(\Omega,\rho^\varepsilon dx)^N$, respectively. Indeed, let us prove boundedness of the sequence $\{y_\varepsilon\}_{\varepsilon>0}$ in the space $H(\Omega,\rho^\varepsilon dx)$ by contradiction. Namely, suppose that $\|y_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)} \to \infty$ as $\varepsilon \to 0$. Then, on the one hand

$$\begin{aligned}
&\langle -\mathrm{div}\left(\rho^\varepsilon \mathcal{U}_\varepsilon(x)[(\nabla\, y_\varepsilon)^{p-2}]\,\nabla y_\varepsilon\right) + |y_\varepsilon|^{p-2}y_\varepsilon, y_\varepsilon - v_\varepsilon\rangle_{H(\Omega,\rho^\varepsilon dx)} \leq \\
&\quad \leq \langle f, y_\varepsilon - v_\varepsilon\rangle_{H(\Omega,\rho^\varepsilon dx)} = \int_\Omega f(y_\varepsilon - v_\varepsilon)\,dx \leq \|f\|_{L^q(\Omega)}\|y_\varepsilon - v_\varepsilon\|_{L^p(\Omega)} \leq \\
&\qquad \leq \|f\|_{L^q(\Omega)}\|y_\varepsilon - v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)},\ \forall\, v_\varepsilon \in K_\varepsilon,\ \forall\,\varepsilon > 0.
\end{aligned} \tag{46}$$

On the other hand, for arbitrary fixed element $v \in K$ let us consider the sequence $\{v_\varepsilon \in K_\varepsilon\}_{\varepsilon>0}$ such that $v_\varepsilon \to v$ weakly in $H(\Omega,\rho^\varepsilon dx)$ (such sequence always exists provided $\widetilde{K} = K_s{-}\lim K_\varepsilon$), and then, using the estimation (see [5, Proposition 3.5])

$$\langle A(\mathcal{U},y), y - v\rangle_H \geq \min\{\alpha,1\}\|y\|_H^p - \max\{\beta,1\}\,\|v\|_H\,\|y\|_H^{p-1},\ v \in H,$$

we obtain the following relations:

$$\begin{aligned}
&\frac{\langle -\mathrm{div}\left(\mathcal{U}_\varepsilon(x)[(\nabla\, y_\varepsilon)^{p-2}]\,\nabla y_\varepsilon\right) + |y_\varepsilon|^{p-2}y_\varepsilon, y_\varepsilon - v_\varepsilon\rangle_{H(\Omega,\rho^\varepsilon dx)}}{\|y_\varepsilon - v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}} \geq \\
&\quad \geq \frac{\langle -\mathrm{div}\left(\mathcal{U}_\varepsilon(x)[(\nabla\, y_\varepsilon)^{p-2}]\,\nabla y_\varepsilon\right) + |y_\varepsilon|^{p-2}y_\varepsilon, y_\varepsilon - v_\varepsilon\rangle_{H(\Omega,\rho^\varepsilon dx)}}{\|y_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)} + \|v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}} \geq \\
&\quad \geq \frac{\|y_\varepsilon\|^p_{H(\Omega,\rho^\varepsilon dx)}\min\{\alpha,1\} - \max\{\beta,1\}\|v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}\|y_\varepsilon\|^{p-1}_{H(\Omega,\rho^\varepsilon dx)}}{\|y_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)} + \|v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}} = \\
&\quad = \|y_\varepsilon\|^{p-1}_{H(\Omega,\rho^\varepsilon dx)} \frac{\left(\min\{\alpha,1\} - \dfrac{\max\{\beta,1\}\|v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}}{\|y_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}}\right)}{\left(1 + \dfrac{\|v_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}}{\|y_\varepsilon\|_{H(\Omega,\rho^\varepsilon dx)}}\right)} \to \infty \text{ as } \varepsilon \to 0,
\end{aligned}$$

since the sequence $\{v_\varepsilon\}_{\varepsilon>0}$ is bounded in $H(\Omega,\rho^\varepsilon dx)$. The obtained contradiction with (46) implies that $\{y_\varepsilon\}_{\varepsilon>0}$ is bounded in $H(\Omega,\rho^\varepsilon dx)$.

Hence, there exists a subsequence $\{\varepsilon_k\}$ of the sequence $\{\varepsilon\}$, converging to 0 and elements $\mathcal{U}^* \in M_p^{\alpha,\beta}(\Omega)$, $y^* \in L^p(\Omega)$, $\vec{v} \in L^p(\Omega,\rho\,dx)^N$, and $\vec{\xi} \in L^q(\Omega,\rho\,dx)^N$ such that $\mathcal{U}_{\varepsilon_k} \overset{*}{\rightharpoonup} \mathcal{U}^*$ in $L^\infty(\Omega;\mathbb{R}^{N\times N})$, $y_{\varepsilon_k} \rightharpoonup y^*$ in $L^p(\Omega)$, $\nabla y_{\varepsilon_k} \rightharpoonup \vec{v}$ in $L^p(\Omega,\rho^{\varepsilon_k}dx)^N$, and

$$\mathcal{U}_{\varepsilon_k}[(\nabla y_{\varepsilon_k})^{p-2}]\nabla y_{\varepsilon_k} := \vec{\xi_{\varepsilon_k}} \rightharpoonup \vec{\xi} \ \text{ in } \ L^q(\Omega,\rho^{\varepsilon_k}dx)^N. \tag{47}$$

By Theorem 3, we have that $y^* \in H$ and $v = \nabla y^*$ and, moreover, we have $y^* \in \widetilde{K}$.

Let us prove that $\mathcal{U}^* \in U_{ad}$. Indeed, since $\{\mathcal{U}_{\varepsilon_k} = [\vec{a}_{1\,\varepsilon_k},\dots,\vec{a}_{N\,\varepsilon_k}]\}_{k\in\mathbb{N}} \subset U_{ad}$, we have that $|\mathrm{div}\ (\rho^{\varepsilon_k}\,\vec{a}_{i\,\varepsilon_k})| \leq \gamma_i$ a. e. in Ω $\forall\, i = 1,\dots,N$. Then passing to the limit as $k \to \infty$ in the relations

$$\int_\Omega (\vec{a}_{i\,\varepsilon_k}, \nabla\varphi)_{\mathbb{R}^N} \rho^{\varepsilon_k}\, dx = -\int_\Omega \varphi \,\mathrm{div}\, (\rho^{\varepsilon_k}\, \vec{a}_{i\,\varepsilon_k})\, dx\, \forall\, \varphi \in C_0^\infty(\Omega),\, \forall\, i = 1, \dots, N,$$

$$-\gamma_i \int_\Omega \varphi\, dx \le \int_\Omega \varphi \,\mathrm{div}\, (\rho^{\varepsilon_k}\, \vec{a}_{i\,\varepsilon_k})\, dx \le \gamma_i \int_\Omega \varphi\, dx, \quad \forall\, i \in \{1, \dots, N\}\,,\ \forall \varphi \ge 0,$$

we may suppose that $|\mathrm{div}\, (\rho\, \vec{a}_i^*)| \le \gamma_i$ a.e. in Ω $\forall\, i \in \{1, \dots, N\}$ and

$$\mathrm{div}\, (\rho^{\varepsilon_k}\, \vec{a}_{i\,\varepsilon_k}) \rightharpoonup \mathrm{div}\, (\rho\, \vec{a}_i^*) \quad \text{in} \;\; L^q(\Omega) \;\; \text{as } k \to \infty. \tag{48}$$

Thus $\mathcal{U}_{\varepsilon_k} \rightharpoonup \mathcal{U}^* = [\vec{a}_1^*, \dots, \vec{a}_N^*]$ weakly-$*$ in $L^\infty(\Omega; \mathbb{R}^{N\times N})$, and $\mathcal{U}^* \in U_{ad}$.

In what follows, we consider the relation (45) for $(\mathcal{U}_{\varepsilon_k}, y_{\varepsilon_k})$ and pass to the limit in it as $k \to \infty$ using the property (14) and the following relations:

$$|y_{\varepsilon_k}|^{p-2} y_{\varepsilon_k} \rightharpoonup |y^*|^{p-2} y^* \quad \text{in } L^q(\Omega) \text{ within a subsequence,} \tag{49}$$

$$\langle -\mathrm{div}\, (\rho^{\varepsilon_k}\, \vec{\xi_{\varepsilon_k}}), y_{\varepsilon_k} \rangle_{H(\Omega, \rho^{\varepsilon_k})} \to \langle -\mathrm{div}\, (\rho\, \vec{\xi}), y^* \rangle_H \tag{50}$$

The latter is valid in view of Lemma 4 and boundedness of the sequence $\{\vec{\xi_{\varepsilon_k}}\}_{k\in\mathbb{N}} \subset X(\Omega, \rho^{\varepsilon_k} dx)$.

Indeed, let us show that $\{\vec{\xi_{\varepsilon_k}}\}_{k\in\mathbb{N}}$ is bounded in $X(\Omega, \rho^{\varepsilon_k} dx)$. To do this we take in Hypothesis 1 $V = H(\Omega, \rho^{\varepsilon_k} dx)$, $X = L^q(\Omega)$, and it is easy to see that the embedding $X \hookrightarrow V^*$ is dense and continuous provided by the fact that the embedding $H(\Omega, \rho^{\varepsilon_k} dx) \hookrightarrow L^p(\Omega)$ is compact and dense (see, for example [8]). Since $f \in L^q(\Omega)$, then, in view of Theorem 1 we have $-\mathrm{div}(\rho \vec{\xi_{\varepsilon_k}}) + |y_{\varepsilon_k}|^{p-2} y_{\varepsilon_k} \in L^q(\Omega)$ $\forall\, k \in \mathbb{N}$, and, obviously, $\mathrm{div}(\rho \vec{\xi_{\varepsilon_k}}) \in L^q(\Omega)$ $\forall\, k \in \mathbb{N}$.

Also, the following relation

$$\int_\Omega \mathrm{div}(\rho^{\varepsilon_k} \vec{\xi_{\varepsilon_k}}) \varphi\, dx = -\int_\Omega (\vec{\xi_{\varepsilon_k}}, \nabla\varphi)_{\mathbb{R}^N}\, \rho^{\varepsilon_k}\, dx \to$$

$$\to -\int_\Omega (\vec{\xi}, \nabla\varphi)_{\mathbb{R}^N}\, \rho\, dx = \int_\Omega \mathrm{div}(\rho \vec{\xi}) \varphi\, dx,\ \forall\, \varphi \in C_0^\infty(\Omega) \text{ as } k \to \infty,$$

means that $\mathrm{div}(\rho^{\varepsilon_k} \vec{\xi_{\varepsilon_k}}) \to \mathrm{div}(\rho \vec{\xi})$ weakly in $L^q(\Omega)$. Therefore the sequence $\{\mathrm{div}(\rho \vec{\xi_{\varepsilon_k}})\}_{k\in\mathbb{N}}$ is bounded in $L^q(\Omega)$ and $\{\vec{\xi_{\varepsilon_k}}\}_{k\in\mathbb{N}}$ is bounded in $X(\Omega, \rho^{\varepsilon_k} dx)$.

Let us prove relation (49). Since $y_{\varepsilon_k} \rightharpoonup y^*$ in $L^p(\Omega)$, $\nabla y_{\varepsilon_k} \rightharpoonup \nabla\, y^*$ in $L^p(\Omega, \rho^{\varepsilon_k} dx)^N$ then from estimates

$$\int_\Omega |y_{\varepsilon_k}|\, dx \le \left(\int_\Omega |y_{\varepsilon_k}|^p dx \right)^{\frac{1}{p}} (|\Omega|)^{\frac{1}{q}} \le C\, (|\Omega|)^{\frac{1}{q}}\,,$$

$$\int_\Omega |\nabla y_{\varepsilon_k}|_p\, dx \le \left(\int_\Omega |\nabla y_{\varepsilon_k}|_p^p \rho^{\varepsilon_k} dx \right)^{\frac{1}{p}} \left(\int_\Omega (\rho^{\varepsilon_k})^{-\frac{1}{(p-1)}}\, dx \right)^{\frac{1}{q}} \le$$

$$\le C \left(\int_\Omega (\rho^{\varepsilon_k})^{-\frac{1}{(p-1)}}\, dx \right)^{\frac{1}{q}},$$

it follows that the sequence $\{y_{\varepsilon_k}\}_{k\in\mathbb{N}}$ is equi-integrable on Ω and bounded in $W^{1,1}(\Omega)$. In view of compact embedding $W^{1,1}(\Omega) \hookrightarrow L^1(\Omega)$, there exists an element $\widetilde{y}$ such that $y_{\varepsilon_k} \to \widetilde{y}$ strongly in $L^1(\Omega)$. However, it is easy to see that $y_{\varepsilon_k} \rightharpoonup y^*$ in $L^1(\Omega)$. Hence $y^* = \widetilde{y}$ a.e. on Ω. It means that up to a subsequence

$y_{\varepsilon_k} \to y^*$ a.e. in Ω and together with boundedness of $\{y_{\varepsilon_k}|^{p-2}y_{\varepsilon_k}\}_{k\in\mathbb{N}}$ in $L^q(\Omega)$ we have $|y_{\varepsilon_k}|^{p-2}y_{\varepsilon_k} \rightharpoonup |y^*|^{p-2}y^*$ in $L^q(\Omega)$ (within a subsequence). Therefore, as a result of limit passage in (45), taking into account (49) and (50), we obtain

$$\langle -\mathrm{div}\,(\rho\,\vec{\xi}), v - y^*\rangle_H + \langle |y^*|^{p-2}y^*, v\rangle_H - \\ - \varlimsup_{k\to\infty} \langle |y_{\varepsilon_k}|^{p-2}y_{\varepsilon_k}, y_{\varepsilon_k}\rangle_{H(\Omega,\rho^{\varepsilon_k}dx)} \geq \langle f, v - y^*\rangle_H,\ \forall\, v \in \widetilde{K}. \quad (51)$$

In order to prove the lemma, it is left to show that $\vec{\xi} = \mathcal{U}^*\left[(\nabla\, y^*)^{p-2}\right]\nabla y^*$. However it can be done in a similar manner as we did it proving [5, Theorem 5.1], using Lemma 4.

Now, let us show that

$$\lim_{k\to\infty} \langle |y_{\varepsilon_k}|^{p-2}y_{\varepsilon_k}, y_{\varepsilon_k}\rangle_{H(\Omega,\rho^{\varepsilon_k}dx)} = \lim_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx = \int_\Omega |y^*|^p\,dx.$$

On the one hand, in view of (15), weak convergence, $y_{\varepsilon_k} \to y^*$ in $L^p(\Omega)$ as $k \to \infty$, implies that:

$$\int_\Omega |y^*|^p\,dx \leq \varliminf_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx.$$

On the other hand, from (51), taking into account the representation of the vector-function ξ, we obtain:

$$\varlimsup_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx \leq \langle -\mathrm{div}\left(\mathcal{U}^*(x)\rho(x)[(\nabla y^*)^{p-2}]\nabla y^*\right) - f, v - y^*\rangle_H + \\ + \langle |y^*|^{p-2}y^*, v\rangle_H,\ \forall\, v \in \widetilde{K}.$$

Having put in the last inequality $v = y^*$, we get

$$\varlimsup_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx \leq \int_\Omega |y^*|^p\,dx.$$

Hence, summing up, the chain of inequalities

$$\int_\Omega |y^*|^p\,dx \leq \varliminf_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx \leq \varlimsup_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx \leq \int_\Omega |y^*|^p\,dx$$

turns into equality $\lim_{k\to\infty}\int_\Omega |y_{\varepsilon_k}|^p\,dx = \int_\Omega |y^*|^p\,dx$, which implies, in view of (16), the strong convergence $y_{\varepsilon_k} \to y^*$ in $L^p(\Omega)$ as $k \to \infty$.

Therefore, variational equality (51) can be represented in the form

$$\langle -\mathrm{div}\left(\mathcal{U}^*(x)\rho(x)[(\nabla y^*)^{p-2}]\nabla y^*\right) + \\ + |y^*|^{p-2}y^*, v - y^*\rangle_H \geq \langle f, v - y^*\rangle_H, \forall\, v \in \widetilde{K} \quad (52)$$

Thus, w-limit pair $(\mathcal{U}^*, y^*)$ is admissible to the problem (1)–(3), (19), hence, $(\mathcal{U}^*, y^*) \in \Xi_H$. The proof is complete. ☐

As an evident consequence of this lemma and the lower semicontinuity property of the cost functional (41) with respect to w-convergence in variable space $\mathbb{Y}(\Omega, \rho^\varepsilon dx)$, we have the following conclusion.

Corollary 1. *Let $\{\varepsilon_k\}$ be a subsequence of indices $\{\varepsilon\}$ such that $\varepsilon_k \to 0$ as $k \to \infty$, and let $\{(\mathcal{U}_k, y_k) \in \Xi_{\varepsilon_k}\}_{k\in\mathbb{N}}$ be a sequence of admissible solutions to corresponding perturbed problems* (41)–(44) *such that $(\mathcal{U}_k, y_k) \overset{w}{\rightharpoonup} (\mathcal{U}, y)$. Then properties* (38) *are valid.*

To discuss properties (39)–(40), we give a result which is reciprocal in some sense to Lemma 11.

Lemma 12. *Let $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ be a "direct" smoothing of a degenerate weight function $\rho(x) \geq 0$ and let $(\mathcal{U}, y) \in \Xi_H$ be any admissible pair. Then there exists a realizing sequence $\left\{(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon) \in \mathbb{Y}(\Omega, \rho^\varepsilon dx)\right\}_{\varepsilon>0}$ such that*

$$(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon) \in \Xi_\varepsilon \ \forall\, \varepsilon > 0, \quad \widehat{\mathcal{U}}_\varepsilon \overset{*}{\rightharpoonup} \mathcal{U} \ \ in \ \ L^\infty(\Omega; \mathbb{R}^{N\times N}); \tag{53}$$

$$\operatorname{div}\left(\rho^\varepsilon\, \vec{a}_{i\,\varepsilon}\right) \rightharpoonup \operatorname{div}\,(\rho\, \vec{a}_i) \ \ in \ \ L^q(\Omega) \quad \forall\, i \in \{1, \dots, N\}, \tag{54}$$

$$\widehat{y}_\varepsilon \to y \ strongly\ in\ L^p(\Omega),\ \nabla y_\varepsilon \rightharpoonup \nabla y \ in\ L^p(\Omega, \rho^\varepsilon\, dx)^N. \tag{55}$$

Proof. To begin with, we assume that a given control $\mathcal{U}$ is such that

$$\mathcal{U} = [\vec{a}_1, \dots, \vec{a}_N] \in M_p^{\alpha,\beta}(\Omega) \ \text{ and } \ |\operatorname{div}\,(\rho\, \vec{a}_i)| < \gamma_i, \text{ a.e. in } \Omega,\ \forall\, i = 1, \dots, N. \tag{56}$$

Further, we construct the sequence $\left\{(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon)\right\}_{\varepsilon>0}$ as follows:

$$\widehat{\mathcal{U}}_\varepsilon(x) = (\mathcal{U})_\varepsilon(x) = [(\vec{a}_1)_\varepsilon(x), \dots, (\vec{a}_N)_\varepsilon(x)] = \int_{\mathbb{R}^N} Q(z)\mathcal{U}(x + \varepsilon z)\, dz, \tag{57}$$

$$\widehat{y}_\varepsilon \in H(\Omega, \rho^\varepsilon dx) \text{ is an } H\text{-solution of (43) corresponding to } \mathcal{U} = \widehat{\mathcal{U}}_\varepsilon. \tag{58}$$

Let us show that for every $\varepsilon > 0$ the pair $(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon)$ is admissible to the corresponding OCP (41)–(44). Indeed, as follows from (57) and properties of the kernel Q, we have

$$\begin{aligned}\alpha|\vec{\xi}\,|_p^p = \alpha \int_{\mathbb{R}^N} Q(x)\, dx\, |\vec{\xi}\,|_p^p \leq \int_{\mathbb{R}^N} Q(z)(\mathcal{U}(x + \varepsilon z)[\vec{\xi}^{p-2}]\vec{\xi}, \vec{\xi}\,)_{\mathbb{R}^N}\, dz = \\ = (\widehat{\mathcal{U}}_\varepsilon(x)[\vec{\xi}^{p-2}]\vec{\xi}, \vec{\xi}\,)_{\mathbb{R}^N} \leq \beta \int_{\mathbb{R}^N} Q(x)\, dx\, |\vec{\xi}\,|_p^p = \beta|\vec{\xi}\,|_p^p \quad \text{a.e. in } \Omega.\end{aligned}$$

Hence, $\widehat{\mathcal{U}}_\varepsilon \in M_p^{\alpha,\beta}(\Omega)\ \forall\, \varepsilon > 0$. Let $T_\varepsilon : L^q(\Omega, \rho\, dx) \to L^q(\Omega, \rho^\varepsilon dx)$ is a "lifting "operator, constructed by (24). For every column $(\vec{a}_k)_\varepsilon$ $(k = 1, \dots, N)$ of the matrix $\widehat{\mathcal{U}}_\varepsilon$ and every $\varphi \in C_0^\infty(\Omega)$, we have

$$\begin{aligned}\int_\Omega (\vec{a}_k)_\varepsilon \nabla\varphi\, dx = \int_\Omega \vec{a}_k (\nabla\varphi)_\varepsilon\, dx = \int_\Omega \rho^{-1}\vec{a}_k(\nabla\varphi)_\varepsilon \rho\, dx = \\ = \int_\Omega T_\varepsilon\left(\rho^{-1}\vec{a}_k\right) \nabla\varphi\rho^\varepsilon\, dx.\end{aligned}$$

Then in view of the main property of lifting operators (see Proposition 2), one gets

$$(\vec{a}_k)_\varepsilon = \rho^\varepsilon T_\varepsilon\left(\rho^{-1}\vec{a}_k\right) \quad \text{and} \quad (\vec{a}_k)_\varepsilon\, (\rho^\varepsilon)^{-1} \to \vec{a}_k\rho^{-1} \ \text{ in } \ L^q(\Omega, \rho^\varepsilon\, dx)^N \tag{59}$$

Therefore, by (59) and the definition of the strong convergence in variable space $L^q(\Omega, \rho^\varepsilon dx)^N$, we obtain

$$\lim_{\varepsilon\to 0}\int_\Omega \left((\vec{a}_k)_\varepsilon, \vec{\phi}\right)_{\mathbb{R}^N} dx = \lim_{\varepsilon\to 0}\int_\Omega (\rho^\varepsilon)^{-1}\left((\vec{a}_k)_\varepsilon, \vec{\phi}\right)_{\mathbb{R}^N} \rho^\varepsilon\, dx = \\ = \int_\Omega \rho^{-1}\left(\vec{a}_k, \vec{\phi}\right)_{\mathbb{R}^N} \rho\, dx = \int_\Omega \left(\vec{a}_k, \vec{\phi}\right)_{\mathbb{R}^N} dx \quad \forall\, \vec{\phi}\in C_0^\infty(\Omega)^N,$$

which implies condition $(53)_2$. Thus, for the sequence $\left\{\widehat{\mathcal{U}}_\varepsilon\right\}_{\varepsilon>0}$ it remains to verify conditions (54) and $(53)_1$.

Since functions $(\vec{a}_k)_\varepsilon$ and ρ^ε are smooth, for every $\varepsilon>0$ and $k=1,\dots,N$ we can define an element $\operatorname{div}(\rho^\varepsilon(\vec{a}_k)_\varepsilon)\in L^q(\Omega)$ by the rule

$$\int_\Omega \varphi \operatorname{div}\,(\rho^\varepsilon(\vec{a}_k)_\varepsilon\,)\,dx = -\int_\Omega ((\vec{a}_k)_\varepsilon, \nabla\varphi)_{\mathbb{R}^N}\, \rho^\varepsilon\, dx \quad \forall\,\varphi\in C_0^\infty(\Omega). \tag{60}$$

Taking into account equality (60), initial assumption on the matrix $\mathcal{U}\in U_{ad}$, and the constrains (57), we come to the relations

$$\lim_{\varepsilon\to 0}\int_\Omega \varphi \operatorname{div}\rho^\varepsilon(\vec{a}_k)_\varepsilon\, dx = -\int_\Omega (\vec{a}_k, \nabla\varphi)_{\mathbb{R}^N}\,\rho\, dx = \\ = \int_\Omega \varphi \operatorname{div}\,(\rho\vec{a}_k)\; dx \quad \forall\, k=1,\dots,N, \tag{61}$$

$$-\gamma_k\int_\Omega \varphi\, dx \le \int_\Omega \varphi \operatorname{div}\rho\vec{a}_k\, dx \le \gamma_k\int_\Omega \varphi\, dx \quad \forall\, k=1,\dots,N,\ 0\le\varphi\in C_0^\infty(\Omega).$$

Since the numerical sequence $\left\{\int_\Omega \varphi \operatorname{div}\rho^\varepsilon(\vec{a}_k)_\varepsilon\, dx\right\}_{\varepsilon>0}$ is bounded, from (61) we have

$$-\gamma_k\int_\Omega \varphi dx \le \int_\Omega \varphi \operatorname{div}\rho^\varepsilon(\vec{a}_k)_\varepsilon dx \le \gamma_k\int_\Omega \varphi dx,\ \forall\, k=1,\dots,N,\ 0\le\varphi\in C_0^\infty(\Omega)$$

which implies

$$|\operatorname{div}\rho^\varepsilon\,(\vec{a}_i)_\varepsilon| \le \gamma_i,\ \text{ a.e. in }\ \Omega,\ \forall\, i=1,\dots,N \quad \text{for } \varepsilon>0 \text{ small enough.} \tag{62}$$

Hence, the sequence $\{\operatorname{div}\rho^\varepsilon(\vec{a}_k)_\varepsilon \in L^q(\Omega)\}_{\varepsilon>0}$ is bounded, and therefore condition (54) immediately follows from (61). Note that the limit case to (56), that is, $\mathcal{U}=[\vec{a}_1,\dots,\vec{a}_N]\in M_p^{\alpha,\beta}(\Omega)$ and $|\operatorname{div}\,(\rho\,\vec{a}_i)|\le\gamma_i$, a. e. in Ω, $\forall\, i=1,\dots,N$, can be considered via the closure procedure of the previous case.

Thus, in view of (62), we conclude that the sequence $\left\{\left(\widehat{\mathcal{U}}_\varepsilon, \widehat{y}_\varepsilon\right)\right\}_{\varepsilon>0}$ is an admissible sequence for the problem (41)–(44). As a result, following arguments of the proof of Lemma 11, we have that $\widehat{y}_\varepsilon\to y$ strongly in $L^p(\Omega)$, $\nabla\widehat{y}_\varepsilon\rightharpoonup\nabla y$ in $L^p(\Omega,\rho^\varepsilon dx)^N$, and $\widehat{\mathcal{U}}_\varepsilon[(\nabla\widehat{y}_\varepsilon)^{p-2}]\nabla\widehat{y}_\varepsilon \rightharpoonup \mathcal{U}[(\nabla y)^{p-2}]\nabla y$ in $L^q(\Omega,\rho^\varepsilon dx)^N$, where $y=y(\mathcal{U})$, for any subsequence of $\{\widehat{y}_\varepsilon\in H(\Omega,\rho^\varepsilon dx)\}_{\varepsilon>0}$ and, hence, for the entire sequence. Here $(\mathcal{U},y)\in\Xi_H$ is a given H-admissible solution to problem (1)–(3), (19).

This concludes the proof. □

Corollary 2. *Lemma 12 implies the equality* $I(\mathcal{U},y)=\lim_{\varepsilon\to 0} I_\varepsilon(\widehat{\mathcal{U}}_\varepsilon,\widehat{y}_\varepsilon)$.

As an obvious consequence of Definition 8, and Lemmas 11–12 with their Corollaries, we can give the following conclusion.

Theorem 5. *Let* $\{\rho^\varepsilon = (\rho)_\varepsilon\}_{\varepsilon>0}$ *be a "direct "smoothing of a degenerate weight function* $\rho(x) > 0$. *Then the minimization problem* (1)–(3), (19) *is a weak variational limit of the sequence* (41)–(44) *as* $\varepsilon \to 0$ *with respect to the* w*-convergence in the variable space* $\mathbb{Y}(\Omega, \rho^\varepsilon dx)$.

As follows from results given above, by Lemma 12 each optimal solution to the problem (1)–(3), (19) can be attained by admissible solutions to perturbed problems (41)–(44), however there exists at least one optimal solution $(\mathcal{U}_0, y_0) \in \Xi_H$ which can be attained by optimal solutions to perturbed problems (41)–(44).

Bybliography

1. Buttazzo G. On the relaxed formulation of some shape optimization problems. / G. Buttazzo, G. Dal Maso, A. Garroni, A. Malusa // Adv. Math. Sci. Appl. — 1997. — **1**, № 7. — P. 1–24.
2. Chiadó Piat V. Some remarks about the density of smooth functions in weighted Sobolev spaces. / V. Chiadó Piat , F. Serra Cassano // J. Convex Analysis. — 1994. — **1**, № 2. — P. 135–142.
3. Gajewski H. Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen. / H. Gajewski, K. Gröger, K. Zacharias — Berlin: Academie-Varlar. — 1974.
4. D'Apice C. Suboptimal boundary control for elliptic equations in critically perforated domains. / C. D'Apice, U. De Maio, P. I. Kogut // Ann. Inst. H. Poincare' Anal. Non Line'aire — 2008. — № 25. — P. 1073–1101.
5. Kupenko O. P. Optimal Control Problems in Coefficients for Degenerate Variational Inequalities of Monotone Type. I. Existence of optimal solutions. / O. P. Kupenko // J. Num. Appl. Math. — 2011, №3(106). — P 88–104.
6. Drabek P. Non linear elliptic equations, singular and degenerate cases. / P. Drabek, A. Kufner, F. Nicolosi — University of West Bohemia. — 1996.
7. Heinonen J. Nonlinear Potential Theory of Degenerate Elliptic Equations. / J. Heinonen, T. Kilpelainen, O. Martio — London: Clarendon Press. — 1993.
8. Ivanenko V. I. Variational Methods in Control Problems for Distributed Systems. / V. I. Ivanenko, V. S. Mel'nik — Kiev: Naukova Dumka. — 1988 (in Russian).
9. Kapustjan V. Ye. Solenoidal controls in coefficients of nonlinear elliptic boundary value problems. / V. Ye. Kapustjan, O. P. Kogut // Computer mathematics (C.M.). — 2010. — **12**, № 1. — P. 138–143 (in Russian).
10. Kogut O. P. On Optimal Control Problem in Coefficients for Nonlinear Elliptic Variational Inequalities. / O. P. Kogut // Visnik DNU., Ser. Problems of Mathematical Modelling and Differential Equations Theory. — 2011. — **3**, № 8. — P. 86–98.
11. Kufner A. Weighted Sobolev Spaces. / A. Kufner — Leipzig:, Band 31, Teubner–Texte Math. — 1980.
12. Lions J.–L. Optimal Control of Systems Governed by Partial Differential Equations. / J.–L. Lions — New York: Springer Verlag. — 1971.
13. Lions J.–L. Some methods of Solving Non–Linear Boundary Value Problems. / J.–L. Lions — Paris:Dunod-Gauthier-Villars. — 1969.

14. Murat F. Compacité par compensation. / F. Murat // Ann. Sc. Norm. Sup. Pisa. — 1978. — № 5. — P. 489–507.
15. Pastukhova S. E. Degenerate equations of monotone type: Lavrent'ev phenomenon and attainability problems. / S. E. Pastukhova // Sbornik: Mathematics. — 2007. — **198**, № 10, — P. 1465—1494.
16. Zhikov V. V. On Lavrentiev phenomenon. / V. V. Zhikov // Russian J. Math. Phys. — 1994. — **3**, № 2. — P. 249–269.
17. Zhikov V. V. Weighted Sobolev spaces. / V. V. Zhikov // Sbornik: Mathematics. — 1998. — **189**, № 8. — P. 27–58.
18. Zhikov V. V. Homogenization of elastitic problems on singular structures. / V. V. Zhikov // Izvestija: Math. — 2002. — **66**, № 2. — P. 299–365.
19. Zhikov V. V. Homogenization of degenerate elliptic equations. / V. V. Zhikov, S. E. Pastukhova // Siberian Math. Journal. — 2006. — **49**, № 1. — P. 80–101.

Кафедра системного аналізу та управління, Вищий навчальний заклад „Національний гірничий університет“, просп. Карла Маркса 19, корп. 7, Дніпропетровськ, 49005, Україна.

Надійшла 16.12.2011

Журнал обчислювальної та прикладної математики | **2012, №1(107)** | Journal of Computational & Applied Mathematics

517.9

ПОШУК НЕРУХОМОЇ ТОЧКИ ЛІПШИЦЕВОЇ НАПІВГРУПИ НЕРОЗТЯГУЮЧИХ ОПЕРАТОРІВ

Ю. В. Маліцький

У роботі запропоновано та обґрунтовано явний алгоритм пошуку нерухомої точки ліпшицевої напівгрупи нерозтягуючих операторів. Дана часткова відповідь на поставлене T. Suzuki питання [T. Suzuki, On strong convergence to common fixed points of nonexpansive semigroups in Hilbert spaces // Proc. Amer. Math. Soc. — 2002. — 131. — P. 2133–2136].

Вступ

Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина. Назвемо оператор $T\colon C \to C$ нерозтягуючим, якщо для всіх $x, y \in C$ виконується нерівність

$$\|Tx - Ty\| \leqslant \|x - y\| .$$

Через $F(T)$ будемо позначати множину нерухомих точок оператора T, тобто $F(T) = \{x \in C : Tx = x\}$. З [1, 2, 3] відомо, що якщо множина $F(T)$ — непорожня, то вона опукла і замкнена.

Теорема 1 (Browder[3]). *Нехай C — замкнена опукла підмножина гільбертового простору H, $T : C \to C$ — нерозтягуючий оператор, $F(T) \neq \emptyset$, (α_n) — послідовність дійсних чисел з проміжку $(0, 1)$, що збігається до 0, y — довільна точка з C. Тоді послідовність (x_n), що визначається як розв'язок рівняння*

$$x_n = \alpha_n y + (1 - \alpha_n)Tx_n,$$

сильно збігається до точки із $F(T)$, що є найближчою до y.

У роботі [4] було запропоновано явний сильно збіжний ітераційний метод пошуку нерухомої точки нерозтягуючого оператора T:

$$y_{n+1} = \alpha_n y + (1 - \alpha_n)Ty_n.$$

Множина $\boldsymbol{T} = \{T(t) : t \in \mathbb{R}^+\}$ — неперервна напівгрупа нерозтягуючих операторів, що діють в замкненій опуклій множині C гільбертового простору H, якщо

(1) для всіх $t \in \mathbb{R}^+$ $\quad T(t)\colon C \to C$ — нерозтягуючий оператор;
(2) $T(0)x = x$ для всіх $x \in C$;
(3) $T(s + t) = T(s) \circ T(t)$ для всіх $s, t \in \mathbb{R}^+$;
(4) для всіх $x \in C$ відображення $T(\cdot)x\colon \mathbb{R}^+ \to C$ — неперервне.

У 2002 році T. Suzuki [5] сформулював теорему про пошук нерухомої точки неперервної напівгрупи нерозтягуючих операторів.

Теорема 2 (Suzuki[5])**.** *Нехай C — замкнена опукла підмножина гільбертового простору H, $\boldsymbol{T} = \{T(t) : t \in \mathbb{R}^+\}$ — неперервна напівгрупа нерозтягуючих операторів, що діють в C, $F(\boldsymbol{T}) = \bigcap_{t\geqslant 0} F(T(t)) \neq \emptyset$, (α_n) і (t_n) — послідовності дійсних чисел таких, що*

$$0 < \alpha_n < 1,\, t_n > 0 \quad i \quad \lim_{n\to\infty} t_n = \lim_{n\to\infty} \frac{\alpha_n}{t_n} = 0,$$

y — довільна точка з C. Тоді послідовність (x_n), що визначається як розв'язок рівняння

$$x_n = \alpha_n y + (1-\alpha_n)T(t_n)x_n,$$

сильно збігається до точки із $F(\boldsymbol{T})$, що є найближчою до y.

Зауваження 1. Якщо в умовах попередньої теореми відомо, що множина C — обмежена, то умову $F(\boldsymbol{T}) \neq \emptyset$ можна забрати.

Алгоритм T. Suzuki є неявним, а отже, незручним для практичних обчислень. В [5] поставлено питання пошуку явного алгоритму. У даній роботі отримана часткова відповідь на це питання. А саме, запропоновано та обгрунтовано явний алгоритм для ліпшицевої напівгрупи нерозтягуючих операторів.

Для подальшої роботи нам знадобиться лема (див., наприклад, [6, 7]).

Лема 1. *Нехай послідовності дійсних чисел (a_n), (b_n) і (c_n) такі, що*

$$0 \leqslant a_{n+1} \leqslant (1-b_n)a_n + c_n,$$
$$a_1 \geqslant 0,\, b_n \in (0,1],\, \sum_{n=1}^{\infty} b_n = +\infty \quad i \quad \overline{\lim}_{n\to\infty} \frac{c_n}{b_n} \leqslant 0.$$

Тоді $\lim_{n\to\infty} a_n = 0$.

1. Явний алгоритм

Розглянемо ліпшицеву напівгрупу нерозтягуючих операторів, тобто в означенні неперервної напівгрупи нерозтягуючих операторів замість умови (4) буде умова

$$(4^*) \quad \exists L > 0 : \quad \forall s, t \geqslant 0 \quad \forall x \in C \quad \|T(s)x - T(t)x\| \leqslant L \cdot |s-t|.$$

Дослідимо асимптотичну поведінку послідовності (u_n), що породжена схемою

$$u_{n+1} = \alpha_n y + (1-\alpha_n)T(t_n)u_n,$$

де $\alpha_n \in (0,1)$, $t_n \in (0,+\infty)$ для всіх $n \in \mathbb{N}$.

Має місце

Теорема 3. *Нехай C — замкнена опукла підмножина гільбертового простору H, $\boldsymbol{T} = \{T(t) : t \in \mathbb{R}^+\}$ — ліпшицева напівгрупа нерозтягуючих операторів, що діють в C, $F(\boldsymbol{T}) \neq \emptyset$, (α_n) і (t_n) — послідовності дійсних чисел таких, що*

$$0 < \alpha_n < 1, t_n > 0, \sum_{n=1}^{\infty} a_n = \infty,$$

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{a_n}{t_n} = \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n-1}}{\alpha_n^2} = \lim_{n\to\infty} \frac{t_n - t_{n-1}}{\alpha_n^2} = 0,$$

y — довільна точка з C. Тоді для довільного $u_1 \in C$ послідовність (u_n), що визначається як

$$u_{n+1} = \alpha_n y + (1 - \alpha_n) T(t_n) u_n,$$

сильно збігається до точки із $F(\boldsymbol{T})$, що є найближчою до y.

Доведення. За теоремою 2 послідовність (x_n), що визначається як

$$x_n = \alpha_n y + (1 - \alpha_n) T(t_n) x_n \tag{1}$$

буде збіжною до точки $\overline{y}$, що є найближчою точкою в множині $F(\boldsymbol{T})$ до точки y. Тоді

$$\begin{aligned} \|u_{n+1} - x_n\| = (1 - \alpha_n) \|T(t_n)u_n - T(t_n)x_n\| \leqslant \\ \leqslant (1 - \alpha_n) \|u_n - x_{n-1}\| + (1 - \alpha_n) \|x_n - x_{n-1}\| . \end{aligned} \tag{2}$$

Доведемо, що послідовність (x_n) обмежена. Нехай $p \in F(\boldsymbol{T})$. Тоді

$$\begin{aligned} \|x_n - p\| \leqslant \alpha_n \|y - p\| + (1 - \alpha_n) \|T(t_n)x_n - T(t_n)p\| \leqslant \\ \leqslant \alpha_n \|y - p\| + (1 - \alpha_n) \|x_n - p\| . \end{aligned}$$

Тобто

$$\|x_n - p\| \leqslant \|y - p\| \quad \forall n \in \mathbb{N}.$$

Спробуємо оцінити $\|x_n - x_{n-1}\|$. Для x_{n-1} можна записати

$$x_{n-1} = \alpha_{n-1} y + (1 - \alpha_{n-1}) T(t_{n-1}) x_{n-1}. \tag{3}$$

Тоді віднявши від (1) (3), матимемо

$$\begin{aligned}\|x_n - x_{n-1}\| \leqslant{}& |\alpha_n - \alpha_{n-1}| \, \|y - T(t_{n-1})x_{n-1})\| + \\ &+ (1-\alpha_n)\,\|T(t_n)x_n - T(t_{n-1})x_{n-1})\| \leqslant \\ &\leqslant |\alpha_n - \alpha_{n-1}| \, \|y - T(t_{n-1})x_{n-1})\| + \\ &+ (1-\alpha_n)\,\|T(t_n)x_n - T(t_n)x_{n-1}\| + \\ &+ (1-\alpha_n)\,\|T(t_n)x_{n-1} - T(t_{n-1})x_{n-1})\| \leqslant \\ &\leqslant |\alpha_n - \alpha_{n-1}| \, \|y - T(t_{n-1})x_{n-1})\| + \\ &+ (1-\alpha_n)\,\|x_n - x_{n-1}\| + \\ &+ (1-\alpha_n)\,\|T(t_n)x_{n-1} - T(t_{n-1})x_{n-1})\| \leqslant \\ &\leqslant |\alpha_n - \alpha_{n-1}| \, \|y - T(t_{n-1})x_{n-1})\| + \\ &+ (1-\alpha_n)\,\|x_n - x_{n-1}\| + \\ &+ (1-\alpha_n) \cdot L|t_n - t_{n-1}|.\end{aligned}$$

Оскільки послідовність (x_n) — обмежена, то нехай

$$M = \max\{\|y - T(t_{n-1})x_{n-1})\|\,, L(1-\alpha_n)|\} < \infty.$$

Тоді

$$\|x_n - x_{n-1}\| \leqslant \frac{|\alpha_n - \alpha_{n-1}| + |t_n - t_{n-1}|}{\alpha_n} M. \tag{4}$$

Повертаючись до (2), можемо записати

$$\begin{aligned}\|u_{n+1} - x_n\| \leqslant (1-\alpha_n)\,\|u_n - x_{n-1}\| + \\ + \frac{|\alpha_n - \alpha_{n-1}| + |t_n - t_{n-1}|}{\alpha_n}(1-\alpha_n)M.\end{aligned} \tag{5}$$

Тоді за лемою 1 послідовність $\|u_{n+1} - x_n\| \to 0$ при $n \to \infty$. А значить і

$$u_n \to P_{F(\boldsymbol{T})}y, \quad n \to \infty,$$

що і треба було довести. □

Зауваження 2. Прикладами послідовностей (α_n), (t_n) можуть бути, наприклад,

$$\alpha_n = \frac{1}{\sqrt{n}}, \quad t_n = \frac{1}{\sqrt[4]{n}}.$$

2. В'язка апроксимація нерухомої точки

Нехай C — замкнена опукла множина гільбертового простору H, $Q : C \to C$ — стискаючий оператор, $F(\boldsymbol{T}) \neq \emptyset$ — множина нерухомих точок ліпшицевої напівгрупи нерозтягуючих операторів $\boldsymbol{T} = \{T(t) : t \in \mathbb{R}^+\}$, що діють в C. Розглянемо задачу

$$\text{знайти } \ x \in F(\boldsymbol{T}) : \ (x - Qx, y - x) \geq 0 \ \ \forall\, y \in F(\boldsymbol{T}). \tag{6}$$

Варіаційна нерівність (6) має єдиний розв'язок $\bar{x}$, який можна охарактеризувати як нерухому точку стискаючого оператора

$$C \ni y \mapsto P_{F(\boldsymbol{T})}Qy \in C,$$

де $P_{F(\boldsymbol{T})}$ — оператор проектування на множину $F(\boldsymbol{T})$.

Незначні модифікації наведених міркувань дозволяють отримати таку теорему.

Теорема 4. *Нехай C — замкнена опукла підмножина гільбертового простору H, $Q : C \to C$ — стискаючий оператор, $\boldsymbol{T} = \{T(t) : t \in \mathbb{R}^+\}$ — ліпшицева напівгрупа нерозтягуючих операторів, що діють в C, $F(\boldsymbol{T}) \neq \emptyset$, (α_n) і (t_n) — послідовності дійсних чисел таких, що*

$$0 < \alpha_n < 1, \ t_n > 0, \ \sum_{n=1}^{\infty} a_n = +\infty,$$

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{a_n}{t_n} = \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n-1}}{\alpha_n^2} = \lim_{n\to\infty} \frac{t_n - t_{n-1}}{\alpha_n^2} = 0.$$

Тоді для довільного $u_1 \in C$ послідовність (u_n), що визначається як

$$u_{n+1} = \alpha_n Q u_n + (1 - \alpha_n) T(t_n) u_n,$$

сильно збігається до єдиного розв'язку варіаційної нерівності (6).

Література

1. Browder F. E. Fixed-point theorem for noncompact mappings in Hilbert space / F. E. Browder // Proc. Nat. Acad. Sci. USA — 1965. — 53. — P. 1272–1276.
2. Browder F. E. Nonexpansive nonlinear operators in a Banach space / F. E. Browder // Proc. Nat. Acad. Sci. USA. — 1965. — 54. — P. 1041–1044.
3. Browder F. E. Convergence of approximates to fixed points of nonexpansive nonlinear mappings in Banach spaces / F. E. Browder // Arch. Ration. Mech. Anal. — 1967. — 24. — P. 82–90.
4. Halpern B. Fixed-points of nonexpanding maps / B. Halpern // Bull. Amer. Math. Soc. — 1967. — 73. — P. 591–597.
5. Suzuki T. On strong convergence to common fixed points of nonexpansive semigroups in Hilbert spaces / T. Suzuki // Proc. Amer. Math. Soc. — 2002. — 131. — P. 2133–2136.
6. Xu H. K. Another control condition in an iterative method for nonexpansive mappings / H. K. Xu // Bull. Austral. Math. Soc. — 2002. — 65 — P. 109–113.
7. Бакушинский А. Б. Некорректные задачи. Численные методы и приложения / А. Б. Бакушинский, А. В. Гончарский. — Москва: Изд-во МГУ, 1989. — 200 с.

Факультет кібернетики, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64, Київ, 01601, Україна.

Надійшла 07.11.2011

Журнал обчислювальної та прикладної математики 2012, №1(107) Journal of Computational & Applied Mathematics

УДК 519.8

КОНЦЕПЦИЯ РАВНОВЕСИЯ ПО НЭШУ И ЕЕ РАЗВИТИЕ

С. О. Мащенко

Резюме. Работа посвящена обзору литературы, в которой отражена концепция равновесия по Нэшу и ее развитие. Рассматривается равновесие по Нэшу в классических играх, в играх с векторными функциями выигрыша игроков (так называемых многокритериальных играх), в обобщенных играх (игры, которые заданы отношениями предпочтения игроков). Значительное внимание уделено развитию концепции равновесия по Нэшу в теории некооперативных игр. Рассматриваются коалиционные равновесия, равновесия по Бержу, индивидуально-оптимальные равновесия, равновесия в условиях неопределенности.

Введение

Некооперативные (бескоалиционные) игры исследуют принятие решений в условиях конфликта в предположении, что игроки действуют независимо один от другого. Соглашения между игроками хотя в принципе возможны, но имеют ограниченный и необязательный характер. Каждый игрок может нарушить соглашение без наказания.

Применение того или иного принципа оптимальности в некооперативных играх существенно зависит от информированности игроков. В условиях полной неинформированности игроков (каждый игрок знает лишь свою функцию выигрыша, игра происходит лишь один раз) известны: недоминируемые и доминирующие стратегии; осторожные стратегии. В условиях “несимметричной информированности” игроков (некоторые игроки полностью информированы, а другие полностью не информированы) используют равновесие по Штакельбергу и его обобщения. В случае полной информированности игроков (каждый игрок знает не только свою функцию выигрыша, но и функции выигрыша других игроков, игра может происходит произвольное количество раз) основным принципом оптимальности является равновесие по Нэшу.

Основы некооперативного поведения полностью информированных игроков изложены в исследованиях лауреатов Нобелевской премии по экономике за 1994 год: Дж. Нэша [59], Д. Харшаньи и Р. Зельтена [36].

Понятие равновесия по Нэшу нашло широкое применение при решении многих прикладных конфликтных задач. Однако потребности практического применения этого принципа оптимальности вызывали необходимость его “улучшения”. Можно выделить три основных направления.

Первое заключается в расширении сферы применения концепции равновесия. В частности, это равновесия в условиях неопределенности как в статических, так и в динамических и дифференциальных играх.

Второе связано с желанием выбрать из равновесий по Нэшу те, которые дополнительно имеют определенные “полезные свойства” (см. например, монографию [36] Д. Харшаньи и Р. Зельтена).

Третье направление связано с рассмотрением новых концепций оптимальности, которые хотя бы частично снимают “негативные свойства” равновесий по Нэшу и более адекватны реальным условиям конфликта. В кооперативных играх это: сильное равновесие по Нэшу [29, с. 134–135], α-ядро [29, с. 160–169] (активные равновесия [11, с. 143–157; 35, с. 70-92]), β-ядро [29, с. 169–175] (равновесия угроз и контр-угроз [5, с. 46–62]), γ-ядро [29, с. 176–178]. В некооперативных играх это: коалиционные равновесия [3, с. 96], равновесия по Бержу [11] и индивидуально-оптимальные равновесия [21; 22].

Рассмотрим игру в, так называемой [29, с. 15], нормальной форме, когда известны лишь множества стратегий игроков и их функции выигрыша.

Пусть $N = \{1, 2, \dots, n\}$ — конечное множество из n ($n \geq 2$) игроков. Игрой G в нормальной форме называется совокупность $(X_i, u_i; i \in N)$, которая задает для каждого игрока $i \in N$:

- множество стратегий X_i (элементы множества стратегий называют стратегиями и обозначают x_i);
- функции выигрыша игроков $u_i(x)$, $i \in N$, которые определены на множестве ситуаций игры $X = \prod_{i \in N} X_i$ и максимизируются (элемент $x = (x_1, \dots, x_n) = (x_i)_{i \in N}$ множества X называется ситуацией игры).

1. Равновесие по Нэшу и его обобщения

Равновесие по Нэшу в классических играх. Если игроки полностью информированы, то им разумно было бы заключить определенное необязательное соглашение, которое ни одному из них не было бы выгодно нарушать. Идея стабильного соглашения приводит к следующему определению [59].

Обозначим $x_{N\setminus i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = (x_i)_{i \in N\setminus\{i\}}$ — набор стратегий, так называемой дополняющей коалиции игрока $i \in N$.

Ситуация x^* называется равновесием (строгим равновесием) по Нэшу игры G в нормальной форме $(X_i, u_i, i \in N)$, если

$$u_i(x^*) \geq u_i(x_i, x^*_{N\setminus i}) \quad \left(u_i(x^*) > u_i(x_i, x^*_{N\setminus i})\right), \quad \forall x_i \in X_i, \forall i \in N,$$

т. е. каждому игроку в отдельности не выгодно изменять свою стратегию x_i^* на другую. Множество равновесий по Нэшу будем обозначать NE, а строгих равновесий по Нэшу — SNE. Справедливо отношение $SNE \subseteq NE$.

Понятие ситуации равновесия в частном случае еще в 19-м столетии использовалось О. Курно при анализе модели дуаполии. Определение ситуации равновесия для игры многих лиц принадлежит лауреату Нобелевской премии по экономики за 1994 год Дж. Нэшу [59]. Критический разбор понятия ситуации равновесия, в частности, анализ игр "семейный спор" и "дилемма узника", сделаны Р. Льюсом и Х. Райфой [14].

Концепция равновесия по Нэшу мотивируется следующими сценариями игры.

Полукооперативный сценарий (по Дж. Нэшу [59]) заключается в совместном выборе игроками ситуации игры, которая станет основой для необязательного соглашения между ними. После этого обмен информацией между игроками прекращается и каждый принимает решение независимо (при этом соглашение можно нарушить). Тогда и только тогда, когда выбранная ситуация будет равновесием Нэша, она будет стабильным соглашением.

Динамический сценарий (по О. Курно [29, с. 69–70]) заключается в нахождении зависимых от предыдущий партий игры стабильных ситуаций, в которых каждый игрок выбирает оптимальную для себя стратегию при условии, что другие игроки не изменяют своих стратегий. Если такая процедура сходится к некоторой ситуации игры, то эта ситуация будет равновесной по Нэшу.

"Метатеоретический сценарий" (Дж. фон Нейман и О. Моргенштерн [30]) заключается в компетентной рекомендации игрокам придерживаться некоторой ситуации игры. Поскольку каждый разумный и полностью информированный игрок может самостоятельно восстановить аргументацию и найти рекомендованную ситуацию, то для того, чтобы эгоистичные игроки прислушивались к рекомендациям теории необходимо, чтоб эта ситуация была равновесием по Нэшу.

Рассмотрим основные свойства равновесий Нэша.

Во-первых, равновесие Нэша удовлетворяют принципу индивидуальной рациональности [29, с. 72]:

$$x^* \in NE \quad \Rightarrow \quad u_i\left(x^*\right) \geq \sup_{y_i \in X_i} \inf_{y_{N\setminus i} \in X_{N\setminus i}} u_i\left(y_i, y_{N\setminus i}\right), \ \forall i \in N,$$

т. е. выигрыш каждого игрока в равновесии Нэша не хуже, чем при выборе ним осторожной стратегии. С другой стороны, стратегия игрока в равновесии Нэша может не быть осторожной. Это утверждение обосновывается многочисленными примерами, в частности игрой "дилемма узника" [14].

Во-вторых, равновесия Нэша не удовлетворяют принципу коллективной рациональности, т. е. они могут не быть оптимальными по Парето [14]. С другой стороны, сосуществование нескольких различных паретовских ситуаций, которые равновесны по Нэшу, приводит к борьбе за лидерство [29, с. 73–74], которая только обостряет конфликт. Лишь только, когда равновесие по Нэшу единственно и оптимально по Парето, его можно считать решением конфликта.

В-третьих, концепция равновесия Нэша обобщает, так называемое, сложное равновесие [29, с. 74–75] в следующем смысле. Если игра G разрешима по доминированию, а множества стратегий игроков конечны, то любое сложное равновесие будет равновесием по Нэшу. Следовательно, сложное поведение всегда приводит к равновесным по Нэшем ситуациям. Обратное утверждение не верно. В общем случае стратегии, которые образуют равновесие Нэша, могут быть доминируемыми (не строго) [29, с. 75].

Равновесия Нэша (в чистых стратегиях) не всегда существуют даже в простых играх с конечными множествами стратегий игроков. Достаточные условия их существования формализует известная теорема Нэша [32].

Пусть $\forall i \in N$ множество стратегий X_i является выпуклым и компактным подмножеством некоторого топологического векторного пространства (вообще говоря, своего для каждого i). Пусть все u_i, $i \in N$, — непрерывные вещественнозначные функции на X такие, что для каждого $x_{N \setminus i} \in X_{N \setminus i}$ функции одной переменной $u_i\left(x_i, X_{N \setminus i}\right)$ вогнуты по $x_i \in X_i$. Тогда множество равновесий по Нэшу будет непустым и компактным.

Доказательство этой теоремы основывается на теореме Какутани [53] из выпуклого анализа о неподвижной точке. В [29] приведен значительно более короткий вариант доказательства теоремы Нэша, который опирается на теорему Брауера [50] и лемму Б. Кнастера, К. Куратовского и С. Мазуркевича [54]. Теорема о существовании ситуации равновесия в игре многих лиц с квазивогнутыми функциями выигрыша доказана Х. Никайдо и К. Йсода [31].

Следствием из теоремы Нэша является аналогичный результат для антагонистических игр — теорема Дж. фон Неймана и О. Моргенштерна [30].

В работе [41] условия теоремы Нэша ослаблены. Достаточным условием непустого множества равновесий по Нэшу игры G в нормальной форме $(X_i, u_i; i \in N)$ является компактность множеств стратегий игроков X_i, $i \in N$, которые могут быть произвольными подмножествами метрических топологических векторных пространств, непрерывность функций выигрыша игроков $u_i(x)$, $i \in N$, и одноэлементность множества $\arg\max_{y_{S_i} \in X_{S_i}} u_i(y_{S_i}, x_{N \setminus S_i})$ для каждого $x_{N \setminus S_i} \in X_{N \setminus S_i}$.

Доказательство этой теоремы опирается на обобщение теоремы о неподвижной точке (О. Арино, С. Гаутьер, Дж. Пино [42]) и на теорему К. Бержа о максимуме [3, с. 82–85]. Как видим, эта теорема, в отличие от теоремы Нэша, не требует выпуклости вниз функций выигрыша, а лишь их непрерывность. С другой стороны, условие одноэлементности множества $\arg\max_{y_{S_i} \in X_{S_i}} u_i\left(y_{S_i}, x_{N \setminus S_i}\right)$ для каждого $x_{N \setminus S_i} \in X_{N \setminus S_i}$ не конструктивно.

В [40] рассматривается игра с бесконечным множеством игроков. Доказывается непустота множества равновесий по Нэшу игры G в нормальной форме $(X_i, u_i, i \in N)$ с конечным или бесконечным множеством игроков в предположениях: множества X_i, $i \in N$, являются компактами некоторых топологических векторных пространств, функции выигрыша $u_i\left(x\right)$, $i \in N$, — непрерывны и квазивогнуты по $x_i \in X_i$ для всех $x_{N \setminus i} \in X_{N \setminus i}$.

Рассмотрим вопрос нахождения равновесий Нэша (если они существуют). Согласно определению для этого необходимо решить систему взаимосвязанных задач оптимизации:

$$u_i(x^*) = \max_{x_i \in X_i} u_i(x_i, x^*_{N \setminus i}), \ \ i \in N,$$

что часто бывает достаточно серьезной проблемой (проблема нахождения глобального максимума).

Если для каждого игрока $i \in N$ функция $u_i\,(x)$ вогнута по $x_i \in X_i$, то нахождение равновесий Нэша сводится к системе задач поиска локального оптимума [29, с. 80]. В частности, если $x^* \in NE$ — внутренняя точка множества ситуаций X, а функции $u_i(x)$, $i \in N$, дифференцируемы, то эта система эквивалентна

$$\partial u_i(x)/\partial x_i = 0, \ \ i \in N.$$

Иллюстрацию описанного метода на примерах можно найти в [29].

Одним из мощных современных инструментов нахождения равновесий по Нэшу являются численные методы равновесного программирования [1; 2]. Этот подход основывается на рассмотрении неравенства, которое замещает отсутствующее у равновесных задач свойство монотонности, и градиентный спуск с управлением в виде прогноза, который компенсирует отсутствующее свойство потенциальности.

Если равновесие Нэша — не единственно, то возникает сложная проблема выбора, анализу которой по критериям "доминирования по выигрышу" и "доминирования по риску" посвящена монография [36].

Наиболее популярным подходом к решению игрового конфликта, который применяют в случае отсутствия равновесий Нэша, является их поиск в смешанных стратегиях [32].

По современной терминологии [13, с. 146] смешанным расширением игры G в нормальной форме $(X_i, u_i; i \in N)$ называется игра $\bar{G}$ в нормальной форме $(M_i, \bar{u}_i; i \in N)$, где:

– M_i — множество, заданных на множестве стратегий X_i, вероятностных мер, которое содержит все простые вероятностные меры и называется множеством смешанных стратегий игрока $i \in N$;
– $\bar{u}_i : M = \prod_{i \in N} M_i \to R^1$ — математическое ожидание выигрыша игрока $i \in N$, которое определяется как интеграл

$$\bar{u}_i(\mu) = \int_X u_i(x) d\mu(x), \mu \in M.$$

В этом определении допускается наличие измеримой структуры на множествах стратегий X_i и соответствующая интегрируемость функций выигрыша u_i, $i \in N$, исходной игры.

Следствием из теоремы Нэша для смешанных расширений игр являются известные теоремы Дж. Нэша о смешанном расширении игры [32], в соответствии с которой в смешанном расширении конечной игры G всегда существуют ситуации равновесия.

Согласно теореме Гликсберга [6], если в игре G множества стратегий игроков X_i, $i \in N$, — компактные подмножества топологических пространств, функции выигрыша u_i — непрерывны по $x \in X$, то в смешанном расширении G существуют равновесия.

Нахождение равновесий по Нэшу в произвольной, даже конечной, игре вызывает определенные трудности. Случай антагонистической игры разработан наиболее полно. Нахождение равновесий по Нэшу в антагонистической игре сводится к паре двойственных задач линейного программирования [33]. В более общем случае не антагонистической игры двух лиц, так называемой биматричной игры, равновесия по Нэшу вычисляется с помощью использования различных линейных методов, в основе которых лежит линейное программирование. Исторически одним из первых подходов является алгоритм Лемке-Хаусона [57] для игр двух лиц. Согласно этому алгоритму в ситуации равновесия игры смешанная стратегия одного игрока уравновешивает выигрыш другого при использовании им чистой стратегии.

Существенными недостатками концепции равновесия Нэша в смешанных стратегиях является доминируемость равновесий по Парето [29, с. 138] и предположение о бесконечном количестве реализаций игры.

Равновесия по Нэшу в играх с векторными функциями выигрыша. Игры с вектор-функциями выигрышей [4; 7–9; 44; 48; 51; 55; 58; 64; 66–69] (многокритериальные, векторные игры) могут более адекватно описать реальные ситуации, поскольку вполне естественно, что участники конфликтов оперируют несколькими критериями. Началом исследования векторных игр можно считать роботы Д. Блеквела [4] об антагонистических играх и Л. Шепли [64], где предложена концепция равновесия многокритериальной игры. Эти исследования были продолжены П. Борном, С. Тиджем, ван дер Арсеном [45]. Для игр с ненулевой суммой аналогичный подход разрабатывали М. Зелени [69], Дж. Найвенхаус [61], Х. Киорли [47]. Теоремы существования для игр с векторными функциями выигрыша сформулировал С. Ванг [66]. В работе Дж. Жао [70] определено равновесие многокритериальной игры как обобщение задачи векторной оптимизации. Он же построил собственно эффективные решения таких игр. Обзор результатов по этой проблематике, изложен в [65].

По аналогии со скалярной игрой, многокритериальная, по существу, представляет собой задачу принятия решений с несколькими критериями в условиях конфликта. Понятно, что методы решения таких задач являются комбинацией методов классической теории игр и многокритериальной оптимизации.

Так, например, широко распространенное применение понятия равновесия Нэша к многокритериальной игре изложено в [45; 47; 58; 66; 61; 65; 67; 70]. Поскольку функции выигрышей — векторные, то используется также принцип оптимальности Парето [34]. Для антагонистических игр с несколькими критериями также обобщается понятие седловой точки [58] и цены игры. Векторный максимин и минимакс используются для описания осторожного поведения в [58]. В [55] рассмотрены многокритериальные игры в развернутой форме, когда задано дерево игры.

Хотя в целом векторная игра не сводится к скалярной, некоторые авторы приводят условия существования равновесий, которые являются аналогичны обычным играм [66; 67], хотя наличие нескольких критериев вносит дополнительные условия и ограничения на существование решения.

В [66] обобщается понятие игры в нормальной форме на многокритериальный случай. Многокритериальной игрой называется совокупность, которая содержит для каждого игрока $i \in N$:

– множество стратегий X_i, которое является произвольным подмножеством нормированного векторного пространства R^{n_i};
– векторы функций выигрыша игроков $U_i(x) = \left(u_i^j(x)\right)_{j \in L_i}, i \in N$, которые определены на множестве ситуаций игры $X = \prod_{j \in N} X_i$, являются непрерывными и максимизируются.

Ситуация $x \in X$ игры называется равновесием Парето–Нэша, если для каждого $i \in N$ не существует $x_i \in X_i$, для которого

$$u_i^j(x_i, \bar{x}_{N \setminus i}) \geq u_i^j(\bar{x}) \quad \forall j \in L_i,$$

и по крайней мере одно неравенство строгое.

Следующий результат [66] представляет собой условия существования равновесия. Он разработан для случая, когда игроки пытаются минимизировать свои функции выигрыша.

Пусть вектор

$$\omega_i \in T_i = \left\{ \omega_i = \left(\omega_i^j\right)_{j \in L_i} \mid \omega_i^j > 0,\ j \in L_i,\ \sum_{j \in L_i} \omega_i^j = 1 \right\},$$

а ситуация $x \in X$. Определим функцию и множество:

$$S^\omega(x, y) = \sum_{i \in N} \omega_i^T U_i(y_i, x_{N \setminus i}),$$

$$M^\omega(x) = \left\{ y^* \in X \mid S^\omega(x, y^*) = \min_{y \in X} S^\omega(x, y) \right\}.$$

В предположении компактности и замкнутости множеств стратегий X_i, $i \in N$, если для всех $i \in N$ существуют такие $\omega_i \in T_i$, что функция $S^\omega(x, y)$ квазивыпукла по $y \in X$ при любом фиксированном $x \in X$, в [66] доказано, что игра имеет хотя бы одно равновесие Парето–Нэша. Доказательство использует обобщенную теорему Какутани о неподвижной точке.

В [48] рассматривается случай, когда множество стратегий каждого игрока зависит от поведения других.

Пусть для каждого $i \in N$ задано отображение $\Gamma_i : X_{N \setminus i} \to 2^{X_i}$. Совокупность $(X_i, \Gamma_i, U_i; i \in N)$ называется игрой на зависимых множествах.

В такой игре отображение Γ_i моделирует зависимость возможностей выбора игрока $i \in N$ от поведения других. Так, например, если заданы стратегии участников $N \setminus \{i\}$, то игрок i может выбирать свою стратегию не из всего множества X_i, а лишь из его части $\Gamma_i(x_{N\setminus i}) \subseteq X_i$. Следовательно, для игры на зависимых множествах не все ситуации из X — допустимы. Такие игры еще называются играми с запрещенными состояниями.

Ситуация $x \in X$ называется допустимой в игре на зависимых множествах, $(X_i, \Gamma_i, U_i; i \in N)$ если $x_i \in \Gamma_i(x_{N\setminus i})$ для $\forall i \in N$. Множество допустимых ситуаций игры обозначим $D = \{x \in X \mid x_i \in \Gamma_i(x_{N\setminus i}), \forall i \in N\}$.

Для формализации понятия равновесия вводятся вспомогательные подмножества пространства $R^k, k \in N : R^k_+ = \{b = (b_1, \dots, b_k) \in R^k \mid b_p \geq 0, p = \overline{1,k}\}$ и множество $int(R^k_+) = \{b = (b_1, \dots, b_k) \in R^k \mid b_p > 0, p = \overline{1,k}\}$. Здесь R^k_+ — положительный ортант пространства R^k, а $int\left(R^k_+\right)$ — его внутренность.

Последующие результаты в [48] получены для случая, когда игроки минимизируют функции выигрыша. Поскольку это не ограничивает общности, приведем их в оригинальном виде.

Для игры на зависимых множествах допустимая ситуация $x^* \in D$ называется обобщенным равновесием Парето–Нэша (слабым обобщенным равновесием Парето–Нэша), если $\forall i \in N \ \nexists x_i \in \Gamma_i(x_{n\setminus i})$, что

$$U_i(x^*) - U_i(x_i, x^*_{N\setminus i}) \in R^{|L_i|}_+ \setminus \{0\}$$

$$\left(U_i(x^*) - U_i(x_i, x^*_{N\setminus i}) \in int(R^{|L_i|}_+) \setminus \{0\}\right).$$

Это определение отличается от общего определения равновесия Парето–Нэша лишь условием, что минимизируя свой выигрыш, игрок i выбирает стратегию учитывая выбор соперников.

Пусть заданы векторы $\omega_i \in R^{|L_i|}_+ \setminus \{0\}, i \in N$. Для игры на зависимых множествах допустимая ситуация $x^* \in D$ называется обобщенным взвешенным равновесием Парето–Нэша относительно $\omega = (\omega_i)_{i\in N}$, если

$$\omega_i \cdot U_i(x^*) = \min_{x_i \in \Gamma_i(x^*_{N\setminus I})} \omega_i \cdot U_i(x_i, x^*_{N\setminus i}),$$

где “$\cdot$” — операция скалярного произведения векторов.

В этом определении, фактически, рассматривается скалярная игра с функциями выигрыша, которые являются линейными свертками критериев игроков.

Определяются следующие отображения:

$$\Phi_\omega : X \times X \to R^1, \quad \Gamma : X \to 2^X,$$

где

$$\Phi_\omega(x, y) = \sum_{i\in N} \omega_i \cdot (U_i(x) - U_i(y_i, x_{N\setminus i})), \quad \Gamma(x) = \prod_{i\in N} \Gamma_i(x_{N\setminus i}).$$

Поскольку ситуация $x \in X$ допустима в игре тогда и только тогда, когда $x \in \Gamma(x)$, т. е. когда является неподвижной точкой отображения Γ, то множество допустимых точек можно представить в виде:

$$D = \{x \in X \mid x \in \Gamma(x)\}.$$

В [48] доказано, что в игре на зависимых множествах с заданными векторами $\omega_i \in R_+^{|L_i|} \setminus \{0\}$, $i \in N$, ($\omega_i \in int(R_+^{|L_i|}), i \in N$) каждое обобщенное (слабо обобщенное) взвешенное равновесие Парето–Нэша относительно ω является также обобщенным равновесием (слабым обобщенным равновесием) Парето–Нэша.

В [48] представлены также условия существования равновесий в многокритериальных играх на зависимых множествах.

Для игры на зависимых множествах с заданными векторами $\omega_i \in R_+^{|L_i|} \setminus \{0\}, i \in N$, для существования обобщенного взвешенного равновесия Парето–Нэша относительно достаточно выполнения для всех $i \in N$ таких условий:

- X_i — компакт некоторого евклидового пространства;
- отображение Γ_i — выпукло, полунепрерывно снизу и не пусто;
- допустимое множество D — замкнуто;
- множество $\{(x, y) \in X \times X \mid \Phi_\omega(x, y) \le 0\}$ — замкнуто;
- для всех $x \in X$ множество $\{y \in \Gamma(x) \mid \Phi_\omega(x, y) > 0\}$ — выпукло.

Условие компактности множеств стратегий можно заменить условием существования компактных подмножеств X_i, декартово произведение которых имеет непустое пересечение с D.

Таким образом, для игры на зависимых множествах с заданными векторами $\omega_i \in R_+^{|L_i|} \setminus \{0\}$, $i \in N$, для существования обобщенного взвешенного равновесия Парето–Нэша относительно ω достаточно выполнения для всех $i \in N$ условий:

- X_i — подмножество некоторого евклидового пространства;
- отображение Γ_i — выпукло, полунепрерывно снизу и не пусто;
- допустимое множество D — замкнуто;
- множество $\{(x, y) \in X \times X \mid \Phi_\omega(x, y) \le 0\}$ — замкнуто;
- для всех $x \in X$ множество $\{y \in \Gamma(x) \mid \Phi_\omega(x, y) > 0\}$ — выпукло;
- $\forall i \in N$ существуют компакты $D_i \subseteq X_i$, $D_i \cap \Gamma_i(x_{N\setminus i}) \neq \emptyset, \forall x_{N\setminus i} \in X_{N\setminus i}$, $\max_{i\in N} \sup_{y_i \in D_i \cap \Gamma_i(x_{N\setminus i})} \omega_i \cdot (U_i(x) - U_i(y, x_{N\setminus i})) > 0, \forall x \in D \setminus \prod_{i\in N} D_i$.

В [48] получены также другие достаточные условия существования обобщенных равновесий Парето–Нэша. Например, для каждого $i \in N$ достаточно, что бы выполнялись условия:

- X_i — компакт евклидового пространства;
- отображение Γ_i — выпукло, полунепрерывно снизу и не пусто;
- функция U_i — непрерывна;
- функция $u_i^j(x_i, x_{N\setminus i})$ выпукла по x_i для $\forall j \in L_i$, $\forall x_{N\setminus i} \in X_{N\setminus i}$;
- допустимое множество D — замкнуто;
- для $\forall x_{N\setminus i} \in X_{N\setminus i} \quad D_i \cap \Gamma_i(x_{N\setminus i}) \neq \emptyset$;
- $\exists i \in N \quad \exists y_i \in D_i \cap \Gamma_i(x_{N\setminus i})$:$U_i(x) - U_i(y_i, x_{N\setminus i}) \in R_+^{|L_i|} \setminus \{0\}$, $\forall x \in D \setminus \prod_{i\in N} D_i$.

Для слабых равновесий Парето–Нэша эти требования аналогичны, но условие $\omega_i \in R_+^{|L_i|} \setminus \{0\}$ нужно заменить на $\omega_i \in int(R_+^{|L_i|})$ для каждого $i \in N$.

Условия непрерывности функций выигрыша игроков можно ослабить. Тогда для игры на зависимых множествах для существования обобщенного взвешенного равновесия Парето–Нэша достаточно выполнения для всех $i \in N$ условий:

- X_i — компакт некоторого топологического пространства;
- отображение Γ_i — выпукло, полунепрерывно снизу и не пусто;
- u_i^j — полунепрерывны снизу на X, $\forall j \in L_i$;
- отображение $x_{N\setminus i} \to u_i^j(y_i, x_{N\setminus i})$ — полунепрерывно сверху на $X_{N\setminus i}$, $\forall y_i \in X_i$;
- отображение $y_i \to u_i^j(y_i, x_{N\setminus i})$ — выпукло на $X_i, \forall x_{N\setminus i} \in X_{N\setminus i}$;
- допустимое множество D — замкнуто;
- существуют компактные замкнутые подмножества $K, X_0 \subseteq X$, что $\forall x \in D \setminus K \ \Gamma(x) \cap conv(X_0 \cap \{x\}) \neq 0$ и $\forall x \in D \setminus K \quad \exists y_i \in \Gamma(x) \cap conv(X_0 \cap \{x\}) \neq 0$ такие, что $\forall i \in N \quad \forall j \in L_i \quad u_i^j(x) > u_i^j(y_i, x_{N\setminus i})$.

Приведенные выше достаточные условия для игр на зависимых множествах обобщают результаты из [67] для обычных векторных игр. Отметим, также, что они более общие, чем классические (непрерывность и квазивыпуклость функций выигрыша).

В [68] для каждого $i \in N$ представленные условия существования слабого равновесия Парето–Нэша в случае нормированного пространства:

- X_i — выпуклое компактное подмножество нормированного пространства;
- u_i^j — полунепрерывны сверху на $\forall j \in L_i$;
- отображение $X_{N\setminus i} \to u_i^j(y_i, x_{N\setminus i})$; полунепрерывно сверху на $X_{N\setminus i}$ $\forall y_i \in X_i$;
- отображение $y_i \to u_i^j(y_i, x_{N\setminus i})$ — выпукло для $\forall x_{N\setminus i} \in X_{N\setminus i}$.

В [44] исследована структура равновесия биматричных многокритериальных игр. Результаты носят негативный характер. Показано, что нахождение оптимальных смешанных стратегий не сводится, как в однокритериальном случае, к решению системы линейных неравенств.

По аналогии со скалярной, многокритериальная биматричная игра задается двумя наборами матриц размерности $m \times n : A = (A_S)_{s\in S}, B = (A_t)_{t\in T}$, где $A_S = [(A_S)_{mn}]_{(m,n)\in M\times N}$ и $B_t = [(A_t)_{mn}]_{(m,n)\in M\times N}$; M и N — множества стратегий первого и второго игрока соответственно. Каждая из матриц $A_S(A_t)$ описывает соответствующий критерий первого (второго) участника. Множества смешанных стратегий обозначаются $\Delta(M)$ и $\Delta(N)$.

Если выбраны смешанные стратегии p и q, то векторы выигрышей принимают вид: $pAq = (pA_Sq)_{s\in S}$ и $pBq = (pB_tq)_{t\in T}$.

Для любого множества $I \subseteq M$ через $\Delta(I)$ обозначается множество смешанных стратегий с носителем I, а через $U(I)$ — множество смешанных стратегий второго игрока, против которых все стратегии из $\Delta(I)$ является наилучшими ответами. Так же для $J \subseteq M$ определяется и $\Delta(J)$ и $U(J)$.

В биматричной однокритериальной игре все множества $\Delta(I)$, $U(I)$, $\Delta(J)$ и $U(J)$ являются линейными многогранниками. Поэтому можно построить систему линейных неравенств, которые их задают. Множество равновесий в смешанных стратегиях будет также линейным многогранником.

В [44] доказано, что во многокритериальной биматричной игре (A, B) равновесие в смешанных стратегиях задается объединением по $\forall I \subseteq M$ и $\forall J \subseteq N$ множеств $(\Delta(I) \cap U(J)) \times (\Delta(J) \cap U(J))$.

Поскольку при переходе к многокритериальности “линейность” решения может теряться, то компоненты $U(I)$ и $U(J)$ могут быть представлены квадратичными функциями. В [44] это иллюстрировано примером игры 3×3. Лишь в частном случае для класса игр $2 \times n$ можно гарантировать, что равновесие $(\Delta(I) \cap U(J)) \times (\Delta(J) \cap U(J))$ будет описываться линейными неравенствами.

Следующий результат [58], показывает, что при выполнении условий теоремы Нэша множество равновесий Парето–Нэша многокритериальной игры будет не пустым. При доказательстве используется метод линейной свертки критериев и для скаляризованной игры применяется теорема Нэша.

В [63] рассмотрен пример практического применения равновесий многокритериальной игры для коалиционной игры двух участников, которыми могут быть политические партии, фирмы или профессиональные союзы.

Таким образом, на сегодняшний день рядом авторов [44; 48; 51; 55; 58; 66–68] проведено исследование многокритериальных игр и проблемы существования их равновесий. Следует отметить, что полученные результаты во многом подобны и отличаются лишь деталями в условиях на функции выигрыша игроков, пространства и множества стратегий. Методы доказательств также подобны. Авторы, в основном, используют известные теоремы о неподвижной точке (или их обобщение), также проводят параллели с аналогичными результатами классической теории игр. Кроме этого, основная масса проведенных исследований направлена на построение достаточных условий существования равновесий многокритериальных игр. Вне поля зрения казались необходимые, необходимые и достаточные условия существования равновесий как в общем случае игры, так и для отдельных классов игр. Также не исследованы свойства равновесий многокритериальных игр. Частично эти проблемы были исследованы в работах [16; 20].

Равновесия по Нэшу в обобщенных играх. В некооперативной игре в нормальной форме $(X_i, u_i; i \in N)$ каждый игрок оценивает ситуацию игры числовым значением своей функции выигрыша, а ситуацию смешанного расширения игры — математическим ожиданием своей функции выигрыша.

С точки зрения теории полезности в смешанном расширении игры каждый игрок знает свое отношение к риску. Это предположение можно поддать критике [13], поскольку: во-первых, в большинстве реальных социально-экономических, спортивных и военных ситуаций игроки могут сравнивать между собой лишь любые две различные ситуации по предпочтению друг к другу; во-вторых, даже в случае, когда игроки могут оценить ситуацию игры функциями полезности, то это еще не означает, что они оценивают ситуацию в смешанных стратегиях значением ожидаемой функции полезности и, тем более, об этом может быть неизвестно другим игрокам. Поэтому в некоторых случаях более адекватными моделями принятия решений в условиях конфликта, чем игры в нормальной форме $(X_i, u_i; i \in N$ с функциями выигрыша игроков являются обобщенные игры [13].

Общей некооперативной игрой называется [13] совокупность $\Gamma = (X_i, R_i;$ $i \in N$, где N — множество игроков; X_i — множество стратегий игрока $i \in N$; R_i — рефлексивное бинарное отношение, которое является отношением предпочтения игрока $i \in N$ на множестве ситуаций игры $X = \Pi_{i \in N} X_i$.

Впервые игры в такой обобщенной постановке были сформулированы еще в роботах К. Бержа [3], но наиболее полно они были исследованы в роботах Е. Яновской [13].

Согласно [13], ситуация x^* обобщенной игры называется равновесием, если для $\forall i \in N$, $\forall y_i \in X_i$, имеет место отношение $(y_i, x^*_{N \setminus i}) \bar{P}_i x^*$, где $P_i = R_i \setminus R_i^{-1}$ — отношение строгого предпочтения игрока $i \in N$, которое порождено отношением R_i.

В [13] также доказано, что достаточными условиями существования равновесий обобщенной игры с выпуклыми и компактными пространствами стратегий являются следующие $\forall i \in N$:

- множество $\{x \mid (y_i, x_{N \setminus i}) P_i x\}$ — открыто для $\forall y_i \in X_i$;
- множество $\{y_i \mid (y_i, x_{N \setminus i}) P_i x\}$ — выпукло для $\forall x \in X_i$;
- отношение P_i — не рефлексивно.

В [13] также исследовано смешанное расширение бескоалиционной обобщенной игры Γ в случаях: с в совершенстве упорядоченными, частично упорядоченными ситуациями и с произвольными асимметричными отношениями предпочтения игроков.

Обозначим M_i — множество вероятностных мер на X_i, $i \in N$.

Согласно [13] смешанным расширением игры Γ называется обобщенная бескоалиционная игра $\bar{\Gamma} = (M_i, R_i; i \in N)$, где R_i — бинарное рефлексивное отношение предпочтения игрока $i \in N$, которое определено на множестве ситуаций в смешанных стратегиях $M = \prod_{i \in N} M_i$.

В случае в совершенстве упорядоченных ситуаций исходной игры Γ (отношение $R_i = P_i \bigcup I_i$, где P_i — слабое упорядочение, I_i — эквивалентность, $i \in N$) рассматриваются $\tilde{P}_i$ и $\tilde{I}_i$ — аффинные транзитивные расширения соответственно отношений P_i и I_i, $i \in N$, на множество всех $\bar{M} \supset M = \prod_{i \in N} M_i$ борелевских вероятностных мер на $X = \prod_{i \in N} X_i$. Отношение $\tilde{R}_i = \tilde{P}_i \bigcup \tilde{I}_i$, $i \in N$. Смешанное расширение $\tilde{\Gamma} = (M_i, \tilde{R}_i; i \in N)$ игры Γ называется аффинным транзитивным расширением игры Γ.

В [13] доказано что, если в игре Γ множества стратегий X_i является компактными топологическими пространствами, отношения R_i, $i \in N$, непрерывны в топологии $X = \prod_{i \in N} X_i$ и пространство X сепарабельно в топологии произведения, то в игре всегда существует ситуация равновесия.

Исследован случай частично упорядоченных ситуаций исходной игры Γ. Обозначим U_i — множество односторонних [13] функций полезности по отношениям R_i, $i \in N$; Γ_H — бескоалиционную игру с функциями выигрыша $H_i \in U_i$; $CP(\tilde{\Gamma})$ — множество равновесий игры $\tilde{\Gamma}$. Тогда согласно [13] справедливо равенство

$$CP(\tilde{\Gamma}) = \bigcup_{H_I \in U_i,\, i \in N} CP(\Gamma_H).$$

Исследован случай асимметричных отношений предпочтения игроков.

В [13] рассматривается множество U_i кососимметричных функций интенсивности предпочтения для отношения P_i, которые отделены от нуля и ограничены по модулю, $i \in N$; произвольный набор $u = \{u_i\}_{i \in N}$, $u_i \in U_i$, $i \in N$, и отображение $T_u : M \to M$, где $\nu \in T_u(\mu)$, если $\forall i \in N$ ν_i будет оптимальной стратегией игрока $i \in N$ в антагонистической симметричной игре $(X_i, X_i, u_i((x_i, \mu_{N \setminus i}), (y_i, \mu_{N \setminus i})))$.

В [13] доказано, что множество ситуаций равновесий игры $\tilde{\Gamma}$ не пусто и задается объединением по всем $u = \{u_i\}_{i \in n}$, $u_i \in U_i$, $i \in N$, множеств неподвижных точек отображений T_u.

Таким образом, обобщенные игры представляют собой важное обобщение классических некооперативных игр. В авторской работе [15] исследования обобщенных игр были продолжены на случай, когда цель каждого игрока может задаваться множеством отношений предпочтения. В [18] рассмотрены и исследованы свойства отношения NE-предпочтения игроков.

2. Развитие концепции равновесия Нэша в некооперативных играх

Коалиционные равновесия. Во многих экономических, социально-экономических и других системах агентами (участниками) игры, которые действуют некооперативно, могут быть не отдельные игроки, а некоторые их группировки (коалиции), которые совместно выбирают коалиционные стратегии и имеют коалиционную цель — множество функций выигрыша членов коалиции.

Понятие некооперативного коалиционного равновесия (точка равновесия для множества игроков P) формализовано еще в монографии К. Бержа [3, с. 96] с целью определения и исследования свойств сильного равновесия по Нэшу и формализации основ кооперативных игр.

Некооперативное коалиционное равновесие как самостоятельное понятие исследовано в [5, с. 60–61]. Рассматривается игра G в нормальной форме $(X_i, u_i; i \in N)$ с разбиением $\{N(k)\}_{k \in K}$ множества игроков N игры G на коалиции, которые не пересекаются, то есть

$$\bigcup_{k\in K}(k) = N;\;\; N(k)\bigcap N(l) = \emptyset;\;\; k, l \in K;\;\; k \neq l.$$

Множество коалиционных стратегий обозначается $X_N(k) = \prod_{i\in N(k)} X_i$, а через $U_{N(k)}(x) = (u_i, (x))_{i\in N(k)}$ обозначается вектор выигрышей коалиции, который состоит из функций выигрыша членов коалиции $k \in K$.

Ситуация X^* называется коалиционным равновесием игры G, если $\forall\ k \in K$ или $\exists i \in N(k)$: $u_i(x^*) > u_i(x_{N(k)}, x^*_{N\setminus N(k)})$, или $u_i(x^*) = u_i(x_{N(k)}$, $X^*_{N\setminus N(k)})$ $\forall i \in N(k)$. Таким образом, ситуация x^* будет коалиционным равновесием, если для любой коалиции $k \in K$ стратегия $x^*_{N(k)}$ максимизирует по Парето вектор выигрыша $U_{N(k)}(x_{N(k)}, x^*_{N\setminus N(k)}) = (u_i(x_{N(k)}$, $x^*_{N\setminus N(k)}))_{i\in N(k)}$ коалиции на множестве коалиционных стратегий $X_{N(k)} = \prod_{i\in N(k)} X_i$.

Если зафиксировать коалиционную структуру, а на коалиции смотреть как на отдельных игроков (с векторными выигрышами), то получим бескоалиционную игру $(X_{N(k)}, U_{N(k)}; k \in K)$. Тогда коалиционное равновесие будет векторным равновесием этой многокритериальной игры. Вилкасом Э. также исследованы коалиционные равновесия в контексте равновесий в угрозах и контр-угрозах. По мнению автора [5, с. 61], понятие коалиционного равновесия исследовано мало и для него известен лишь очевидный результат: если в игре G множества стратегий игроков компактны, а функции выигрыша непрерывны, то существует коалиционное равновесие хотя бы для одной коалиции — всего множества игроков.

В авторских работах [21; 22; 24; 25; 27] понятия коалиционного равновесия обобщено применением слабой аксиомы Парето и на случай обобщенных игр в отношениях предпочтения игроков. Оно также используется для исследования свойств индивидуально-оптимальных равновесий.

Равновесия по Бержу. Применение концепции равновесия по Нэшу для решения реальных социально–экономических, политических конфликтов выявило ряд случаев, в которых ее применение приводило к парадоксальным результатам. Впервые на это обратил внимание К. Берж в [3] предложил новое понятие равновесия, согласно которому одна коалиция игроков может максимизировать функции выигрыша игроков другой коалиции.

Пусть $R \subset N$ — произвольное подмножество индексов функций выигрыша игроков, а $S \subset N$ — коалиция игроков. Ситуация $x^* \in X$ называется равновесием игры G в нормальной форме $(X_i, u_i; i \in N)$ для множества индексов R относительно коалиции S, если:

$$u_i(x^*) \geq u_i(x_S, x^*_{N\setminus S}), \quad \forall x_S \in X_S,\ \forall i \in R.$$

Это понятие не было развито К. Бержем далее и только спустя некоторое время к нему обратились авторы работ [11; 62]. Определенный интерес представляет вариант альтруистического равновесия по Бержу, который был предложен В. Жуковским [11, с. 119–124].

Ситуация x^* называется равновесием по Бержу игры G в нормальной форме $(X_i, u_i, i \in N)$, если

$$u_i(x^*) \geq u_i(x_i^*, x_{N\backslash i}), \quad \forall x_{N\backslash i} \in X_{N\backslash i},\ \forall i \in N,$$

т. е. в равновесии по Бержу дополняющая коалиция $N\backslash\{i\}$ каждого игрока $i \in N$ максимизирует его выигрыш (“один — за всех, все — за одного”).

На первый взгляд, такое поведение сложно назвать рациональным, но примеры из реальной жизни доказывают противоположное (альтруистические взгляды имеют место, например, в родственных отношениях, в христианских общинах и т.п.). Эти идеи позже были развиты в работе [62].

Это равновесие может использоваться также как альтернативное решение конфликта, когда равновесия по Нэшу отсутствуют или наоборот, когда их много. В равновесии по Бержу каждый игрок получает свой максимальный выигрыш, если ситуация благоприятна для него обязательством или готовностью других игроки выбирать стратегии, оптимальные для него.

В [11, с. 14–126] получены достаточные условия существования равновесия по Бержу. В предположении конечного количества игроков, компактности множеств стратегий игроков и непрерывности их функций выигрыша автором рассматривается вещественнозначная функция $\varphi : X \times \hat{X} \to R^1$, которая определена таким образом:

$$\varphi(x, \hat{y}) = \sum_{i \in N} (u_i(x_i, \hat{y}_{N\backslash\{i\}}) - u_i(x)),$$

где $\hat{y} = (x_{N\backslash\{i\}}, \ldots, x_{N\backslash\{n\}}) \in \hat{X} = \prod_{i \in N} X_{N\backslash i}$. В соответствии с определением этой функции справедливы неравенства

$$\max_{\hat{y} \in X} \varphi(x, \hat{y}) \geq 0, \quad \forall x \in X.$$

В [11, с. 124-126] показано, что игра имеет равновесие по Бержу тогда и только тогда, когда $\min_{x \in X} \max_{\hat{y} \in \hat{X}} \varphi(x, \hat{y}) = 0$.

В статье [62] рассматривается специальный случай равновесия Бержа, где использовано изменение индексов стратегий для нахождения полного множества равновесий. Для случая двух игроков это означает, что они просто обмениваются стратегиями. В работах [43; 49; 52] исследуются соответственно: теорема существования, применение в разностно-дифференциальных и биматричных играх. В работах [37–41; 46] применяется различный математический аппарат для доказательства существования равновесия Бержа и его сравнения с равновесием Нэша. Например, теорема о неподвижной точке применяется в [38], а теория оптимизации в [46]. Авторы статьи [56] изучали основные свойства строгого равновесия Бержа и доказали теорему существования, которая основывается на “Ky Fan неравенствах”. Они также впервые предложили алгоритм поиска равновесий по Бержу. В монографии [11] исследованы свойства равновесий по Бержу, проведен их сравнительный анализ с равновесиями Нэша для дифференциальных линейно-квадратичных игр.

В работах [37; 40; 41] обобщается равновесие по Бержу.

Пусть $R = \{R(k)\}_{k\in K}$ — разбиение N на множества, которые не пересекаются, т. е. $\bigcup_{k\in K} R(k) = N$; $R(m)\bigcap R(l) = \emptyset$; $m, l, \in K$; $m \neq 1$. Пусть также $S = \{S(k)\}_{k\in K}$ — произвольный набор попарно различных подмножеств N, которые покрывают все множество игроков, т. е. $\bigcup_{k\in K} S(K) = N$; $S(m) \neq S(l)$; $m, l, \in K$; $m \neq l$.

Ситуация $x^* \in X$ называется обобщенным равновесием по Бержу игры G в нормальной форме $(X_i, u_i, i \in N)$ для разбиения относительно набора подмножеств S, если $\forall x_{S(k)} \in X_{S(k)}$, $\forall i \in R(k)$, $\forall k \in K$ $u_i(x^*) \geq$ $\geq u_i(x_{S(k)}, x^*_{N\setminus S(k)})$.

Обратим внимание, что в этом определении имеется два набора подмножеств. Набор $R = \{R(k)\}_{k\in K}$ является разбиением множества функций выигрыша игроков на подмножества, которые не пересекаются, а набор $S = \{S(k)\}_{k\in K}$ является совокупностью коалиций игроков, которые совместно избирают свои коалиционные стратегии. Эти коалиции могут пересекаться, но должны быть попарно различными. Таким образом, коалиции игроков, которые могут пересекаться между собой, преследуют коалиционные цели, которые не обязательно пересекаются и могут состоять из функций выигрыша, даже им не принадлежащих.

Определение обобщенного равновесия по Бержу дает понятие равновесия в очень широком смысле. В частности, для $R(i) = \{i\}$, $S(i) = \{i\}$, $i \in N$ получим определение равновесия по Нэшу; для $R(i) = \{i\}$, $S(i) = N\setminus\{i\}$, $i \in N$ получим определение альтруистического равновесия по Бержу в смысле В. Жуковского; для $R(k) = S(k)$, $k \in K$, $k \in K$ при условиях: $\bigcup_{k\in K} S(k) = N$; $S(k)\bigcap S(l) = \emptyset$; $k, l \in K$; $k \neq l$ получим определение коалиционного равновесия. В определенном смысле этот принцип оптимальности можно считать попыткой параметризации равновесий посредством совокупности подмножеств $R(k)$, $S(k)$, $k \in K$.

В [37] К. Абало и М. Кострева формулируют также теоремы существования обобщенного равновесия по Бержу. Следует отметить, что эти теоремы основаны на более ранних результатах, которые принадлежат М. Раджифу [62] и формализуют условия существования равновесия по Бержу в смысле В. Жуковского [11].

После исследования проблемы существования равновесий по Бержу, авторами работы [60] были сделаны выводы, что предположение теорем К. Абало и М. Кострева, не являются достаточными для существования равновесия по Бержу, как в смысле В. Жуковского, так и в общем виде. Такие же замечания были сделаны и применительно к теореме М. Раджифа. Причиной проблем, связанных с равновесием по Бержу, как в смысле В. Жуковского, так и К. Абало и М. Кострева, является зависимость множеств коалиционных стратегий игроков. В [60] была приведена исправленная версия одной из теорем К. Абало и Г. Кострева. Для ее формулировки вводится такое понятие.

Пусть в игре G в нормальной форме $(X_i, u_i, i \in N)$ множества коалиций S_i, $i \in N$ являются непустыми, попарно различными ($S_i \neq S_j$, $i, j \in N$, $i \neq j$) и покрывают все множество игроков ($N = \bigcup_{i\in N} S_i$). Множество

$S = \{S_i\}_{i \in N}$ называется S — системой игры G, если для всех $i \in N$ множество $\arg\max_{y_{S_i} \in X_{S_i}} u_i(y_{S_i}, x_{N \setminus S_i}) \neq \emptyset$.

Пусть в игре G в нормальной форме $(X_i, u_i, i \in N)$ с конечным или бесконечным множеством игроков N, множества стратегий X_i являются компактами некоторых метрических топологических векторных пространств $E_i, i \in N$. Пусть $S = \{S_i\}_{i \in N}$ является S–системой игры G и для всех $i \in N$ выполняются условия:

1) $u_i(x)$ — непрерывна на множестве ситуаций X;
2) $\forall x \in X \quad \exists y \in X$ такой что

$$\arg \max_{z_{S_i} \in X_{S_i}} u_i(z_S, x_{N \setminus S_I}) = y_{S_i}, \quad \forall i \in N.$$

Тогда, как это показано в [60], множество равновесий по Бержу будет непустым.

Другими словами, условие 2) означает, что $\forall x \in X$ и $\forall i, j \in N$ соответствующие компоненты векторов

$$\arg \max_{z_{S_i} \in X_{S_i}} u_i(z_S, x_{N \setminus S_I}) \text{ и } \arg \max_{z_{S_j} \in X_{S_j}} u_j(z_{S_j}, x_{N \setminus S_j})$$

должны совпадать.

Вполне понятно, что условие 2) не только не конструктивно, но и показывает, что задача существования равновесия по Бержу гораздо более сложная, чем проблема существования равновесий по Нэшу. Более того, равновесие по Бержу в играх более, чем трех лиц, существует в исключительных случаях. Поэтому попытка использования "каркаса" некооперативной игры для взаимозависимых стратегий игроков является неудачной.

С другой стороны, оригинальной идеей, которая заложена в основу этого равновесия, является возможность моделирования конфликтных ситуаций, в которых в области интересов одних игроков могут оказаться цели других игроков. Идея учета игроками не только своих интересов, но и целей других игроков перекликается с концепцией индивидуально-оптимальных равновесий и служит для нее дополнительной аргументацией актуальности.

Индивидуально–оптимальные равновесия. Принцип индивидуальной оптимальности, развитый в работах [17; 19; 21–23; 25], предоставляет возможность каждому игроку выбирать свои стратегии индивидуально (некооперативно), но при этом учитывать интересы всех других игроков (компромисс ради разрешения конфликта). Этот принцип обоснован в, так называемых, одноцелевых играх, где у всех игроков цель одна, но она характеризуется для каждого игрока своей функцией выигрыша. В идеале, эта цель, заключается в выборе игроками своих стратегий так, чтобы сложилась наиболее предпочтительная ситуация для всех игроков. Поскольку такие ситуации могут не существовать, то игроки согласны на компромисс. В силу того, что игроки действуют некооперативно, каждый видит этот компромисс по своему, что приводит к конфликту между ними.

Одноцелевая игра принципиально отличается от классической многоцелевой игры, которая имеет такую же самую нормальную форму, но характеризуется тем, что каждый игрок независимо от того, действует он кооперативно с другими игроками, или нет, имеет свою личную цель. Одноцелевая некооперативная игра G также принципиально отличается от многокритериальной задачи принятия решений, которая может иметь такую же нормальную форму, но специфична коллективным поведением игроков.

Принцип индивидуальной оптимальности базируется на специальном отношении доминирования по Нэшу.

Ситуация y находится в отношении сильного *NE–доминирования* игрока $i \in N$ к ситуации x и обозначается это $y \overset{NE(i)}{\succ\succ} x$, если $y = (y_i, x_{N\setminus i})$ и $u_j(y_i, x_{N\setminus i}) > u_j(x)$, $\forall j \in N$.

Ситуация x^* называется слабым индивидуально-оптимальным равновесием, а множество этих равновесий обозначается $WIOE$, если не существует такого игрока $i \in N$, и ситуации $y \in N$, которая бы сильно доминировала по Нэшу x^*, т. е. $\forall i \in N \quad \nexists y \in X : y \overset{NE(i)}{\succ\succ} x^*$.

Применение слабых индивидуально-оптимальных равновесий мотивируется следующим сценарием одноцелевой игры. Игроки заключают необязательное соглашение (игроки прислушиваются к научно обоснованным рекомендациям компетентных лиц) о том, что они будут придерживаться ситуации $x^* = (x_i^*)_{i\in N}$. Потом они независимо один от другого принимают решение о выборе своих стратегий. В том и только в том случае, когда основой соглашения будет слабое индивидуально-оптимальное равновесие, изменение любым игроком $i \in N$ согласованной с другими игроками стратегии x_i^* на другую, всегда приведет к ситуации, которая не будет лучше x^* хотя бы для одного игрока. Поскольку эта ситуация может не быть наилучшей для всех игроков вместе, то она является компромиссом. Поэтому цель игрока $i \in N$, которая заключается в максимизации всех функций выигрыша $u_j(x)$, $j \in N$ может быть не удовлетворена и достигнутый в ходе предыдущих переговоров компромисс разрушен.

В [21; 22; 24; 25] установлены условия индивидуальной-оптимальности для различных классов игр.

Как показано в [22], множество индивидуально-оптимальных равновесий содержит в себе множества: равновесий по Нэшу (NE), коалиционных равновесий ($WKNE_Q$) и оптимальных по Слейтеру (SO) ситуаций игры, равновесий по Бержу (BE), т. е. $NE, WKNE_Q, SO, BE \subseteq WIOE$. Это свойство дает возможность построить оценку максимальной стабильности индивидуально-оптимального равновесия:

$$\hat{\mu}^{\max}(x) = \max_{\mu} \sum_{i\in N} \mu_i^i;$$

$$\min_{j\in N}(u_i(y_i, x) - \mu_i^j S) \le \min_{j\in N}(u_j(x) - \mu_i^j S), \;\; \forall y_i \in X_i, \;\; i \in N;$$

$$\mu_i^j \ge 0, \;\; i, j \in N; \;\; \sum_{j\in N} \mu_i^j = 1, \;\; i \in N,$$

где $S = \sum_{j\in N} \sup_{x\in X} \; u_j(x)$.

Оценка максимальной стабильности ситуации $x \in WIOE$ характеризует основные типы равновесий некооперативных игр следующим образом: $x \in NE \Leftrightarrow \hat{\mu}^{\max}(x) = n; x \in KNE \Rightarrow \hat{\mu}^{\max}(x) = |K|$, где $|K|$ — количество коалиций в игре G; $x \in SO \Rightarrow \hat{\mu}^{\max}(x) = 1$; $x \in BE \Rightarrow \hat{\mu}^{\max}(x) = 0$. Поэтому одним из подходов к выбору конкретного индивидуально-оптимального равновесия, как основы соглашения между игроками, может быть выбор наиболее стабильного из них по критерию $\hat{\mu}^{\max}(x) \to \max$. Вполне понятно, что в случае существования равновесий по Нэшу, они и будут иметь максимальные оценки стабильности. Поэтому этот подход актуален в случае их отсутствия в игре.

В [28] рассматривается еще один подход к выбору индивидуально-оптимальных равновесий. Оказывается, что индивидуально-оптимальное равновесие x будет стабильным для любого игрока $i \in N$ по функции полезности $\nu_i(x, \xi_i) = \min_{j\in N}(u_j(x) - \xi_i^j)$, $\xi_i \in L_i = \{\xi_i = (\xi_i^j)_{j\in N} \in E^n \mid \sum_{j\in N} \xi_i^j = 0\}$.

Ситуацию $\hat{x} \in X$ называется равновесием в предпочтениях игроков игры G, если существует вектор параметров $\hat{\xi} = (\hat{\xi}_i)_{i\in N} \in L = \prod_{i\in N} L_i$, для которого выполняются неравенства:

$$\nu_i(\hat{x}, \hat{\xi}_i) \geq \nu_i((x_i, \hat{x}_{N\setminus i}), \xi_i) \quad \forall x_i \in X_i,\ \forall \xi_i = (\xi_i^j)_{j\in N} \in L_i,\ i \in N.$$

Обозначим RE — множество равновесий в предпочтениях игроков.

Если $\hat{x} \in RE$, то для каждого игрока $i \in N$ существует $\hat{\xi}_i \in L_i$, что ему не выгодно с точки зрения функции совокупной полезности ν_i изменять одновременно свою стратегию $\hat{x}_i$ и вектор параметров $\hat{\xi}_i$, который характеризует его предпочтение на множестве функций выигрыша других игроков.

В [28] установлены: условия равновесия в предпочтениях игроков, условия согласованности ($\hat{\xi}_i^j = \hat{\xi}_j^j, i, j \in N$) компонент векторов параметров $\hat{\xi}_i = (\hat{\xi}_i^j)_{j\in N}$, $i \in N$ равенство между собой значений функций совокупной полезности игроков ($\nu(\hat{x}) = \nu_i(\hat{x}, \hat{\xi}_i) = \sum_{j\in N} u_j(\hat{x})/n, i \in N$).

Установлены условия существования равновесий в предпочтениях.

Пусть в игре G множества стратегий игроков X_i, $i \in N$, — компактны, а функции выигрыша u_i, $i \in N$, — полунепрерывны сверху. Тогда множество RE равновесий в предпочтениях игры G будет непустым компактом.

Равновесия в условиях неопределенности. Игры в условиях неопределенности безусловно представляют собой актуальную и практически полезную постановку задачи принятия решений в условиях конфликта.

В [12, с. 180] рассматривается бескоалиционная игра UG в условиях неопределенности в нормальной форме $(Y; X_i, u_i; i \in N)$ где:

- $N = \{1, 2, \ldots, n\}$ — конечное множество из $n \geq 2$ игроков;
- Y — множество состояний внешней среды;
- X_i — множество стратегий (элементы множества стратегий называют стратегиями и обозначают x_i) игрока $i \in N$;

– $u_i(x,y)$ — функция выигрыша игрока $i \in N$ которая определена на декартовом произведении множества ситуаций игры $X = \prod_{i\in N} X_i$ (элемент множества X называется ситуацией игры) и множества Y состояний внешней среды, и максимизируются.

Под решением игры в условиях неопределенности автор понимает пару, которая состоит из ситуации и векторной гарантии [12, с. 180], которую обеспечивают себе игроки выбором своих стратегий.

Векторной гарантией игры UG в ситуации $x^* \in X$ называется [12, с. 181] вектор $U^* = (u_i^*)_{i\in N}$ такой, что существует состояние $y^* \in Y$ внешней среды, при котором $u_i^* = u_i(x^*, y^*)$, $i \in N$, и $\forall y \in Y$ будет несовместной система неравенств: $u_i^* > u_i(x^*, y), i \in N$.

Различные виды векторных гарантий игры UG, их свойства и методы построения изложены в [12, с. 91–104].

В частности, для векторных гарантий по Слейтеру [11, с. 233] пара (x^*, U^*) называется равновесием Нэша-Слейтера игры UG, если существует состояние внешней среды $y^* \in Y$, в котором $u_i^* = u_i(x^*, y^*), i \in N$ и выполняются неравенства: $u_i(x_i, x^*_{N\setminus i}, y^*) \le u_i(x^*, y^*)$, $\forall x_i \in X_i$, $i \in N$; также $\forall y \in Y$ будет несовместной система неравенств: $u_i^* > u_i(x^*, y)$, $i \in N$.

В [11, с. 234–241] рассмотрены свойства и достаточные условия существования равновесия Нэша-Слейтера.

Согласно свойству индивидуальной рациональности, если в игре UG существует равновесие Нэша-Слейтера (x^*, U^*) и максмины:

$$\bar{u}_i = \max_{x_i\in X_i} \min_{(x_{N\setminus i}, y)\in X_{N\setminus u}\times Y} u_i(x,y), \;\; i \in N,$$

то $u_i(x^*, y^*) \ge \bar{u}_i$, $i \in N$.

Согласно свойству компактности, если в игре UG множества X_i, $i \in N$, и Y компактны, а функции выигрыша $u_i(x,y)$, $i \in N$ непрерывны на $X \times Y$, то множество всех равновесий Нэша-Слейтера будет компактным (может быть и пустым) на $X \times Y$.

Согласно свойству инвариантности относительно аффинного преобразования, если $\alpha_i, \lambda_i > 0$, $i \in N$, — константы, то игры UG и $(Y; X_i, \alpha_i, +\lambda_i u_i;$ $i \in N)$ имеют одно и то же множество Нэша-Слейтера.

Пусть в игре UG множества X_i, $i \in N$, и Y является выпуклыми компактами, а функции выигрыша $u_i(x,y)$, $i \in N$ непрерывны на $X \times Y$, выпуклы по y при $\forall x \in X$ и вогнуты по x_i при фиксированных $(x_{N\setminus i}, y) \in X_{N\setminus i} \times Y$, $i \in N$. Согласно этими условиями в [11, с. 241] доказано, что в игре UG существует равновесие Нэша-Слейтера.

Недостатком равновесий Нэша-Слейтера является внутренняя неустойчивость их множества [12, с. 214] (могут существовать по крайней мере две ситуации x' и x'', которые образуют равновесия Нэша-Слейтера, причем: $u_i(x') > u_i(x'')$, $i \in N$.

Аналогичные результаты получены для других типов векторных гарантий: по Парето, по Борвейну, по Джеоффриону, по конусу A.

В [12, с. 217] рассматривается понятие гарантированного по риску решения.

Пара (x^L, Φ^L) называется гарантированным по L равновесием игры UG, если существует состояние внешней среды $y^L \in Y$, при котором $\Phi^L = \Phi(x^L, y^L) = (\Phi_i(x^L, y^L))_{i \in n}$, и выполняются следующие условия:

- ситуация x^L является равновесием по Нэшу в игре $(X_i, u_i(x, y^L); i \in N)$;
- состояние внешней среды $y^L \in Y$ является L-минимальным [12, с. 217] в многокритериальной задаче $(Y, \Phi_i(x^L, y); i \in N)$, где функция риска игрока $i \in N$ имеет вид

$$\Phi_i(x, y) = \max_{z_i \in X_i} u_i(z_i, x_{N \setminus i}, y) - u_i(x, y).$$

При этом вектор Φ^L называется L-гарантированным риском игры UG, а x^L — ситуацией, которая гарантирует этот риск.

Это определение формализует пять различных гарантированных по риску решений игры UG [12, с. 187–192]:

- при $L = A$ — гарантированное по конусу A равновесие;
- при $L = G$ — гарантированное по Джеоффриону равновесие;
- при $L = B$ — гарантированное по Борвейну равновесие;
- при $L = P$ — гарантированное по Парето равновесие;
- при $L = S$ — гарантированное по Слейтеру равновесие.

В [12, с. 219–227] рассмотрены основные свойства гарантированных по $L = A, G, B, P, S$ равновесий игры UG: индивидуальная рациональность, компактность, инвариантность относительно аффинных преобразований. Показана внутренняя устойчивость множества гарантированных по L равновесий игры UG. Условия существования гарантированных по L равновесий игры UG рассмотрены в [12, с. 231–245].

В [10] рассмотрен случай игры в условиях неопределенности, когда игроки могут использовать информацию о состояниях внешней среды, которые реализовались. Такая "дискриминация" неопределенности позволяет игроку формировать свои стратегии, как реакцию на появление того или другого состояния внешней среды. Это позволяет сохранить стабильность ситуации равновесия относительно отклонений отдельных игроков при реализации конкретного состояния внешней среды.

Рассматривается бескоалиционная игра $\overline{UG}$ в условиях неопределенности в нормальной форме $(Y; B(Y, X_i), u_i; i \in N)$, где:

- $B(Y, X_i)$ — множество контр-стратегий (элементы этого множества называют контр-стратегиями и обозначают $x_i(y)$) игрока $i \in N$;
- N; Y; X_i, $u_i(x, y)$, $i \in N$, имеют такой же смысл, как и в игре UG.

Тройка $(x^L(y), U^L, \Phi^L)$ называется L-гарантированным по выигрышу и риску равновесием игры $\overline{UG}$ в контр-стратегиях [10, с. 267] ($L = A$, G, B, P, S), если существует состояние внешней среды $y^L \in Y$, в котором имеют место равенства:

$$U^L = U(x^L(y^L), y^L) = (u_i(x^L(y^L), y^L))_{i \in N},$$

$$\Phi^L = \Phi(x^L(y^L), y^L) = (\Phi_i(x^L(y^L), y^L))_{i \in N}$$

и выполняются следующие условия:

– ситуация в контр-стратегиях $x^L(y)$ является равновесием по Нэшу для $\forall y \in Y$ в игре $(B(Y, X_i), u_i(x(y), y); i \in N)$;
– состояние внешней среды $y^L \in Y$ является L-максимальным в $2n$–критериальной задаче $(Y; u_i(x^L(y), y), -\Phi_i(x^L(y), y), i \in N)$, где функция риска игрока $i \in N$ имеет вид

$$\Phi_i(x, y) = \max_{z_i \in X_i} u_i(z_i, x_{N \setminus i}, y) - u_i(x, y).$$

При этом вектор U^L называется L-гарантированным выигрышем игры UG, а Φ^L — L-гарантированным векторным риском.

В [10, с. 269-276] рассматриваются свойства и условия существования L-гарантированных по выигрышу и риску равновесий игры $\overline{UG}$ в контр-стратегиях.

Следует заметить, что сценарий стабильного соглашения между игроками, который базируется на L-гарантированному по выигрышу и риску равновесию игры $\overline{UG}$ в контр-стратегиях имеет в виду выбор игроками контр-стратегий (реакций на появление того или другого состояния внешней среды), что не снимает вопрос выбора стратегий игроками в условиях неопределенности.

В работе [26] введенное понятие, исследованные свойства и условия существования "виртуального равновесия" игры в условиях неопределенности, которое не обязательно должно быть равновесием по Нэшу при некотором состоянии внешней среды. В таком равновесии неопределенность является гарантией стабильности соглашения между игроками.

Заключение

Рассмотренные выше направления развития концепции равновесия по Нэшу в некооперативных играх показывают, что за последние годы она интенсивно развивается как в сторону расширении сферы ее применения и обобщения (многокритериальные игры, игры, заданные отношениями предпочтения игроков), так и в глубь, в сторону более адекватного описания реальных конфликтов (коалиционные равновесия, равновесия по Бержу и индивидуально-оптимальные равновесия).

Литература

1. Антипин А. С. Равновесное программирование: методы градиентного типа / А. С. Антипин // Автоматика и телемеханика. — 1997. — № 8. — С. 125–137.
2. Антипин А. С. Равновесное программирование: проксимальные методы / А. С. Антипин // ЖВМиМФ. — 1997. — 37, № 11. — С. 1327–1339.
3. Берж К. Общая теория игр нескольких лиц / К. Берж; пер. с франц. И. В. Соловьева; под ред. В. Ф. Колчина. — М.: Физматгиз, 1961. — 126 с.
4. Блекуэлл Д. Аналог теоремы о минимаксе для векторных выигрышей / Д. Блекуэлл // Матричные игры: [сб. переводов / под ред. Н. Н. Воробьева]. — М.: Физматгиз, 1961. — С. 156–166.
5. Вилкас Э. Й. Оптимальность в играх и решениях / Э. Й. Вилкас. — М.: Наука, 1990. — 256 с.

6. Гликсберг И. Л. Дальнейшее обобщение теоремы Какутани с приложением к ситуациям равновесия в смысле Нэша. // Бесконечные антагонистические игры: [сб. переводов / под ред. Н. Н. Воробьева]. — М.: Физматгиз, 1963. — С. 47–53.
7. Ерошов С. А. Иерархические игры двух лиц с векторной функцией выигрыша / С. А. Ерошов // Некоторые вопросы прикладной математики и программного обеспечения ЭВМ. — М.: Изд-во МГУ, 1982. — С. 91–93.
8. Ерошов С. А. Модель Гросса с векторным критерием / С. А. Ерошов // Математические методы в исследовании операций. — М.: Изд-во МГУ, 1981. — С. 93–103.
9. Ерошов С. А. Об играх двух лиц с полной информацией и векторной функцией выигрышей / С. А. Ерошов // Прикл. матем. и матем. обеспечение ЭВМ. — М.: Изд-во МГУ, 1979. — С. 99–105.
10. Жуковский В. И. Конфликты и риски / В. И. Жуковский; под ред. В. С. Молоствова. — М.: ГУП МО "Орехово-Зуевская типография", 2007. — 456 с.
11. Жуковский В. И. Линейно-квадратичные дифференциальные игры / В. И. Жуковский, А. А. Чикрий; отв. ред. В. А. Плотников. — К.: Наукова думка, 1994. — 320 с.
12. Жуковский В. И. Риск в многокритериальных и конфликтных системах при неопределённости : монография / В. И. Жуковский, Л. В. Жуковская; под ред. В. С. Молоствова. — М.: УРСС; Едиториал УРСС, 2004. — 272 с.
13. Кирута А. Я. Оптимальный выбор распределений в сложных социально-экономических задачах (вероятностный подход) / А. Я. Кирута, А. М. Рубинов, Е. Б. Яновская; отв. ред. Н. Н. Воробьев. — Л. : Наука, 1980. — 167 с.
14. Льюс Р. Л. Игры и решения. Введение и критический обзор / Р. Л. Льюс, Х. Райфа; перев. с англ. И. В. Соловьева; под ред. Д. Б. Юдина; предисл. А. А. Ляпунова. — М.: Изд-во иностр. лит., 1961. — 642 с.
15. Мащенко С. О. Багатоцільові ігри у відношеннях переваги гравців / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2008. – № 4. — С. 135–140.
16. Мащенко С. О. Векторні рівновги у змішаних стратегіях / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2006. — № 1. — С. 171–177.
17. Мащенко С. О. Вибір індивідуально-оптимальних рівноваг за критерієм стабільності / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2009. — № 4. — С. 113–118.
18. Мащенко С. О. Відношення NE-переваги і домінування в некооперативних іграх та їх властивості / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2009. — № 1. — С. 115–120.
19. Мащенко С. О. Достатні умови індивідуальної оптимальності / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2009. — № 2. — С. 119–124.
20. Мащенко С. О. Загальні умови векторної рівноваги за Нешем / С. О. Мащенко, О. Г. Павлюченко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2004. — № 4. — С. 212–216.
21. Мащенко С. О. Индивидуально-оптимальные равновесия некооперативных игр в отношениях предпочтения / С. О. Мащенко // Кибернетика и системный анализ. — 2009. — № 1. — С. 171–179.

22. Мащенко С. О. Исследование стабильности равновесий на основе принципа индивидуальной оптимальности / С. О. Мащенко // Кибернетика и системный анализ. — 2007. — № 4. — С. 162–169.
23. Мащенко С. О. Індивідуальна раціональність індивідуально-оптимальних рівноваг / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2010. — № 2. — С. 124–129.
24. Мащенко С. О. Індивідуально-оптимальні рівноваги в некооперативних опуклих іграх / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2008. — № 2. — С. 105–110.
25. Мащенко С. О. Локальні умови слабкої індивідуальної оптимальності рівноваг / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2008. — № 3. — С. 142–147.
26. Мащенко С. О. Один подход к равновесиям в играх в условиях неопределенности / С. О. Мащенко // XIII-th International Conference KDS–2007 "Knowledge — Dialogue — Solution": June, 2007, Varna, Bulgaria. Proceedings. — Sofia: ITHEA. — 2007. — 1. — P. 129–137.
27. Мащенко С. О. Слабкі індивідуально-оптимальні рівноваги / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2006. — № 2. — С. 216 –223.
28. Мащенко С. О. Стабільні за перевагами рівноваги в одноцільових некооперативних іграх / С. О. Мащенко // Вісник Київського університету. Серія: фіз.-мат. науки. — 2009. — № 3. — С. 152–157.
29. Мулен Э. Теория игр с примерами из математической экономики / Э. Мулен; пер с франц. О. Р. Меньшиковой и И. С. Меньшикова; под ред. Н. С. Кукушкина. — М.: Мир, 1985. — 200 с.
30. Нейман Дж. фон. Теория игр и экономическое поведение / Дж. фон Нейман, О. Моргенштерн; пер. с англ.; под ред. и с доб. Н. Н. Воробьева. — М.: Наука, 1970. — 708 с.
31. Никайдо Х. Заметки о бескоалиционных выпуклых играх / Х. Никайдо, К. Йосода // Бесконечные антагонистические игры: [сб. переводов / под ред. Н. Н. Воробьева]. — М.: Физматгиз, 1963. — С. 449 — 458.
32. Нэш Дж. Бескоалиционные игры / Дж. Нэш // Матричные игры : [сб. переводов / под ред. Н. Н. Воробьева]. — М.: Физматгиз, 1961. — С. 205–221.
33. Партхасаратхи Т. Некоторые вопросы теории игр двух лиц / Т. Партхасаратхи, Т. Радхаван ; пер. с англ. В. К. Доманского, Г. Н. Дюбиина и А. Н. Ляпунова; под. ред. и с ком. Е. Б. Яновской. — М.: Мир, 1974. — 295 с.
34. Подиновский В. В. Парето-оптимальные решения многокритериальных задач: монография / В. В. Подиновский, В. Д. Ногин. — 2-е изд., испр. и доп. — М.: Физматлит, 2007 (Чебоксары). — 255 с.
35. Равновесия в динамических системах и их приложения / С. Жаркынбаев, В. И. Жуковский, Б. А. Омарова, Б. Д. Шарипова; под рец. д. т. н., проф. Т. Н. Биярова. — Алматы, 2006. — 272 с.
36. Харшаньи Дж. Общая теория выбора равновесий в играх / Джон Харшаньи, Рейнхард Зельтен; пер. с англ. под ред. Н. А. Зенкевича. — СПб.: Экономическая школа, 2001. — 424 с.
37. Abalo K. Y. Berge equilibria: some recent results from fixed–point theorems / K. Y. Abalo, M. M. Kostreva // Appl. Math. and Comp. — 2005. — 169 . — P. 624–638.

38. Abalo K. Y. Equi-well-posed games / K. Y. Abalo, M. M. Kostreva // J. Optim. Theory Appl. — 1996. — 89, № 1. — P. 89–99.
39. Abalo K. Y. Fixed points, Nash games and their organizations / K. Y. Abalo, M. M. Kostreva // Topol. Meth. Nonlinear Anal. — 1994. — № 8. — P. 205–215.
40. Abalo K. Y. Berge equilibria: some recent results from fixed–point theorems / K. Y. Abalo, M. M. Kostreva // Appl. Math. and Comp. — 2005. — 169 . — P. 624–638.
41. Abalo K. Y. Some existence theorems of Nash and Berge equilibria / K. Y. Abalo, M. M. Kostreva // Appl. Math. Letters. — 2004. — 17. — P. 569–573.
42. Arino O. A Fixed point theorem for sequentially continuous mappings with applications to ordinary differential equations / O. Arino, S. Gautier, G. P. Penot // Funkc. Ekvacioj. — 1984. — 27. — P. 273–279.
43. Boribekova K. A. Equilibrium of Berge in one differential-difference game, multicriteria dynamical problems under uncertainty / K. A. Boribekova, F. Jarkynbayev // Collection of Scientific Works. — Orekhovo-Zuevo, 1991. — P. 83–86.
44. Borm P. The structure of the set of equilibria for two person multicriteria games / P. Borm, D. Vermeulen, M. Voorneveld // Eur. J. Oper. Res. — 2003. — 148. — P. 489–493.
45. Born P. Pareto-equilibrium in multiobjective games / P. Born, S. Tijs, J. Van der Aarssen // Meth. and Oper. Res. — 1988. — 60. — P. 302–312.
46. Browder F.E. The fxed point theory of multi–valued mappings in topological vector spaces / F.E. Browder // Math. Annal. — 1968. — 177. — P. 283–301.
47. Corly H. N. Games with vector payoff / H. N. Corly // J. Optimiz. Theory and Appl. — 1985. — 47. — P. 491–498.
48. Cubiotti P. Existence of generalized Pareto equilibria for constrained multiobjective games / P. Cubiotti // Intern. Game Theory Review. — 2000. — 2, № 4. — P. 329–344.
49. Dinovsky V. Existence d'un point d'equilibre de Berge, multicriteria dynamical problems under uncertainty / V. Dinovsky // Collection of Scientifc Works. — Orekhovo–Zuevo, 1991. — P. 75–77.
50. Fan K. Fixed point and minimax theorems in locally convex linear spaces / K. Fan // Proc. Nat. Acad. Sci. U.S.A. — 1952. — 38. — P. 121–126.
51. Fernandez F. R. Multicriteria goal games / F. R. Fernandez, L. Monroy, J. Puerto // J. Opt. Theory Appl. — 1998. — 99, № 2. — P. 403–421.
52. Gintchev I. A. Method to obtain Berge equilibrium in bi–matrix games, multicriteria dynamical problems under uncertainty / I. A. Gintchev // Collection of Scientifc Works. — Orekhovo-Zuevo, 1991. — P. 78–82.
53. Kakutani S. A generalization of Brower's fixed point theorem / S. Kakutani // Duke Math. J. —1941. — 8, № 3. — P. 457–459.
54. Knaster B. Ein beweis des fixpunktsatzes fur n–dimensionale simplexe / B. Knaster, C. Kuratowski, S. Masurkievicz // Found Math. — 1929. — 14. — P. 132–137.
55. Krieger T. On Pareto equilibria in vector–valued extensive form games / T. Krieger // Math. Meth. Oper. Res. — 2003. — 58. — P. 449–458.
56. Larbani M. Sur l'equilibre fort selon Berge / M. Larbani, R. Nessah // RAIRO Oper. Res. — 2001. — 35. — P. 439–451.
57. Lemke C. E. Equlibrium points of bimatrix games / C. E. Lemke, J. J. Houson // J. Soc. Idust. Appl. Math. — 1964. — №12. — P. 412–423.

58. Lozovanu D. Multiobjective games and determining Pareto–Nash equilibria / Lozovanu D., Solomon D., Zelikowsky A. // Bul. Acad. de Stiinte a Rep. Moldova "Matematica" — 2005. — 49, № 3. — P. 115–122.
59. Nash J. F. Equilibrium points in n–person games / J. F. Nash // Proc. Nat. Acad. Sci. U.S.A. — 1950. — 36. — P. 48–49.
60. Nessah R. A note on Berge equilibrium / R. Nessah, M. Larbani, T. Tazdait // Appl. Math. Letters. — 2007. — 20. — P. 926–932.
61. Nieuwenhuis J.'W. Some minimax theorems in vector–valued functions / J. W. Nieuwenhuis // J. Optimiz. Theory and Appl. — 1983. — 40. — P. 463–475.
62. Radjef M. S. Sur l'existence d'un equilibre de Berge pour un Jeu diferentiel a "N" personnes / M. S. Radjef // Cahiers Math. Universite d'Oran. — 1988. — № 1. — P. 89–93.
63. Roemer J. Games with vector-valued payoffs and their application to competition between organizations / J. Roemer // Econ. Bull. — 2005. — 3, № 16. — P. 1–13.
64. Shapley L. S. Equilibrium points in games with vector playoff / L. S. Shapley // Naval Res. Log. Quart. — 1959. — 6. — P. 57–61.
65. Van Megen F. A preference concept for multicriteria game / F. Van Megen, P. Born, S. Tijs // Math. Meth. of Oper. Res. — 1999. — 49, № 3. — P. 401–412.
66. Wang S. An Existence theorem of a Pareto equilibrium / S. Wang. // Appl. Math. Lett. — 1991. — 4, № 3. — P. 61–63.
67. Wang S. Existence of a Pareto equilibrium / S. Wang // J. Optim. Theory Appl. — 1993. — 79. — P. 373–384.
68. Yang H. Essential components of the set of weakly Pareto–Nash equilibrium points / H. Yang, J. Yu // Appl. Math. Lett. — 2002. — 15. — P. 552–560.
69. Zeleny M. Games with multiple playoffs / M. Zeleny. // Intern. J. of Game Theory. — 4. — P. 179–191.
70. Zhao J. The equilibrium of a multiple objective game / J. Zhao // Int. J. of Game Theory. — 1991. — 20. — P. 171–182.

Факультет кибернетики, Киевский национальный университет имени Тараса Шевченко, ул. Владимирская, 64, Киев, 01601, Украина.

Поступила 10.11.2011

Журнал обчислювальної та прикладної математики | 2012, №1(107) | Journal of Computational & Applied Mathematics

УДК 519.854

ОПТИМАЛЬНІ НАБЛИЖЕНІ АЛГОРИТМИ РЕОПТИМІЗАЦІЇ УЗАГАЛЬНЕНИХ ПРОБЛЕМ ПРО ВИКОНУВАНІСТЬ З ПРЕДИКАТАМИ РОЗМІРНОСТІ 2

І. В. Сергієнко, В. О. Михайлюк

Резюме. Показано, що якщо для проблеми $Max - E2CSP - P$ існує поліноміальний оптимальний (пороговий) ρ-наближений алгоритм, то для проблеми $Ins - Max - E2CSP - P$ (реоптимізація $Max - E2CSP - P$ з додаванням одного обмеження) існує поліноміальний оптимальний (пороговий) $\varphi(\rho)$-наближений алгоритм, де $\varphi(\rho) = \frac{1}{2-\rho}$. Цей результат застосовується для проблем з $P = XOR$ (реоптимізація Max Cut) і з $P = OR$ (реоптимізація Max 2-Sat) при виконанні унікальної ігрової гіпотези (UGC).

Вступ

Важливий клас оптимізаційних проблем, який містить багато цікавих і вже добре вивчених задач дискретної оптимізації, представляють собою проблеми про узагальнену виконуваність або CSP проблеми (Constraint Satisfaction Problems). В таких проблемах є множина з n змінних і множина обмежень (що задаються предикатами), кожне з яких залежить від деякого числа змінних, і мета полягає в тому, щоб знайти такі приписування значень змінних, які виконують максимальне число обмежень. Предикати від 2 змінних є основними при вивченні властивостей узагальнених проблем про виконуваність. Визначення максимального числа виконаних обмежень в таких проблемах є NP-складним (NP-hard) і така проблема відома як $Max-2CSP$. Ця проблема має два частинних найбільше вивчених випадки: Max Cut і Max 2-Sat проблеми. В Max Cut проблемі кожне обмеження задається предикатом $x_i \oplus x_j$, тобто воно істинно, якщо точно один із входів є істинним. В Max 2-Sat проблемі кожне обмеження має вигляд $l_i \vee l_j$, тобто диз'юнкція двох літералів, кожний з яких або змінна, або її заперечення.

Розв'язати NP-складні оптимізаційні проблеми точно за поліноміальний час навряд чи представляється можливим. Тому розглядаються ефективні наближені алгоритми для розв'язування таких задач. Для максимізаційної проблеми кажуть, що алгоритм є C-наближеним алгоритмом, якщо він для довільного екземпляра дає розв'язок зі значенням цільової функції не меншим, ніж $C \cdot OPT$, де $OPT-$ глобальний оптимум. При цьому C називають відношенням апроксимації. Подібне означення можна дати для минімізаційних проблем.

Фундаментальне питання для заданої NP-складної проблеми Q — визначити, для яких значень C можна надіятись на ефективний (поліноміальний) C-наближений алгоритм. Це велика дослідницька область в теоретичній информатиці зі своїми позитивними та негативними результатами. Кажуть, що для проблеми Q встановлена верхня оцінка відношення апроксимації C, якщо існує поліноміальний C-наближений алгоритм для розв'язання Q. Для проблеми Q встановлена нижня оцінка відношення апроксимації c, якщо для довільного $\epsilon > 0$ не існує поліноміального наближеного алгоритма для Q, на якому досягається відношення апроксимації $c + \epsilon$. Якщо $C = c$, то для проблеми Q встановлений **поріг відношення апроксимації** (рівний $C = c$). Відповідний алгоритм називається **пороговим** або **оптимальним**.

Проблема встановлення нижніх оцінок відношення апроксимації (як і довільна проблема отримання нижніх оцінок складності) є дуже важкою задачею. Для такої проблеми існує назва неапроксимованість (inapproximability) або складність апроксимації (hardness of approximation). Великий вплив на развиток методів отримання нижніх оцінок здійснила відома PCP теорема [1] і дискретний аналіз Фур'є для тестування властивостей проблем (property testing) [2]. Поруч зі звичайними детермінованими алгоритмами розглядають випадкові алгоритми, де оцінюється очікувані значення розв'язку. Протягом десятиліть випадковий алгоритм для проблем Max Cut і Max 2-Sat був найкращим і лише тільки в 1995 році Гоеманс і Уільямсон (Goemans and Williamson) [3] запропонували 0,87856-наближені алгоритми для Max Cut і Max 2-Sat. Вони застосували техніку напіввизначеного програмування (semidefinite programming, SDP) для знаходження оптимального розв'язку з високою точністю (алгоритм внутрішньої точки), а потім "заокруглили" розв'язок до дискретного варіанту початкової проблеми. Цей підхід потім був успішно застосований до інших комбінаторних оптимізаційних проблем.

Як вже зазначалося, проблема неапроксимованості успішно була розв'язана для багатьох проблем завдяки PCP теоремі. Зокрема, Хастад (Hastad) [4] показав, що узагальнення Max Cut і Max 2-Sat від 2 до 3 змінних, тобто Max-E3-Lin-2 і Max 3-Sat є NP-складними для апроксимації з відношеннями $1/2 + \epsilon$ і $7/8 + \epsilon$ відповідно. Це означає, що найпростіший випадковий алгоритм приписування є найкращим для цих проблем, якщо $P \neq NP$ або, що відношення $1/2$ і $7/8$ є пороговими. В [5] показано (також з допомогою PCP теореми), що задача про покриття множинами має поріг відношення апроксимації, що дорівнює $\ln n$.

При вивченні проблеми неапроксимованості для проблем про узагальнену виконуваність з предикатами розмірності 2 такий прогрес не був досягнутий. Найкращі результати по складності апроксимації для Max 2-CSP, Max 2-Sat і Max Cut є відповідно $9/10 + \epsilon \approx 0,900$, $21/22 + \epsilon \approx 0,955$ і $16/17 + \epsilon \approx 0,941$ [4, 6]. Найперспективніший підхід до отримання сильних результатів (порогів відношень апроксимації) — це так звана унікальна ігрова гіпотеза (Unique Games Conjecture, UGC), введена С. Кхотом

(S. Khot) [7]. Унікальна ігрова гіпотеза (UGC) — це одна з найважливіших в сучасній теоретичній інформатиці відкритих проблем за кількістю сильних результатів по неапроксимованості, які одержані завдяки UGC.

Для Max 2-CSP проблем отримані такі результати. Кхот [8] показав, що з UGC випливає $\alpha_{GW} + \epsilon$ результат по складності апроксимації для Max Cut, де $\alpha_{GW} \approx 0,87856$ відношення апроксимації алгоритма Гоеманса-Уільямсона [3]. В [9, 10] показано, що з UGC випливає $\alpha^{-}_{LLZ} + \epsilon$ результат по складності апроксимації для Max 2-Sat, де $\alpha^{-}_{LLZ} \approx 0,94016$ відношення апроксимації алгоритма Левіна-Лівната-Звіка (Lewin-Livnat-Zwick) [11]. Відомі інші роботи, де найкращі результати по неапроксимованості, що задовольняють UGC, збігаються з відношенням апроксимації наближених алгоритмів напіввизначеного програмування.

Поняття реоптимізації [12]–[19] полягає в такому. Нехай Q — деяка NP-складна (можливо, NP-повна) проблема, I — початковий екземпляр проблеми Q, оптимальне рішення якого відомо. Пропонується новий екземпляр I' задачі Q, отриманий деякими “незначними” змінами екземпляра I. Виникає питання: як можна ефективно використати знання про оптимальне рішення I для обчислення точного або наближеного рішення екземпляра I'? Мета реоптимізації при використанні наближених методів — застосування знань про рішення початкового екземпляра I при умові: або досягнення кращої якості наближення (апроксимаційного відношення) I'; або створення ефективнішого (за часом) алгоритму визначення оптимального або близького до нього розв’язку I'; або виконання першого і другого пунктів.

Відомі такі результати по реоптимізації дискретних задач оптимізації. При вставці элементарної диз’юнкції реоптимізація Max Weighted Sat (виважена задача про виконуваність на максимум) — апроксимована з відношенням 0,81, хоча Max Weighted Sat — апроксимована з відношенням 0,77 [18]. При вставці вершини в граф реоптимізація Min Vertex Cover (мінімальне вершинне покриття графа) апроксимована з відношенням 1,5, Min Vertex Cover — з відношенням 2 [18]. При вставці вершини (термінальної або ні) реоптимізація Min Steiner Tree (мінімальне дерево Штейнера) апроксимована з відношенням 1,5, Min Steiner Tree — апроксимована з відношенням $1 + \ln 3/2 \approx 1,55$ [15]. При вставці або видаленні елемента з множини, задача про покриття множинами апроксимована з відношенням $2 - \frac{1}{\ln m + 1}$, де m — число елементів множини. Подібний результат має місце при вставці або видаленні довільного числа $1 < p < m$ елементів з множини [19]. Треба відмітити цілий цикл робіт по реоптимізації задачі про комівояжера (TSP — Travelling Salesman Problem) [12]–[14], [16]. Наприклад, задача Minimum Metric TSP (Min TSP — задача про комівояжера на мінімум з метричними відстанями) апроксимована з відношенням 1,5, її реоптимізація при вставці нового вузла — з відношенням 1,34, реоптимізація цієї задачі при зміні відстаней — апроксимована з відношенням 1,4 [18]. Для загальної задачі про комівояжера (Min TSP) невідомі оцінки апроксимації як для неї самої, так і для різноманітних версій реоптимізації.

Основні результати даної роботи полягають в такому. Досліджувались реоптимізаційні варіанти задач про узагальнену виконуваність з предикатами розмірності 2. Доведені такі твердження.

Теорема 1. *Якщо для проблеми $Max - E2CSP - P$ існує поліноміальний пороговий (оптимальний) ρ-наближений алгоритм, то для проблеми $Ins - Max - E2CSP - P$ (реоптимізація $Max - E2CSP - P$) існує поліноміальний пороговий (оптимальний) $\varphi(\rho)$-наближений алгоритм, де $\varphi(\rho) = \frac{1}{2-\rho}$.*

Теорема 2. *Припустимо, що справедлива унікальна ігрова гіпотеза (UGC). Тоді для проблеми $Ins - Max - E2CSP - XOR$ (реоптимізація Max Cut) існує поліноміальний пороговий (оптимальний) $\varphi(\alpha_{GW})$-наближений алгоритм, де $\varphi(\alpha_{GW}) = \frac{1}{2-\alpha_{GW}}$.*

Теорема 3. *Припустимо, що справедлива унікальна ігрова гіпотеза (UGC). Тоді для проблеми $Ins - Max - E2CSP - OR$ (реоптимізація Max 2-Sat) існує поліноміальний пороговий (оптимальний) $\varphi(\alpha^-_{LLZ})$ -наближений алгоритм, де $\varphi(\alpha^-_{LLZ}) = \frac{1}{2-\alpha^-_{LLZ}}$.*

Згідно з теоремою 2 поріг відношення апроксимації для реоптимізації Max Cut (при додаванні довільного ребра в граф) дорівнює $\varphi(\alpha_{GW}) \approx 0,891716$. Згідно з теоремою 3 поріг відношення апроксимації для реоптимізації Max 2-Sat (при додаванні довільної диз'юнкції) дорівнює $\varphi(\alpha^-_{LLZ}) \approx 0,943544$.

1. Основні означення

Наведемо необхідні позначення і означення [4, 21]. Під предикатом P розмірності k будемо розуміти відображення $P : \{-1,1\}^k \to \{0,1\}$. Для зручності позначень вхідні дані зі значенням –1 інтерпретуємо як "істина", зі значенням 1 — як "хибність". Якщо предикат P *приймає вхідне значення* y, то $P(y) = 1$, інакше $P(y) = 0$. Таким чином, множина значень, що приймається предикатом P, позначається як $P^{-1}(1)$. Літерал — це булева змінна або її заперечення.

Означення 1. Нехай $P : \{-1,1\}^k \to \{0,1\}$ є предикат. Екземпляр проблеми $Max - CSP - P$ складається з m обмежень з вагами, кожне з яких є k-кортеж літералів $(z_{i_1}, ..., z_{i_k})$ з множини $\{x_1, ..., x_n, \overline{x}_1, ..., \overline{x}_n\}$. Всі змінні в цьому кортежі вважаються різними. Обмеження виконано тоді і тільки тоді, коли P *приймає* цей кортеж. Рішення є приписування значень істинності до $\{x_1, ..., x_n\}$. Значення рішення є $\sum_{i=1}^{n} w_i P(z_{i_1}, ..., z_{i_n})$, де w_i є (невід'ємна) вага i-ого обмеження. Мета полягає в максимізації цього значення. Коли P залежить від не більше ніж k літералів, $Max - CSP - P$ будемо називати $Max - kCSP - P$, якщо в P дорівнює k літералів — то $Max - EkCSP - P$.

Означення 2. Два предикати P і P' мають однаковий тип тоді і тільки тоді, коли існує перестановка π на $[k] = \{1, ..., k\}$ і $a \in \{-1,1\}^k$ таке, що $P(x_1, ..., x_k) = P'(a_1 x_{\pi(1)}, ..., a_k x_{\pi(k)})$ для всіх $x \in \{-1,1\}^k$.

Якщо P і P' мають однаковий тип, то екземпляр $Max - CSP - P$ може бути представлений як екземпляр $Max - CSP - P'$, переставляючи кортежі згідно масці. Тобто ці проблеми еквівалентні.

Означення 3. Проблема $Max - kCSP - P$, де кожне обмеження є диз'юнкція не більше k літералів, є проблемою $Max - k - SAT$. Якщо кожне обмеження містить рівно k літералів, то це проблема $Max - Ek - SAT$.

Означення 4. Проблема $Max - kCSP - P$, де кожне обмеження є добутком не більше k літералів дорівнює константі, є проблемою $Max - k - LIN$. Якщо кожне обмеження містить рівно k літералів, то це проблема $Max - Ek - LIN$.

Нехай $w_{opt}(I)$ — значення оптимального рішення екземпляра I.

Означення 5. Алгоритм A є C-наближений алгоритм для максимізаційної проблеми, якщо для всіх екземплярів I проблеми $w(A, I) \geq C \cdot w_{opt}(I)$, де $w(A, I)$ — значення розв'язку алгоритма A на вході I. При цьому кажуть, що A має апроксимаційне відношення C. Для ймовірнісних алгоритмів $w(A, I)$ чекаєме значення (математичне сподівання) серед випадкових виборок алгоритма A.

Кажуть, що предикат P є апроксимаційно стійким (як і відповідна проблема $Max - CSP - P$), якщо знайти розв'язок $Max - CSP - P$, який значно кращий, ніж очікуване значення при випадковому приписуванні, є NP-складним. Поскільки випадкове приписування виконує довільне P-обмеження з ймовірністю $d(P) = 2^{-k}|P^{-1}(1)|$, маємо таке означення.

Означення 6. Предикат $P : \{-1, 1\}^k \to \{0, 1\}$ називається апроксимаційно стійким, якщо для довільної константи $\epsilon > 0$ знайдеться розв'язок x екземпляра I проблеми $Max - CSP - P$ такий, що значення x не менше, ніж $(d(P) + \epsilon)w_{opt}(I)$, є NP-складним.

Означення 7. Проблема $Max - CSP - P$ завжди апроксимована, якщо для довільного $\delta > 0$ існує $\epsilon_\delta > 0$ і ефективний алгоритм, який виходячи з екземпляра, де $(d(P)+\delta)$ частина обмежень може бути одночасно виконана, знаходить приписування, яке виконує не менше, ніж $(d(P) + \epsilon_\delta)$ частину обмежень.

Якщо проблема не є завжди апроксимованою, то вона апроксимаційно стійка.

В [20] встановлений поріг відношення апроксимації для алгоритмів реоптимізації проблем $Max - EkCSP - P$ при $k = O(\log n)$ з апроксимаційно стійкими предикатами.

Будемо розглядати випадок $k = 2$. Введемо предикати $AND(x_1, x_2) = x_1 \wedge x_2$, $XOR(x_1, x_2) = x_1 \oplus x_2$ і $OR(x_1, x_2) = x_1 \vee x_2$. Можна показати, що вони описують з точністю до однакових типів (означення 2) всі предикати розмірності 2.

Означення 8 (Max Cut). Для даного неорієнтованого графа $G = (V, E)$ із множиною вершин V і ребер E Max Cut є проблема знаходження розбиття $C = (V_1, V_2)$ вершин $V(V = V_1 \cup V_2, V_1 \cap V_2 = \emptyset)$, яке максимізує розмір множини $(V_1 \times V_2) \cap E$. Для заданої вагової функції $w : E \to R^+$ Max Cut проблема з вагами полягає в максимізації $\sum_{e\in(V_1\times V_2)\cap E} w(e)$.

Означення 9 (Max 2-Sat). Екземпляр Max 2-Sat проблеми є множина змінних і множина диз'юнкцій від двох літералів, де літерал є змінна або її заперечення. Проблема полягає в приписуванні змінним значень так, щоб число виконаних диз'юнкцій було максимальним. Для заданої невід'ємної вагової функції над множиною диз'юнкцій Max 2-Sat проблема з вагами полягає в максимізації сум вагів всіх виконаних диз'юнкцій.

Згідно з введеними означеннями проблема Max 2-Sat в нашому випадку буде позначатися як $Max - E2 - SAT$, або $Max - E2CSP - OR$. Розглянемо детальніше проблему Max Cut. Для графа $G = (V, E)$ із множиною вершин V і ребер E ця проблема (максимальний розріз в графі) визначається так: знайти таке разбиття V на V_1 і V_2, щоб максимізувати число ребер, які утворюють розріз, тобто лежать між двома частинами. Якщо кожній вершині поставити у відповідність булеву змінну $x_i (x_i = 1, i \in V_1, x_i = -1, i \in V_2)$, то дану проблему можна розглядати як $Max - E2CSP - XOR$ або $Max - E2 - LIN$ з рівняннями вигляду $x_i x_j = -1$. Відмітимо, що не існує предикатів розмірності $2 (k = 2)$, які є апроксимаційно стійкими. Таким чином, проблеми $Max - E2CSP - P$ завжди апроксимовані.

2. Напіввизначене програмування (SDP) і проблеми Max Cut та Max 2-Sat

Основна техніка для отримання ефективних наближених алгоритмів для $Max - CSP - P$ — це напіввизначене програмування, запропоноване Гоемансом та Уільямсоном [3]. Представимо коротко базовий алгоритм. Формалізуємо Max Cut як задачу квадратичного цілочисельного програмування:

$$max_{x\in\{-1,1\}^n} \sum_{(i,j)\in E} \frac{1 - x_i x_j}{2}. \tag{1}$$

Ця сума точно відповідає розміру максимального розрізу, поскільки кожний терм 1, якщо ребро належить розрізу и 0 інакше. Послабимо (1) введенням змінних y_{ij} для добутків $x_i x_j$, отримаємо

$$max_y \sum_{(i,j)\in E} \frac{1 - y_{ij}}{2}.$$

Припустимо, що змінні y_{ij} формують додатню симетричну напіввизначену матрицю з одиницями на діагоналі. Запишемо це таким чином

$$max_{y\geq 0, y_{ii}=1} \sum_{(i,j)\in E} \frac{1 - y_{ij}}{2}. \tag{2}$$

Відмітимо, що це релаксація початкової проблеми, так як $y_{ij} = x_i x_j$ і, дійсно, матриця задовольняє умовам. Оптимізаційна проблема (2) може бути розв'язана поліноміальним алгоритмом внутрішньої точки з довільною точністю. Для простоти будемо вважати, що знайдено точний оптимум. Запишемо (2) інакше. Відомо, що $n \times n$ матриця Y симетрично додатньо напіввизначена тоді і тільки тоді, коли існує інша $n \times n$ матриця V така, що $Y = V^T V$. Звідси випливає, що знайдуться вектори (стовпці V) такі, що $y_{ij} = (v_i, v_j)$ і вимога $y_{ii} = 1$ еквівалентна тому, що v_i — одиничний вектор. Таким чином, (2) еквівалентно

$$max_{(\|v_i\|=1)_{i=1}^n} \sum_{(i,j)\in E} \frac{1-(v_i,v_j)}{2}. \tag{3}$$

В такій постановці (3) строге узагальнення (1), поскільки можна інтерпретувати x_i як одиничний вектор. В даній постановці розв'язок з високою розмірністю використовується для визначення розв'язку розмірності 1. Пропонується така процедура заокруглення, використовуючи випадкові гіперплощини. Нормальні вектори цих випадкових гіперплощин рівномірно разподілені на n-вимірній сфері. Таким чином, вектор $r \in R^n$ вибираємо випадково і рівномірно і покладемо $x_i = sign((r, v_i))$ (при $(r, v_i) = 0$ покладемо $x_i = 0$). Проаналізуємо цю процедуру заокруглення. Припустимо, що кут між v_i і v_j є θ_{ij}. Тоді вклад в цільову функцію буде $\frac{1-\cos\theta_{ij}}{2}$, а ймовірність, що ребро належить розрізу, тобто $sign((r, v_i)) \neq sign((r, v_j))$ є $\frac{\theta_{ij}}{\pi}$. Визначимо дійсне число α_{GW} так

$$\alpha_{GW} = min_{0\le\theta\le\pi} \frac{\theta/\pi}{(1-\cos\theta)/2}.$$

Числове значення α_{GW} приблизно дорівнює 0,87856. Маємо такий ланцюжок нерівностей:

$$E[\sum_{(i,j)\in E} \frac{1-x_i x_j}{2}] = \sum_{(i,j)\in E} \frac{\theta_{ij}}{\pi} \ge \alpha_{GW} \sum_{(i,j)\in E} \frac{1-\cos\theta_{ij}}{2} \ge \alpha_{GW} \cdot OPT$$

Остання нерівність випливає з того, що максимум релаксації не менший ніж дійсний максимум. Таким чином, даний алгоритм є α_{GW}-наближеним алгоритмом. Цей алгоритм може бути дерандомізований [23] з довільно малою втратою якості отриманого розв'язку. Формалізуємо Max 2-Sat (або $Max{-}E2CSP{-}OR$) як задачу квадратичного цілочисельного програмування. Екземпляр ψ цієї задачі на множині n змінних складається з множини дужок, де кожна дужка $\psi \in \Psi$ є диз'юнкція $l_i \vee l_j$ двох літералів, а кожний літерал є змінна або її заперечення, тобто має вигляд $b \cdot x_i$ для $b \in \{-1, 1\}$ і деякої змінної x_i. Кожна дужка ψ має невід'ємну вагу $w(\psi)$ (невиважений варіант, якщо $w(\psi) = 1$). Треба знайти таке приписування $x \in \{-1,1\}^n$ змінним, щоб сума ваг виконаних дужок була максимальною. Арифметизуємо кожну дужку $(b_1 x_i \vee b_2 x_j)$ як $\frac{3-b_1x_i-b_2x_j-b_1b_2x_ix_j}{4}$, де останній вираз дорівнює 1, якщо дужка виконана і 0 інакше. Тоді значення приписування $x \in \{-1,1\}^n$ для Ψ є

$$Val_\Psi(x) = \sum_{\psi=(b_1x_i \vee b_2x_j)\in\Psi} w(\psi) \cdot \frac{3-b_1x_i-b_2x_j-b_1b_2x_ix_j}{4}$$

і можна записати Max 2-Sat екземпляр Ψ як задачу квадратичного цілочисельного програмування виду:

$$max\{Val_{\Psi}(x)\}, x_i \in \{-1, 1\}. \qquad (4)$$

Побудуємо релаксацію напіввизначеного програмування (SDP релаксацію) для задачі Max 2-Sat. Покладемо $x_{n+i} = \overline{x}_i$ для $1 \leq i \leq n$ і $x_0 = 0$. Тоді кожна дужка буде мати вигляд $x_i \vee x_j, 0 \leq i, j \leq 2n$. Екземпляр Max 2-Sat може бути представленим як масив $(w_{ij}), 0 \leq i, j \leq 2n$, де w_{ij} вага дужки $x_i \vee x_j$. Мета приписати змінним $x_1, ..., x_{2n}$ такі булеві значення, що $x_{n+i} = \overline{x}_i$ і загальна вага виконаних дужок — максимальна. Одиничный вектор v_i припишемо кожному $x_i, 0 \leq i \leq 2n$. Через те, що $x_0 = 0$, вектор v_0 відповідає константі 0 (хибність). Для забезпечення узгодженості покладемо $v_i \cdot v_{n+i} = -1$ для $1 \leq i \leq n$ ($v_i \cdot v_{n+i}$ позначає скалярний добуток (v_i, v_{n+i})). Через те, що розв'язки залежить від скалярного добутку, припустимо $v_0 = (1, 0, ..., 0) \in R^{n+1}$. Відмітимо, що обмеження "нерівність трикутника" має місце для всіх $1 \leq i, j \leq 2n$ і поскільки $v_i = -v_{n+i}$ для $1 \leq i \leq n$ маємо, якщо v_0, v_i, v_j будуть допустимими розв'язками SDP релаксації, то

$$\begin{aligned}
v_o \cdot v_i + v_0 \cdot v_j + v_i \cdot v_j &\geq -1, \\
-v_0 \cdot v_i - v_0 \cdot v_j + v_i \cdot v_j &\geq -1, \\
-v_o \cdot v_i + v_0 \cdot v_j - v_i \cdot v_j &\geq -1, \\
v_o \cdot v_i - v_0 \cdot v_j - v_i \cdot v_j &\geq -1.
\end{aligned}$$

Можна посилити релаксацію, вимагаючи $v_i \cdot v_j + v_i \cdot v_k + v_j \cdot v_k \geq -1$ для $1 \leq i, j, k \leq 2n$ (що еквівалентно нерівності трикутника вигляду $\|v_i - v_j\|^2 + \|v_j - v_k\|^2 \geq \|v_i - v_k\|^2$). Таким чином, SDP релаксація Max 2-Sat буде мати вигляд:

$$\begin{gathered}
max\{(1/4) \sum_{i,j} w_{ij}(3 - v_0 \cdot v_i - v_0 \cdot v_j - v_i \cdot v_j)\}, \\
v_0 \cdot v_i + v_0 \cdot v_j + v_i \cdot v_j \geq -1, 1 \leq i, j \leq 2n, \qquad (5) \\
v_i \cdot v_{n+i} = -1, 1 \leq n, \\
v_i \in R^{n+1}, v_i \cdot v_i = 1, 0 \leq i \leq 2n.
\end{gathered}$$

Задача (5) розв'зується з аддитивною відносною помилкою ϵ за час поліноміальний від $\log 1/\epsilon$ [22]. Після цього здійснюється обертання вектора v_i відносно вектора v_0 і потім застосовується техніка заокруглення. Причому випадкові гіперплощини вибираються вже не рівномірно розподіленими на n-вимірній сфері, а асиметрично розподіленими (skewed distributions). Для техніки заокруглення застосовуються цілі схеми (конфігурації) заокруглення, які залежать від вибору функції заокруглення. В результаті був отриманий α^{-}_{LLZ}-наближений алгоритм [11, 9] для розв'язання Max 2-Sat, де $|\alpha^{-}_{LLZ} - 0,94016567245| \leq 5 \cdot 10^{-11}$ (слід відмітити, що для визначення α^{-}_{LLZ} використовувався чисельний алгоритм; аналітичні доведення не приводились).

3. Унікальна ігрова гіпотеза (UGC)

Унікальна ігрова гіпотеза (Unique Games Conjecture, UGC) була введена Кхотом (Khot) [7] як можливий спосіб отримання нових сильних результатів по неапроксимованості. Сформулюємо цю гіпотезу в термінах проблеми про покриття міток (Label Cover problem).

Означення 10. Екземпляр $X = (V, E, w, [L], \{\sigma_e^v, \sigma_e^w\}_{e=(u,v)\in E})$ унікальної проблеми про покриття міток (Unique Label Cover, ULC) визначається таким чином: заданий граф $G = (V, E)$ з ваговою функцією $w : E \to [0, 1]$, множина допустимих міток $[L] = \{1, ..., L\}$ і для кожного ребра $e = (v, w)$ дві перестановки $\sigma_e^v, \sigma_e^w \in \Sigma_L$ такі, що $\sigma_e^w = (\sigma_e^v)^{-1}$, тобто вони взаємно обернені. Будемо говорити, що функція $\ell : V \to [L]$ (яка називається маркіровкою вершин) задовольняє ребро $e = (v, w)$, якщо $\sigma_e^v(\ell(v)) = \ell(w)$ або еквівалентно, якщо $\sigma_e^w(\ell(w)) = \ell(v)$. Значення ℓ є загальна вага ребер, які вона задовольняє, тобто $Val_X(\ell) = \sum_e w(e)$, де сума береться по всім ℓ таким, що ℓ задовольняє e. Значення X є максимальна частина ребер, що задовольняються, по всім маркіровкам, тобто $Val(X) = max_\ell\{Val_X(\ell)\}$.

ULC проблема, де G є двохдольний граф може бути інтерпретована як одно-раундова гра двох гравців, в якій приймаючий предикат перевіряючого такий, що для даної відповіді одного з гравців завжди існує *унікальна* відповідь для іншого гравця така, що перевіряючий приймає. Ймовірність, що перевіряючий приймає, якщо гравці притримуються оптимальної стратегії, тоді дорівнює $Val(X)$. Звідси термінологія "унікальна ігрова гіпотеза". Будемо притримуватися *розривної версії ULC проблеми (gap version)*, яку можна визначити так.

Означення 11. $Gap - ULC_{\eta,\gamma,L}$ є проблема для даного екземпляра X ULC проблеми з множиною міток $[L]$ визначити: або $Val(X) \geq 1 - \eta$, або $Val(X) \leq \gamma$.

Унікальна ігрова гіпотеза (UGC) [9]**.** Для довільних $\eta > 0, \gamma > 0$ існує константа $L > 0$ така, що $Gap - ULC_{\eta,\gamma,L}$ є NP-складною. Таким чином, UGC стверджує, що розривна версія ULC важкорозв'язна для довільних достатньо малих η і γ з достатньо великим L. Справедливі такі результати.

Теорема 4. [3] *Припустимо, що справедлива унікальна ігрова гіпотеза (UGC). Тоді для довільного $-1 < \rho < 0$ і $\epsilon > 0$ є NP-складним відрізнити екземпляри Max Cut, які не менше ніж $(1/2 - (1/2)\rho)$-виконані, від екземплярів, які не більше ніж $((\arccos\rho)/\pi + \epsilon)$-виконані. Зокрема, вибираючи $\rho = \rho^*$, де*

$$\rho^* = argmin_{-1<\rho<0}\{\frac{(\arccos\rho)/\pi}{1/2 - (1/2)\rho}\} \approx -0,689$$

отримаємо, що NP-складно апроксимувати Max Cut з довільним відношенням більшим, ніж Гоеманса-Уільямсона константа $\alpha_{GW} \approx 0,878567$.

Оскільки $\alpha_{GW} = min_{-1<\rho<0}\{\frac{(\arccos\rho)/\pi}{1/2-(1/2)\rho}\} = min_{0\leq\theta\leq\pi}\{\frac{\theta/\pi}{(1-\cos\theta)/2}\}$, то оптимальний вибір θ — розв'язок рівняння $\theta = \tan(\theta/2)$,$\theta^* \approx 2,33 \approx 134°$, $\alpha_{GW} = \frac{2}{\pi\sin\theta^*}$.

Теорема 5. [9] *Припустимо, що справедлива унікальна ігрова гіпотеза (UGC).Тоді для довільного $\epsilon > 0$ є NP-складним апроксимувати Max 2-Sat з відношенням не меншим ніж $\alpha^-_{LLZ} + \epsilon$, де $\alpha^-_{LLZ} \approx 0,94017$ (апроксимаційне відношення алгоритма Левіна-Літвана-Звіка).*

Згідно з теоремою 1 можна стверджувати, що, приймаючи унікальну ігрову гіпотезу (UGC), відношення апроксимації α_{GW} є пороговим для проблеми Max Cut (алгоритм Гоеманса-Уільямсона — оптимальний). Те ж саме можна стверджувати (згідно з теоремою 2) відносно відношення апроксимації α^-_{LLZ}. Воно є пороговим для Max 2-Sat (алгоритм Левіна-Літвана-Звіка — оптимальний).

4. Поріг відношення апроксимації для реоптимізації Max Cut і Max 2-Sat

Нехай $Max - E2CSP - P$ — довільна безвагова CSP проблема з m обмеженнями ($w_i = 1, i \in [m]$). Нехай I — довільний екземпляр проблеми $Max-E2CSP - P$, екземпляр I' проблеми отримається з екземпляра I додаванням довільного $(m+1)$-ого обмеження $z^{(m+1)} = (z^{(m+1)}_{i_1}, z^{(m+1)}_{i_2})(z^{(m+1)}_{i_j} \in \{x_1, ..., x_n, \overline{x}_1, ..., \overline{x}_n\}, j = 1, 2)$. Визначимо реоптимізаційний варіант проблеми $Max - E2CSP - P$.

Проблема $Ins-Max-E2CSP-P$**. Вхідні дані.** Довільний екземпляр I проблеми $Max - E2CSP - P$, x^* — оптимальний розв'зок екземпляра I.

Результат. Знайти оптимальний розв'зок екземпляра I' (отриманого виходячи з I, як описано вище) проблеми $Max-E2CSP-P$, використовуючи x^*.

Мета. Знайти x, яке максимізує число виконаних обмежень екземпляра I'.

Оскільки $Max - E2CSP - P$ є NP-складною, то можна показати, що такою буде і $Ins - Max - E2CSP - P$.

Теорема 6. *Якщо для проблеми $Max - E2CSP - P$ існує поліноміальний ρ-наближений алгоритм, то для проблеми $Ins - Max - E2CSP - P$ (реоптимізація $Max - E2CSP - P$) існує поліноміальний $\varphi(\rho)$-наближений алгоритм, де $\varphi(\rho) = \frac{1}{2-\rho}$.*

Доведення. Застосуємо підхід, розглянутий в [19]. Нехай I — екземпляр проблеми $Max - E2CSP - P$, який складається з системи обмежень $C = \{z^{(i)}, i \in [m]\}$ і оптимального розв'зку x^*; $w(x^*)$ є число виконаних обмежень в системі C розв'зком x^*. До системи додається обмеження $z^{(m+1)}$, в результаті отримаємо екземпляр I' проблеми $Max - E2CSP - P$, нехай $x^*_{I'}$ його оптимальний розв'зок. Якщо $x^*_{I'}$ не виконує обмеження $z^{(m+1)}$, то x^* є оптимальний розв'зок екземпляра I' проблеми

$Ins - Max - E2CSP - P$, звідси

$$w(x^*) \geq w(x^*_{I'}) - 1 \tag{6}$$

(в лівій частині записана умова, що x^* оптимальний розв'зок I', а в правій, що оптимальний розв'зок не виконує обмеження $z^{(m+1)}$). Нехай $x^*_{I'}$ виконує обмеження $z^{(m+1)}$ і є l варіантів, при яких буде виконано обмеження $z^{(m+1)}$(очевидно $l < 2^2 = 4$). Побудуємо l наближених розв'зків $x^i(i \in [l])$ таким чином. Беремо i-е приписування, яке виконує $z^{(m+1)}$. Із системи обмежень видаляємо $z^{(m+1)}$ та до обмежень, які залишились (враховуючи результат приписування), застосовуємо деякий поліноміальний ρ-наближений алгоритм, отримаємо наближений розв'зок x^i. В результаті отримаємо

$$w(x^i) \geq \rho(w(x^*_{I'}) - 1) + 1 = \rho w(x^*_{I'}) + 1 - \rho \tag{7}$$

Домножуючи (6) на $1 - \rho$ і додаючи до (7), отримаємо

$$(1-\rho)w(x^*) + w(x^i) \geq (1-\rho)w(x^*_{I'}) - (1-\rho) + \rho w(x^*_{I'}) + 1 - \rho = w(x^*_{I'})$$

Серед розв'зків x^* і x^i вибираємо найкращий (тобто з найбільшим значенням цільової функції w) і позначаємо $\overline{x}$. Маємо

$w(x^*_{I'}) \leq (1 - \rho + 1)max\{w(x^*), w(x^i)\} = (2-\rho)w(\overline{x})$

звідки $w(\overline{x}) \geq \frac{1}{2-\rho}w(x^*_{I'})$. Таким чином, в результаті виконання описаного алгоритму отримано наближений розв'зок $\overline{x}$ екземпляра I' з відношенням апроксимації $\frac{1}{2-\rho}$. Ясно, що завжди $\frac{1}{2-\rho} > \rho(\rho \neq 1)$. $\square$

Теорема 7. *Якщо для проблеми $Max - E2CSP - P$ існує поліноміальний пороговий (оптимальний) ρ-наближений алгоритм, а для проблеми $Ins - Max - E2CSP - P$ (реоптимізація $Max - E2CSP - P$) існує поліноміальний γ-наближений алгоритм, то $\gamma \leq \varphi(\rho)$.*

Доведення. Нехай I — екземпляр проблеми $Max - E2CSP - P$, який складається із системи обмежень $C = \{z^{(i)}, i \in [m]\}$ і оптимального розв'зку x^*. До системи додається обмеження $z^{(m+1)}$, в результаті отримаємо екземпляр I' проблеми $Ins - Max - E2CSP - P$. Нехай $\overline{x}$ — розв'зок проблеми $Ins - Max - E2CSP - P$, отриманий в результаті алгоритма з доведення теореми 6. Розв'зок $\overline{x}$ є кращим (більшим за значенням цільової функції) з розв'зків x^* та $x^i(i \in [l], l < 4)$, він отримається поліноміальним наближеним алгоритмом з відношенням апроксимації $\varphi(\rho) = \frac{1}{2-\rho}$. Доведення проведемо від супротивного. Нехай $\gamma > \varphi(\rho)$ і ρ^* таке, що $\varphi(\rho^*) = \gamma$. Оскільки функція $\varphi(\rho)$ є зростаючою по ρ і $\varphi(\rho^*) = \gamma > \varphi(\rho)$, то звідси випливає, що $\rho^* > \rho$. А це суперечить тому факту, що для проблеми $Max - E2CSP - P$ існує поліноміальний пороговий (оптимальний) ρ-наближений алгоритм (тобто для отримання розв'зку x^i повинен застосовуватись поліноміальний алгоритм з відношенням апроксимації ρ^* більшим ніж ρ, що неможливо). $\square$

З теорем 6 та 7 випливає *доведення* теореми 1.

Доведення теореми 2. З теореми 4 випливає, що алгоритм Гоеманса-Уільямсона є поліноміальним пороговим (оптимальним) α_{GW}-наближеним алгоритмом для проблеми Max Cut ($Max - E2CSP - XOR$). Застосовуючи теорему 1, отримаємо, що для проблеми $Ins - Max - E2CSP - XOR$ (реоптимизація Max Cut) існує поліноміальний пороговий (оптимальний) $\varphi(\alpha_{GW}$-наближений алгоритм, де $\varphi(\alpha_{GW}) = \frac{1}{2-\alpha_{GW}}$. □

Доведення теореми 3. З теореми 5 випливає, що алгоритм Левіна-Лівната-Звіка є поліноміальним пороговим (оптимальним) α^{-}_{LLZ}-наближеним алгоритмом для проблеми Max 2-Sat ($Max - E2CSP - OR$). Застосовуючи теорему 1, отримаємо, що для проблеми $Ins - Max - E2CSP - OR$ (реоптимізація Max 2-Sat) існує поліноміальний пороговий (оптимальний) $\varphi(\alpha^{-}_{LLZ})$-наближений алгоритм, де $\varphi(\alpha^{-}_{LLZ}) = \frac{1}{2-\alpha^{-}_{LLZ}}$. □

Висновки

В даній роботі показано, що при виконанні унікальної ігрової гіпотези (UGC) для реоптимізації Max Cut (при додаванні довільного ребра в граф) і для реоптимізації Max 2-Sat (при додаванні довільної диз'юнкції) існують поліноміальні порогові (оптимальні) наближені алгоритми. Відмітимо, що хоча для проблеми $Max - E2CSP - AND$ (або Max 2-And) існує поліноміальний пороговий (оптимальний) наближений алгоритм (при виконанні UGC та ще деяких додаткових вимог) [10], однак обчислити відповідне відношення апроксимації представляється непростою задачею. В [10] показано, що оцінка складності апроксимації (нижня оцінка) є 0,87435, тоді як апроксимаційне відношення кращого алгоритму (верхня оцінка) — 0,87401 (тобто співпадають три значущі цифри після коми). Якщо вважати ці три цифри вірними, то згідно результатам цієї роботи для реоптимізації Max 2-And при додаванні нової кон'юнкції отримаємо поліноміальний пороговий (оптимальний) 0,888-наближений алгоритм. Результати цієї роботи суттєво залежать від істинності унікальної ігрової гіпотези (UGC). Поряд з проблемами взаємовідношення класів складності проблем по включенню (наприклад, $P \stackrel{?}{=} NP$) це одна з основних відкритих проблем сучасної теоретичної інформатики. Навіть якщо UGC хибна, може виявитись, що $Gap - ULC_{\eta,\gamma,L}$ складна в смислі нерозв'язності за поліноміальний час і така (слабка) складність може бути застосована до всіх проблем, де складність демонструвалась виходячи з UGC.

Література

1. Arora S. Proof verification and intractability of approximation problems/ S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy // Journal of the ACM. — 1998. — 45, №3. — P. 501–555.
2. Goldreich O. Locally testable codes and PCPs of almost-linear length / O. Goldreich and M. Sudan // Journal of the ACM. — 2006. — 53, №4. — P. 558–655.
3. Goemans M. X. Improved approximation algorithms for Maximum Cut and Satisfiability Problems using Semidefinite Programming/ M. X. Goemans and D. P. Williamson // Journal of the ACM. — 1995. — 42. — P. 1115–1145.

4. Hastad J. Some optimal inapproximability results/ J. Hastad // Journal of the ACM. — 2001. — 48, №4. — P. 798–859.
5. Feige U. A threshold of ln n for approximating set cover/ U. Feige // Journal of the ACM. — 1998. — 45, №4. — P.634–652.
6. Trevisan L. Gadgets, Approximation, and Linear Programming/ L. Trevisan, D. B. Sorkin, M. Sudan, D. P. Williamson // SIAM Journal on Computing. — 2000. — 29, №6. — P. 2074–2097.
7. Khot S. On the power of unique 2-prover 1-round games. / S. Khot // In Proc. 34th Annual ACM Symposium on the Theory of Computing (STOC). — 2002. — P. 767–775.
8. Khot S. Optimal Inapproximability Results for Max-Cut and Other 2-Variable CSPs?/ S. Khot, G. Kindler, E. Mossel, and R. O'Donnell // SIAM Journal on Computing. — 2007. — 37(1). — P. 319–357.
9. Austrin P. Balanced Max 2-Sat Might Not be the Hardest/ P. Austrin // In STOC. — 2007. — P. 189–197.
10. Austrin P. Towards Sharp Inapproximability For Any 2-CSP/ P. Austrin // In IEEE FOCS. — 2007. — P. 307–317.
11. Lewin M. Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems / M. Lewin, D. Livnat, U. Zwick // Lecture Notes in Computer Science. — 2002. — 2337. — P. 67–82.
12. Ausiello G. Reoptimization of minimum and maximum traveling salesman's tours / G. Ausiello, B. Escoffier, J. Monnot, and V. Th. Paschos // Algorithmic theory. — SWAT 2006, Lect. Notes Comput. Sci. — Berlin: Springer, 2006. — 4059. — P. 196—207.
13. Bockenhauer H.J. On the approximability of TSP on local modifications of optimal solved instances / H. J. Bockenhauer, L. Forlizzi, J. Hromkovic // Algorithmic Oper. Res. — 2007. — 2(2). — P. 83—93.
14. Bockenhauer H. J. On the hardness of reoptimization/ H. J. Bockenhauer, J. Hromkovic, T. Momke, and P. Widmayer // Proc. of the 34 th Intern. Conf. on Current Trends in Theory and Practice of Computer Science (SOF-SEM 2008); Lect. Notes Comput. Sci. — Springer: Berlin, 2008. — 4910. — P. 50–65.
15. Escoffier B. Simple and fast reoptimizations for the Steiner tree problem / B. Escoffier, M. Milanic, and V. Paschos // Algorithmic Operations Research. — 2009. — 4(2). — P. 86–94.
16. Archetti C. Reoptimizing the travelling salesman problem/ C. Archetti, L. Bertazzi, and M. G. Speranza // Networks. — 2003. — 42(3). — P. 154–159.
17. Archetti C. Reoptimizing the 0-1 knapsack problem/ C. Archetti, L. Bertazzi, and M. G. Speranza // Discrete Applied Mathematics. — 2010. — 158(17). — P. 1879–1887.
18. Ausiello G. Complexity and approximation in reoptimization/ G. Ausiello, V. Bonifaci, and B. Escoffier // Computability in Context: C omputation and Logic in the Real World (ed. S. Barry Cooper and Andrea Sorbi). — 2011. — London: Imperial College Press. — P. 101–130.
19. Михайлюк В. А. Реоптимизация задачи о покрытии множествами/ В. А. Михайлюк // Кибернетика и системный анализ. — 2010. — 46, №6. — С. 27–31.
20. Михайлюк В. А. Реоптимизация обобщенных проблем о выполнимости с аппроксимационно-устойчивыми предикатами / В. А. Михайлюк, И. В. Сергиенко // Кибернетика и системный анализ — 2012. — 48, №1. — С. 89–104.

21. Hast G. Beating a random assignment. — 2005. — Doctoral Thesis. — Royal Institute of Technology. — Stockholm, Sweden. — 102 p.
22. Alizadeh F. Interior point methods in semidefinite programming with applications to combinatorial optimization/ F. Alizadeh // SIAM Journal on Optimization — 1995. — 5. — P. 13–51.
23. Mahajan S. Derondomizing approximation algorithms based on semidefinite programming/ S. Mahajan, H. Ramesh // SIAM Journal on Computing. — 1999. — 28. — P. 1641–1663.

Інститут кібенетики ім. В.М. Глушкова НАН України, просп. Академіка Глушкова, 40, Київ 187, ГСП, 03680, Україна.

Надійшла 01.01.2012

Журнал обчислювальної та прикладної математики | 2012, №3(106) | Journal of Computational & Applied Mathematics

UDC 681.3 513, 004.6(075.8)

SOME APPROACHES TO SOLVING THE BICHROMATIC CLOSEST-PAIRS PROBLEM ON METRICS L_1

V. M. TERESHCHENKO, D. SUVOROV

Abstract. In this paper we present an original algorithm and data structures to solve the Bichromatic Closest Pair problem in metrics L_1. We got $O(n)$ time for the case of sets, each of which is coincided with the set of its maximums regarding to dominance, and *O(nlogn)* time in general. In addition, we offer a modification randomized Khuller and Matthias algorithm [3], which improves performance. So, Khuller and Matias algorithm approximates minimal distance with the closeness of multiplicative constant, that is if d — minimal distance, and d^* — approximation, then $d^* \leq d \leq ad^*$, where $a > 1$. In standard situation, $a = 4$, but this result can be perfected to $1+e$, for any $e > 0$, keeping the linear time algorithm.

Introduction

The closest pair problem in classical arrangement has been studied rather fully and is considered to be the central in the class of proximity problems. One of the cases of its stating is the problem about smallest distance (and couple of points) between two different sets. This problem is known as the bichromatic closest-pair problem (BCP). In particular, it is actual in metrics L_1 , because distance between two chips on the board is usually determined by this metric. Among the major works devoted to study the problem of nearest pair in its different variations are the following.

In [1–5] are presented a series of effective the deterministic and randomized algorithms for solving the bichromatic closest-pair problem (BCP) in plane for the metrics L_1, L_2. In a paper [2] authors represented simpler than Khuller and Mathias [3] algorithm, which combines the use of the function and randomization to solve the problem in $O(n)$ expected time. Much of the work reduces to maintaining a dynamic dictionary. In addition, the paper gives some expansion to the fully dynamic closest pair problem, and to the k closest pair problem. In a paper [4] presented a modification Rabin's algorithm so that only very few random bits are needed, but still a polynomial reliability is maintained, and in [5] shown a simple algorithm for the 1-dimensional case with $O(n^{3/4}log^2n)$ time on queries and $O(n)$ running time and proved the quantum lower bound for the problem. Works [6–8] are devoted to solving BCP-problem in R^d spaces ($d > 2$). In [6] is presented a new data structure as the *Leafary tree*. It is use for constructing an efficient randomized algorithm of solving problem in R^d space with $O(nlog(n)/log(log(n)))$ expected time. In [7, 8] are investigated the problems of finding the closest pair in R^d and the problems of maintaining the approximate BCP and the possibility of developing effective algorithms for its solution using reduction to NNS problems. Thus, Indyk and others in [7] proposed the first better-than- naive-solutions for the problem of finding the closest pair among a set S of n points in R^d. In [8] for the closest pair problem he gives a roughly n^{1+p} time Las Vegas algorithm with approximation factor $O(1/plog(1/p))$.

Approaches to solving the k-Closest-Pairs problem are presented in [9–12]. In [9] authors examines in detail the problem of finding the k closest pairs between two spatial data sets, where each set is stored in a structure belonging in the R-tree family. It is offered five different algorithms (four recursive and one iterative) for solving this problem and proved their high performance, and given experimental comparison with existing algorithms. To solve the problems of the k closest pairs in [10] Chan gives simpler randomized and deterministic algorithms with $O(nlog(n + k))$ running time in E^d space and $O(nlog(n) + n^{2/3}k^{1/3}log^{5/3}(n))$ time in Euclidean plane. The paper [11] shows that the k-closest pairs in set S of n, points in R^d can be found in $O(sort(n+k))$ I/Os and uses $O((n + k)/B)$ blocks of external memory. Angiulli and Pizzuti in [12] presented an $O(d^{1+1/t})$ approximate algorithm to solve efficiently the k-Closest-Pairs problem under the L_t metrics on large high-dimensional data sets in time $O(d^2nk)$ and linear space. The algorithm exploits the order induced on the data by the Hilbert space-filling curve to eliminate points early from the input data set. In a paper [13–16] presented algorithms to compute a Euclidean minimum spanning tree of a given set S of n points in E^d in time $O(F_d(n, n)log^d n)$, where $F_d(n, m)$ is the time required to compute a bichromatic closest pair among n red and m green points in E^d. In addition, there is described a randomized algorithm to compute a bichromatic closest pair in expected time $O((nmlog(n) \cdot log(m))^{2/3} + mlog^2(n) + nlog^2(m))$ in E^3. For $d \geq 4$ dimensions had been obtained $O((nm)^{1-1/(\lceil d/2\rceil+1)+\varepsilon} + mlogn + nlogm)$ time for the bichromatic closest pair problem for any $\varepsilon > 0$. And in [13] showed that the problem of computing a planar bichromatic minimum spanning tree for a set of n points in the plane is NP-hard, and gave an $O(\sqrt{n})$-approximation algorithm that runs in $O(nlog(n) \cdot log(log(n)))$ time.

1. Formulation and solving problem

Formally the bichromatic closest-pair problem according to dominance can be represented next way.

Problem.(*The Bichromatic Closest-Pair Problem* (*by dominance*)). Given two sets of points A and B in plane, each of which is coincided with the set of its maximums regarding to dominance (i.e. each set composes "stairs Fig. 1). Need to find points when $a \in A$ and $b \in B$, for which $dist(a, b) = inf(dist(a', b'))$, $a' \in A$, $b' \in B$. When a is Manhattan Distance (L_1- metrics), we will get a bichromatic distance problem on metrics L_1.

We consider two approaches to solving this problem. The first approach represents the original effective algorithm, and the second is the improved modification of known Khullera and Matthias ideas [3]. Consider the first approach.

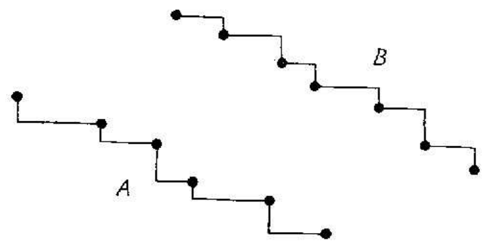

FIG. 1. Format sets A and B

2. ALGORITHM

The paper proposes an original algorithm for solving the problem with the expected linear time performance.

Theorem 1. *The bichromatic closest pair between two sets of points size n on the plane in the L_1 metrics can be found in $O(n)$ time.*

Proof. Both entering sets are given in a certain order: points are ordered as on the OX — axis and on the OY — axis respectively.

Lemma 1. *If the set of points in the plane defined as the ordered list $U = (P_1, P_2, ..., P_n)$ so that each next point list is situated more to the right and below the previous one, then this set ordered by $x - y$ (x "minus" y).*

Proof. Lets take point a_i and a_{i+1} and define $d_x(a_i, a_{i+1})$ and $d_y(a_i, a_{i+1})$ as the distances between projections of points a_i, a_{i+1} to axis OX and OY respectively. After that we can state that $d_x(a_i, a_{i+1}) \geq 0$, and $d_y(a_i, a_{i+1}) \leq 0$. If point a_i has "value"$x_i - y_i = d_i$, then point a_{i+1} has value $d_{i+1} = x_{i+1} - y_{i+1} = = x_i + d_x - (y_i + d_y) = x_i - y_i + (d_x - d_y)$. From $d_x \geq 0$, $dy \leq 0 => d_x - d_y \geq \geq 0 => d_{i+1} \geq d_i$.

Points of every set keep the order. Lets bring in structure of data us put data which will help us to solve the problem in $O(n)$ time. Doing this requires points to keep the order $x - y$, as well as the order $x + y$.

2.1 DATA STRUCTURE

Consider the input data and their description. Points from one set marked with grey and points from another — black respectively, Fig. 2.

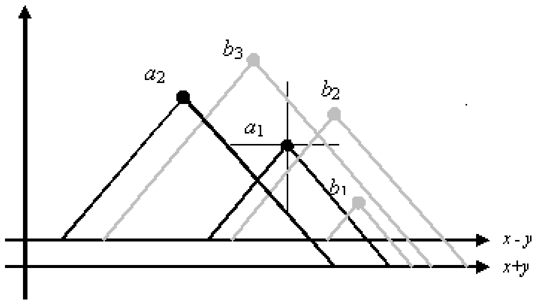

FIG. 2. $A = \{a_1, a_2\}$ — a set of black color points, $B = \{b_1, b_2, b_3\}$ — a set of grey color points

Their projections are also shown in the direction (-1,-1) and (1, -1) to axis OX, which are defined from $x - y$, and $x + y$. Regarding to this first keeps the initial order. First projection can be obtained by integration of two arrays. The second is kept in the input data structure.

The data is represented by the following structure: 4 lists, two of which contain two input arrays and the next two have integrated and sorted arrays, by $x - y$ and $x + y$ respectively. Two first lists denote as A, B, and the last two as XpY and XmY. Moreover, every element from lists A and B has index of its representative in lists XpY and XmY, and every element from XpY and XmY has index of its representative in A, B (Fig. 3).

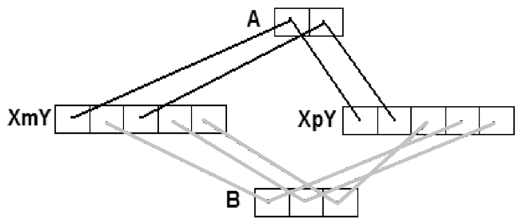

FIG. 3. The connection between lists A, B and XpY, XmY

Hence for every point from the first list its position in the second list can be indicated in time $O(1)$. For every black point its grey neighbor can be found in lists XmY and XpY in constant time. There is the following assertion.

Lemma 2. *For each black point closest to her on metrics L_1 grey point will be her neighbor in XmY or in XpY.*

Proof. As it was shown above the projection $x - y$ keeps the order of points in domination set. Let's examine for our black point (b) those grey points which are situated in the upper-left quadrant (UL) and in the base-right (BR) of our point, Fig. 4. Point from UL, distance from b to which is the smallest, will be the first grey point to the left from b on the projection $x - y$. It is proved by the fact that the nearest points d_x and d_y are the smallest and it follows that

$dx + dy$ is the smallest, that is why it is the nearest of all on $x - y$. It is similar for BR, here point is the nearer, if it is the first right one from b. Projection $x + y$ is analyzed for base-left and upper-right. Everything is proved similar to the previous case.

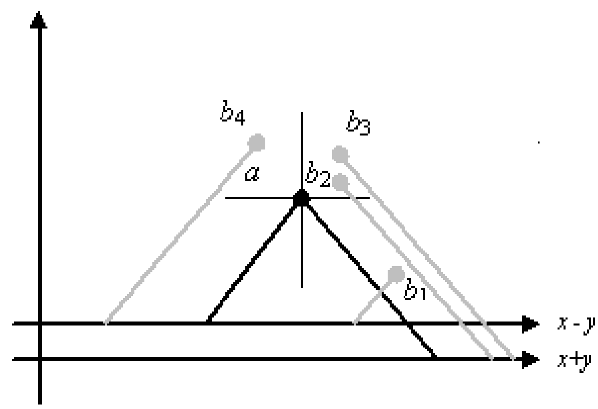

FIG. 4. Example for Lemma 2. $A = \{a\}$ — a set of black color points, $B = \{b_1, b_2, b_3\}$ — a set of grey color points

Therefore the nearest point to the point b is the one from (maximum) four points. It means that for every black point only 4 grey points have to be tested. According to this we have the following (line) algorithm.

2.2 Algorithm

1. Let $min = d(a_0, b_0)$, $min(a) = a_0$, $min(b) = b_0$.
2. Take further element a from the list A.
3. Find its representative a'' in XmY, and a'' in XpY.
4. For a' find the first from its left grey point b_1', and the first right b_2'. Find its representatives in B: b_1 and b_2 respectively.
5. For a'' find the first from its left grey point b_3', and first right b_4'. Find its representatives in B: b_3 and b_4, respectively.
6. Compute $d' = min(d(a, b_1), d(a, b_2), d(a, b_3), d(a, b_4))$.
7. If $d' < min$, then suppose $min = d'$, $min(a) = a$, $min(b) = b_i$, where i such that $d(a, b_i) = d'$.
8. Go to step 2.

As the algorithm performs one underpass according to the list in time $O(n)$, and all other steps are done in constant time, the general complexity of the algorithm will be $O(n)$. This algorithm can be used and for the case of any sets of data. But in that case we have to perform sorting by two projections ($x + y$, $x - y$) , that requires $O(nlogn)$ time in general and is optimal for this case, since all the other steps of the algorithm do not exceed $O(n)$.

Alternative solution is very interesting which approximates minimal distance with given closeness. Modification of the known Khuller and Matias algorithm which improves its efficiency is suggested in the work. We use Manhattan distance instead Euclidean distance.

3. Modified Khuller-Matias algorithm

Khuller and Matias in the [3] proposed the randomized algorithm, which approximates minimal distance with the accuracy of multiplicative constant. That is if d — minimal distance, and d^* — approximation, then $d^* \leq d \leq ad^*$, where $a > 1$. The expected time of algorithm is $O(n)$. In the paper [2] presented simpler than Khuller and Mathias [3] algorithm, which combines using of the floor function with randomization to achieve an $O(n)$ expected time. Much of the work greyuces to maintaining a dynamic dictionary. For standard situation, $a = 4$, but this result can be perfected to $1 + e$, for any $e > 0$, keeping the line time of algorithm work.

First, let's have a look at the algorithm for looking for the closest pair of points in one set. Let S be the input set, where $d(S)$ — minimal distance between two points in this set. First of all we need to find approximations $d(S)$, and after that use it for finding $d(S)$. Approximation can be reached in several steps on each we filter current set excluding some points from it. Filtering continuing until the set is empty. In the end we reach the approximation.

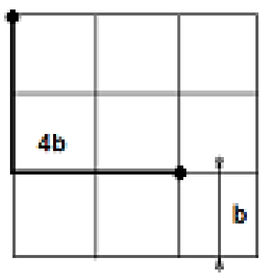

Fig. 5. The maximum distance between neighbors in a grid box b

3.1 Filtration

Let $d(x)$– the distance between x to the neighbor in the current set S_i. Take random point x_1, and find $d(x_1)$, then exclude all points x from the set Si, for which $(x) \geq d(x_1)$. Let $i*$ — the last iteration and x_i^*- the point which we choose during the last iteration. Then $d(x_i^*)/4 \leq d(S) \leq d(x_i^*)$. $d(x_i^*)$ — approximation. Let's put in the definition of "neighborhood"to show how to do it. Consider at the frame size b, then neighborhood of point x — is a cell which contains and other 8 neighboring cells, Fig. 5. Denote the set of neighboring points trough as $N(x)$. The furthest neighbor point can have distance of $4b$(biggest). So:

1) all points, that have the distance to is more than $4b$, are not neighbors;

2) all points, that have the distance which is less than b, are neighbors.

On every step we built the frame with the size of $d(x_i)/4$, and exclude from S_i all points which don't have neighbors. It can be done in time $O(|S_i|)$. It follows that every step of filtration is lineal according to the set of filtration. It is easy to show that on the last step: $d(x_i^*)/4 \leq d(S) \leq d(x_i^*)$. After that for the whole set S a frame is built with the size $b = d(x_i^*)$. It is obvious that for each point only constant amount of points can be neighbors. During the next step minimal distance between every point and its neighbors is defined. As the amount of these points is limited by the constant, time for each point is equal $O(1)$, and time for all points will be in $O(n)$ period. It is left to prove that the step of filtration also requires the time $O(n)$.

It is clear to derive from the following statement: the amount of points on every step is greyuced at least in geometrical progression. According to the fact that we exclude point for which $d(x) \geq d(x_i)$, follows that we exclude half of points in average, and: $\sum |S_i| \leq 2n$. Referring to the formula of decreasing geometrical progression and from the fact that $|S_{i+1}| \leq |S_i|/2$. So, time $O(n)$ is necessary.

3.2 e -approximation of minimal bichromatic distance

Let $d(x)$ — be the distance from to the nearest point of another color. During the filtration, when we choose random point x_i, we take away all such points, that $d(x) \geq d(x_i)$. We exclude at least half of the points with every step, so, the total amount of points will be $O(n)$, same as in the previous case. It is necessary to know that there is a neighbor of another color for each point. Let p, q — the pair of the nearest points, then for the distance $d(x_i^*)$ mentioned above the inequality $d(x_i^*)/4 \leq dist(p, q) \leq d(x_i^*)$ is held. Examining the frame size $d(x_i^*)$, points p and q will be neighbors but despite the previous case it is impossible to guarantee constant amount of points of another color. Browsing all such points the estimation value close to $O(n^2)$ will be obtained. We can come up with the following.

Let us look at the frame size $b = e * d(x_i^*)/16$. One point from all the points of one color is chosen as a representative of this cell for every cell. The distance from any point in this cell to the representative is less or equals $2b = e * d(x_i^*)/8$. So if p' and q' — are representatives of p and q respectively, then $dist(p', q') - dist(p, q) \leq e * d(x_i^*)/4 \leq e * dist(p, q)$. In that case $dist(p', q') \leq (1+e) * dist(p, q)$. The closest distant between the representatives is left to be found, and we will reach approximation with the closeness of $1+e$ of minimal distance. As there is no more than 1 representative in every cell, the amount of neighboring representatives for every point is constant. It follows that the time of the algorithm is $O(n)$ for every fixed e. If all points belong to the closed set and the distance between any two points is more than the certain constant d (for example pixels on the screen), such e can be chosen, so the distance between the representatives will bring closer the minimal distance with additive closeness of $d/2$ and so can be calculated concretely.

Conclusion

In this paper we developed an algorithm of finding minimal distance between two sets according to the metrics L_1, that gives $O(n)$ time in the case of sets, each of which is coincided with the set of its maximums regarding to dominance, and $O(nlogn)$ time in general. Further improvement of the algorithm is impossible, because we reached the optimal estimation for the general case [17]. In addition, all algorithm steps are trivial and require very little time. Access is sequential, so the algorithm can be easily and efficiently implemented. It could be interesting to look at it in comparison with the standard method "divide and conquer".

Bybliography

1. Preparata F. P. Computational Geometry:An introduction. / F. P. Preparata, M. I. Shamos — Berlin: Springer-Verlag, 1985. — 478 p.
2. Golin M. Simple Randomized Algorithms for Closest Pair Problems. / M. Golin, R. Raman, Ch. Schwarz, M. Smid // Nordic. J. Comput. — 1995. — 2. — P. 3–27.
3. Khuller S. A simple randomized sieve algorithm for the closest-pair problem. / S. Khuller, Y. Matias // Inf. Comput. — 1995. — 118(1). — P. 34–37.
4. Dietzfelbinger M. A reliable randomized algorithm for the closest-pair problem. / M. Dietzfelbinger, T. Hagerup, J. Katajainen, M. Penttonen // Journal of Algorithms. — 1997. — 1(25). — P. 19–51.
5. Volpato N. An $O(N^{3/4}log^2N)$ quantum algorithm for the 1- dimensional closest-pair problem. / N. Volpato, A. Moura // In WECIQ 2006: Workshop-Escola de Computacao e Informatica Quantica, October 2006: proceedings/ Pelotas, Brazil. — 2006. — P. 102–108.
6. Kamakoti V. Efficient randomized incremental algorithm for the closest pair problem using Leafary trees. / V. Kamakoti, K. Krithivasan, C. P. Rangan // Comp.and Comb. — 2006. — V. 959. — P. 71–80.
7. Indyk P. Closest Pair Problems in Very High Dimensions. / P. Indyk, M. Lewenstein, O. Lipsky, E. Porat // The International Colloquium on Automata, Languages and Programming (ICALP'04), July 12-16, 2004: Proceedings / Turku, Finland. — P. 782–792.
8. Indyk P. Dimensionality Greyuction Techniques for Proximity Problems. / P. Indyk // In Proc. 11th Annu. ACM-SIAM Sympos. Discrete Algorithms. — 2000. — P. 371–378.
9. Corrali A. Closest Pair Queries in Spatial Databases. / A. Corrali, Y. Manolopoulos, Y. Theodoridis, M. Vassilakopoulos // ACM SIGMOD. — 2000. — V. 29 , Is. 2. — P. 189–200.
10. Chan T. M. On enumerating and selecting distances. / T. M. Chan // Proc. 14th Annu. ACM Sympos. Compu. Geom. — 1998. — P. 279–286.
11. Govindarajan S. I/O-efficient well-separated pair decomposition and its applications. / S. Govindarajan, T. Lukovszki, A. Maheshwari, N. Zeh // Algorithmica. — 2006. — 45. — P. 585–614.
12. Angiulli F. Approximate k-Closest-Pairs in Large High-Dimensional Data Sets. / F. Angiulli, C. Pizzuti // Journal of Mathematical Modelling and Algorithms. — 2005. — V. 4, №2. — P. 149–179.
13. Agarwal P. Euclidean MST and bichromatic closest pairs. / P. Agarwal, H. Edelsbrunner, O. Schwarzkopf, E. Welzl // Discrete and Comp. Geom. — 1991. — 6. — P. 407–422.

14. Arora S. Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. / S. Arora // ACM (JACM). — 1998. — V. 45. — P. 753–782.
15. Borgelt M.G. Planar Bichromatic Minimum Spanning Trees. / M.G. Borgelt, M. van Kreveld, M. Loffler, J. Luo, D. Merrick, R. I. Silveira, M. Vahedi // Discrete Algorithms. — 2009. — V. 7, Is. 4. — P. 469–478.
16. D. Eppstein. Dynamic Euclidean minimum spanning trees and extrema of binary functions. / D. Eppstein // Discrete. Comput. Geom. — 1995. — 13(1). — P. 111–122.
17. D. T. Lee. Computational Geometry. In "Algorithms and Theory of Computation Handbook"edited by M. J. Atallah. — 1999.

Факультет кібернетики, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64, Київ, 01601, Україна.

Надійшла 03.04.2012

Журнал обчислювальної та прикладної математики | 2012, №1(107) | Journal of Computational & Applied Mathematics

УДК 517.9

ІТЕРАЦІЙНО-РІЗНИЦЕВІ МЕТОДИ ДЛЯ РОЗВ'ЯЗУВАННЯ НЕЛІНІЙНИХ ОПЕРАТОРНИХ РІВНЯНЬ

С. М. Шахно

Резюме. У роботі досліджуються ітераційні алгоритми для розв'язування нелінійних операторних рівнянь, які загалом не вимагають аналітичного задання похідних. Похідні апроксимуються поділеними різницями. Вивчаються питання локальної та напівлокальної збіжності методів, єдиності розв'язку рівняння при послаблених умовах до нелінійного оператора.

Вступ

Нехай задано нелінійне операторне рівняння

$$F(x) = 0, \tag{1}$$

де F — нелінійний оператор, визначений у відкритій області D банахового простору X зі значеннями в банаховому просторі Y.

Для розв'язування рівняння (1) широко використовують ітераційно-різницеві методи. Перевагою цих методів є те, що вони використовують у своїх ітераційних формулах тільки значення нелінійного оператора і не вимагають аналітично заданих похідних. Найпростішим методом такого типу є метод хорд [5, 14, 18, 25]. Однак метод хорд збігається до розв'язку лише зі швидкістю з порядком 1,618... У праці [4] запропоновано ітераційний різницевий метод лінійної інтерполяції з квадратичним порядком збіжності. Обидва методи використовують в ітераційних формулах значення нелінійного оператора з двох попередніх ітерацій. Ф.Потра розглянув різницевий метод [20], який використовує значення оператора з трьох попередніх ітерацій та має порядок збіжності 1,839... У працях [3, 28] запропоновано двокроковий різницевий ітераційний метод з порядком збіжності $1 + \sqrt{2}$, у якому на двох кроках використовується одне значення поділеної різниці. Його природно вважати двокроковою модифікацією саме методу хорд, а не методу Ньютона.

У праці [27] для дослідження методу Ньютона введено узагальнені умови Ліпшиця для оператора похідної, в яких замість константи Ліпшиця використано деяку додатну інтегровну функцію. У цій праці ми вводимо аналогічну узагальнену умову Ліпшиця для оператора поділеної різниці першого порядку і при цій умові досліджуємо збіжність різницевих методів. Відзначимо, що умова Ліпшиця з константою Ліпшиця є частковим випадком узагальненої умови Ліпшиця, а з отриманих нижче результатів випливають вже відомі теоретичні твердження.

Основні результати цієї статті — з єдиної точки зору досліджені основні ітераційні різницеві методи розв'язування нелінійних операторних рівнянь: метод хорд та його модифікація з квадратичною швидкістю збіжності — метод лінійної інтерполяції Курчатова, методи з порядками збіжності 1,839... та 2,41.... Методи вивчаються при послаблених вимогах до нелінійного оператора, що розширює область застосування методів. Також запропоновано та обґрунтовано двопараметричний метод, який містить згадані методи як часткові випадки. Досліджено локальну та напівлокальну збіжність методів, отримано апріорні та апостеріорні оцінки похибки, встановлено область єдиності розв'язку, введено узагальнені умови Ліпшиця. Головні результати цієї статті детально викладені в статтях [7, 10, 11, 12, 13, 14, 15, 17, 23, 25, 26].

Зокрема, у статтях [10, 12, 17] за звичайних умов Ліпшиця для поділених різниць досліджено локальну збіжність методів хорд та методу лінійної інтерполяції Курчатова, застосовано теорію мажорант Канторовича для вивчення напівлокальної збіжності, отримано апріорну та апостеріону оцінки похибки методів. Проведено порівняння цих методів з методом Потра.

У статтях [14, 25] введено поняття узагальненої умови Ліпшиця і за цих умов досліджено локальну збіжність методу хорд та встановлено область єдиності розв'язку.

У статтях [13, 26] вивчено збіжність методу Стеффенсена для операторних рівнянь за узагальнених умов Ліпшиця для поділених різниць першого порядку нелінійного оператора $F(x)$.

У статті [7] запропоновано двопараметричний метод типу хорд та вивчено локальну збіжність методу за узагальнених умов Ліпшиця для поділених різниць. Частковими випадками методу є відомі методи хорд, Курчатова, Ньютона.

У статтях [15, 23] досліджено метод з надквадратичною збіжністю за слабких умов Гьольдера та узагальнених умов Ліпшиця, визначено поділені різниці для конкретних операторів. Побудовано квадратичну нелінійну мажоранту для нелінійного оператора, відповідно до накладених на нього умов. Встановлено локальну та напівлокальну збіжність методу, знайдено апріорні і апостеріорні оцінки похибки методу та область єдиності розв'язку. У статті [7] запропоновано двокроковий параметричний метод, який містить деякі відомі методи з надквадратичною збіжністю. Отримані результати застосовані до чисельного розв'язування нелінійної крайової задачі другого порядку.

У статті [16] доведено збіжність методів Курчатова та Потра за умов Гьольдера для поділених різниць другого порядку та встановлено залежність порядку збіжності від сталої Гьольдера. На кінець, стаття [11] містить теоретичні результати з дослідження неточних різницевих методів.

Опишемо коротко структуру статті.

У першому пункті наведені означення поділених різниць та сформульовано різницеві методи розв'язування нелінійних рівнянь. Другий пункт присвячений обґрунтуванню локальної та напівлокальної збіжності методу хорд. У третьому пункті наведено основні результати збіжності методу лінійної інтерполяції Курчатова. У четвертому пункті вивчено збіжність ітераційного алгоритму із надквадратичною збіжністю в умовах неперервності за Гьольдером поділених різниць. У п'ятому пункті введено узагальнені умови Ліпшиця для поділених різниць першого порядку та за цих умов обгрунтувано збіжність різницевих методів. Всі теореми сформульовані в термінах поділених різниць. У шостому пункті зроблено короткі висновки.

1. Поділені різниці та різницеві ітераційні процеси

Для опису ітераційно-різницевих методів попередньо введемо деякі означення.

Нехай X, Y — два банахові простори, D — підмножина простору X, а x, y та z — три точки з D. Лінійний оператор з X в Y, позначуваний $F(x,y)$, називається *поділеною різницею першого порядку* від F за точками x і y, якщо він задовольняє умову

$$F(x,y)(x-y) = F(x) - F(y). \tag{2}$$

Поділеною різницею другого порядку від функції F за точками x, y та z називатимемо оператор $F(x,y,z)$, який задовольняє умову

$$F(x,y,z)(y-z) = F(x,y) - F(x,z). \tag{3}$$

Вважатимемо, що для $F(x,y)$ і $F(x,y,z)$ виконуються умови типу Ліпшиця у формі:

$$\|F(x,y) - F(x,z)\| \le p\,\|y-z\|, \quad x,y,z \in D, \tag{4}$$

$$\|F(y,x) - F(z,x)\| \le \bar{p}\,\|y-z\|, \quad x,y,z \in D, \tag{5}$$

$$\|F(x,y,z) - F(u,y,z)\| \le q\,\|x-u\|, \quad u,x,y,z \in D. \tag{6}$$

Якщо поділена різниця $F(x,y)$ від F задовольняє (4) або (5), тоді F диференційовний за Фреше на D і $F'(x) = F(x,x)$. Більше того, якщо (4) і (5) виконуються, тоді похідна Фреше неперервна за Ліпшицем на D з константою Ліпшиця $k = p + \bar{p}$ [20].

Будемо розглядати ітераційні процеси, які можна представити у вигляді

$$x_{n+1} = x_n - A_n^{-1}F(x_n), \quad n = 0, 1, \ldots. \tag{7}$$

В методі хорд

$$A_n = F(x_n, x_{n-1}). \tag{8}$$

Запропонований В.А. Курчатовим [4] метод містить

$$A_n = F(2x_n - x_{n-1}, x_{n-1}). \tag{9}$$

В дослідженому Потра методі [20] A_n є лінійною комбінацією поділених різниць від F у вигляді

$$A_n = F(x_n, x_{n-1}) + F(x_{n-2}, x_n) - F(x_{n-2}, x_{n-1}). \tag{10}$$

Зауважимо, що метод хорд і метод Курчатова вимагають два початкові наближення x_0, x_{-1}, а метод Потра — три початкові наближення x_0, x_{-1}, x_{-2}.

2. Збіжність методу хорд

При вивченні збіжності ітераційних процесів розглядають теореми двох типів — про локальну та напівлокальну збіжність. Перша теорема — теорема про локальну збіжність методу хорд

$$x_{n+1} = x_n - (F(x_n, x_{n-1}))^{-1} F(x_n), \quad n = 0, 1, \ldots \tag{11}$$

сформульована з умовами, накладеними на оператор в околі розв'язку, або, інакше, *з умовами типу Коші*. Таку збіжність називають *локальною*.

Теорема 1. *Нехай F — нелінійний оператор, який визначений у відкритій області D банахового простору X зі значеннями в банаховому просторі Y. Припустимо, що рівняння $F(x) = 0$ має розв'язок $x^* \in D$, і існує оборотна похідна Фреше $F'(x^*)$. Нехай F має в D поділену різницю першого порядку, яка задовольняє умову Ліпшиця*

$$\left\| F'(x^*)^{-1} \left(F(x, y) - F(u, v) \right) \right\| \le p_* \left(\|x - u\| + \|y - v\| \right). \tag{12}$$

Припустимо також, що відкрита куля $U = U(x^*, r_*)$ з центром x^* і радіусом

$$r_* = \frac{1}{3p_*} \tag{13}$$

така, що $U \subset D$. Тоді ітераційний процес (11) є коректно визначеним, і генерована ним послідовність, яка належить U, збігається до x^* і задовольняє нерівність

$$\|x_{n+1} - x^*\| \le \frac{p_* \|x^* - x_{n-1}\|}{1 - p_* \left(\|x^* - x_n\| + \|x^* - x_{n-1}\| \right)} \|x_n - x^*\|. \tag{14}$$

Інший тип збіжності — *напівлокальна збіжність*, яка передбачає формулювання більшості умов теореми в околі початкової точки. Такі умови називають ще *умовами типу Канторовича.*

Нами обґрунтовано напівлокальну збіжність методу хорд з використанням принципу мажорант Л. В. Канторовича. Досліджено метод хорд при різних умовах, накладених на нелінійний оператор. Так, при виконанні умови Ліпшиця для поділених різниць першого порядку побудована квадратична мажорантна функція однієї змінної, а при виконанні умови Ліпшиця для оператора другої поділеної різниці — кубічна мажорантна функція. Метод хорд, застосований до цих функцій, дає числові послідовності, які мажорують за нормою ітераційну послідовність, утворену від застосування методу хорд до нелінійного оператора. В обох випадках отримано апріорну та апостеріорну оцінки похибок методу хорд. Сформулюємо теорему про напівлокальну збіжність цього методу.

Теорема 2. *Нехай F — нелінійний оператор, який визначений на відкритій множині D банахового простору X зі значеннями в банаховому просторі Y. Нехай $F(\cdot, \cdot)$ поділена різниця першого порядку від F на множині D. Припустимо, що лінійний оператор*

$$A_0 = F(x_{-1}, x_0),$$

де $x_0, x_{-1} \in D$, *має обернений і існують невід'ємні числа* a, c *такі, що*

$$\|x_0 - x_{-1}\| \le a, \quad \|A_0^{-1}F(x_0)\| \le c. \tag{15}$$

Нехай в D *виконується така умова Ліпшиця*

$$\|A_0^{-1}(F(x,y) - F(u,v))\| \le p_0(\|x-u\| + \|y-v\|). \tag{16}$$

Нехай $p_0 a \le 1$, $r = \dfrac{1-p_0 a}{2p_0}$ *і* h — *дійсний поліном*

$$h(t) = -p_0 t^2 + (1-p_0 a)t.$$

Якщо задовольняється нерівність

$$c \le h(r) = \frac{(1-p_0 a)^2}{4p_0} \tag{17}$$

і замкнена куля $U_0 = U(x_0, r_0) \subset D$, *де* r_0 *є єдиним коренем рівняння* $h(t) = c$ *на* $(0, r]$, *тоді ітераційний процес (11) є коректно визначений і генерована ним послідовність збігається до розв'язку* x^* *рівняння* $F(x) = 0$. *Більше того, задовольняється нерівність*

$$\|x_n - x^*\| \le t_n, \qquad n = -1, 0, 1, 2, \ldots, \tag{18}$$

де

$$t_0 = r_0, \quad t_{-1} = r_0 + a, \quad p_0(a + 2r_0) < 1, \tag{19}$$

$$t_{n+1} = t_n \cdot \frac{p_0 t_{n-1}}{1 - p_0 a - 2p_0 r_0 + p_0(t_n + t_{n-1})}, \quad n = 0, 1, 2, \ldots. \tag{20}$$

Якщо відомі константи a, c, p_0, q_0, то ми можемо обчислити послідовність $\{t_n\}_{n\ge 0}$ перед отриманням послідовності $\{x_n\}_{n\ge 0}$ за ітераційним алгоритмом хорд. З допомогою нерівностей (18) даються апріорні оцінки похибки методу хорд. Нижче ми отримаємо апостеріорні оцінки похибки, які точніші за апріорні.

Теорема 3. *Нехай виконуються умови теореми 2. Позначимо*

$$e_n = p_0(\|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\|)\|x_n - x_{n-1}\|,$$

$$g_n = 1 - p_0 a - 2p_0\|x_n - x_0\|.$$

Тоді справедлива оцінка для $n = 1, 2, 3, \ldots$

$$\|x_n - x^*\| \le \frac{2e_n}{g_n + (g_n^2 - 4p_0 e_n)^{\frac{1}{2}}} \le t_n.$$

3. Збіжність методу лінійної інтерполяції Курчатова

Менш дослідженим є ітераційний метод лінійної інтерполяції, запропонований в [4] В. А. Курчатовим. Використовуючи два попередні наближення, як і метод хорд, метод Курчатова володіє квадратичною швидкістю збіжності. Однак дослідження методу в [4] проведене при досить жорстких умовах, зокрема вимагається обмеженість за нормою третьої похідної від нелінійного оператора. В цьому пункті ми вивчаємо локальну збіжність методу Курчатова; використовуючи принцип мажорант Л. В. Канторовича, доводимо напівлокальну збіжність методу, встановлюємо квадратичний порядок збіжності, знаходимо апріорну та апостеріорну оцінки похибки методу. При цьому вимагаємо виконання умови Ліпшиця для поділених різниць другого порядку. Всі дослідження проводяться в термінах поділених різниць. Подібні результати пізніше отримані в [19]. Відзначимо, що при таких умовах Ф. А. Потра дослідив різницевий метод [20], який використовує інформацію з трьох попередніх ітерацій, однак порядок збіжності його нижчий і дорівнює 1,839. . . .

Розглянемо теорему про локальну збіжність методу лінійної інтерполяції Курчатова.

Теорема 4. *Нехай F — нелінійний оператор, який визначений у відкритій опуклій області D банахового простору X зі значеннями в банаховому просторі Y. Припустимо, що рівняння $F(x) = 0$ має розв'язок $x^* \in D$ і існує оборотна похідна Фреше $F^{'}(x^*)$. Нехай F має в області $V = \{x : \|x - x^*\| < 3r_*\} \subseteq D$ поділені різниці першого та другого порядку, які задовольняють умови Ліпшиця*

$$\|F^{'}(x^*)^{-1}(F(x, y) - F(u, v))\| \leq p_*(\|x - u\| + \|y - v\|), \tag{21}$$

$$\|F^{'}(x^*)^{-1}(F(u, x, y) - F(v, x, y))\| \leq q_*\|u - v\|, \tag{22}$$

де

$$r_* = \frac{2}{3p_* + \sqrt{9p_*^2 + 32q_*}}. \tag{23}$$

Тоді для всіх $x_0, x_{-1} \in U = \{x : \|x - x^*\| < r_*\}$ ітераційний процес

$$x_{n+1} = x_n - (F(2x_n - x_{n-1}, x_{n-1}))^{-1}F(x_n), \; n = 0,1, \ldots \tag{24}$$

коректно визначений і генерована ним послідовність $\{x_n\}_{n\geq0}$, яка належить U, збігається до x^* і задовольняє нерівність

$$\|x_{n+1} - x^*\| \leq \frac{p_*\|x_n - x^*\| + q_*\|x_n - x_{n-1}\|^2}{1 - 2p_*\|x_n - x^*\| - q_*\|x_n - x_{n-1}\|^2}\|x_n - x^*\|. \tag{25}$$

Наслідок 1. *Порядок збіжності ітераційної процедури Курчатова (24) квадратичний.*

Встановимо напівлокальну збіжність методу Курчатова.

Теорема 5. *Нехай F — нелінійний оператор, який визначений на відкритій опуклій множині D банахового простору X зі значеннями в банаховому просторі Y. Нехай $F(\cdot, \cdot)$ і $F(\cdot, \cdot, \cdot)$ поділені різниці першого і другого порядку від F на множині $V_0 = \{x : \|x - x_0\| \leq 3r_0\} \subset D$. Припустимо, що лінійний оператор $A_0 = F(2x_0 - x_{-1}, x_{-1})$, де $x_0, x_{-1} \in U_0 = \{x : \|x - x_0\| \leq \leq r_0\}$, є оборотний і задовольняє такі умови Ліпшиця*

$$\|A_0^{-1}(F(x, y) - F(u, v))\| \leq p_0(\|x - u\| + \|y - v\|), \tag{26}$$

$$\|A_0^{-1}(F(x, y, z) - F(u, y, z))\| \leq q_0\|x - u\|. \tag{27}$$

Визначимо два невід'ємні числа a і c, такі що

$$\|x_0 - x_{-1}\| \leq a, \qquad \|A_0^{-1} F(x_0)\| \leq c. \tag{28}$$

Припустимо, що $2q_0a^2 \leq 1$ і нехай

$$s = \left\{(p_0 + q_0 a)^2 + 3q_0(1 - q_0 a^2)\right\}^{1/2}, \qquad r = \frac{1 - q_0 a^2}{p_0 + q_0 a + s}$$

та h — дійсний поліном

$$h(t) = -q_0 t^3 - (p_0 + q_0 a)t^2 + (1 - q_0 a^2)t.$$

Якщо задовольняється нерівність

$$c(1 - 2q_0 a^2) \leq h(r) = \frac{1}{3} \cdot (p_0 + q_0 a + 2s)\Big(\frac{1 - q_0 a^2}{p_0 + q_0 a + s}\Big)^2 \tag{29}$$

і замкнена куля $V_0 \subset D$, де $r_0 \in (0, r]$ є коренем рівняння $h(t) = c(1-2q_0a^2)$, то ітераційний процес (24) коректно визначений і генерована ним послідовність збігається до розв'язку x^* рівняння (1). Більше того, справедлива нерівність

$$\|x_n - x^*\| \leq t_n, \qquad n = -1,0,1,2,\cdots, \tag{30}$$

де

$$t_0 = r_0,\ t_{-1} = r_0 + a, \tag{31}$$

$$a_0 = p_0 + 3q_0 r_0 + q_0 a,\ b_0 = 3q_0 r_0^2 - 2a_0 r_0 - q_0 a^2 + 1, \tag{32}$$

$$t_{n+1} = t_n \cdot \frac{a_0 t_n - q_0(t_n - t_{n-1})^2 - 2q_0 t_n^2}{b_0 + 2a_0 t_n - q_0(t_n - t_{n-1})^2 - 3q_0 t_n^2},\ n = 0,1,2,\ldots \tag{33}$$

Аналогічно встановлено локальну та напівлокальну збіжність методу Курчатова за умов Ліпшиця на поділені різниці першого порядку та умов Гьольдера на поділені різниці другого порядку, які мають вигляд

$$\left\|F'(x^*)^{-1}\left(F(x,y) - F(u,v)\right)\right\| \leq p_* \left(\|x-u\| + \|y-v\|\right), \tag{34}$$

$$\left\|F'(x^*)^{-1}\left(F(u,x,y) - F(v,x,y)\right)\right\| \leq q_* \left\|u - v\right\|^{p}, \quad p \in (0,\ 1]. \tag{35}$$

З теореми про локальну збіжність випливає

Наслідок 2. *Порядок збіжності методу (24) дорівнює єдиному додатному кореню рівняння $m^2 - m - (p+1) = 0$: $m_K = \dfrac{1+\sqrt{4p+5}}{2}$.*

Також за умов (34) та (35) вивчено збіжність методу Потра

$$x_{n+1} = x_n - \left[F(x_n, x_{n-1}) + F(x_{n-2}, x_n) - F(x_{n-2}, x_{n-1})\right]^{-1} F(x_n), \quad n = 0,\ 1,\ \ldots\ . \tag{36}$$

З теореми про локальну збіжність методу (36) отримуємо [16]

Наслідок 3. *Порядок збіжності ітераційного процесу (36) m_P є коренем рівняння $m^3 - m^2 - m - p = 0$.*

Зауважимо, що при $p = 1$ порядок збіжності методу (36) дорівнює 1,839... та одержуються відповідні теореми з [20].

З табл. 1 видно, що зі зменшенням сталої Гьольдера порядок збіжності ітераційних методів знижується. В таблиці m_S — порядок збіжності методу хорд (11), який дорівнює $(1+\sqrt{1+4p})/2$, де p — стала з умови Гьольдера для поділених різниць першого порядку.

Табл. 1. Залежність порядку збіжності методів (24),(36) і (11) від сталої Гьольдера p

p	m_K	m_P	m_S
0,0010	1,6184	1,6183	1,000
0,0625	1,6456	1,6350	1,059
0,1250	1,6726	1,6513	1,112
0,2500	1,7247	1,6826	1,207
0,5000	1,8228	1,7399	1,366
0,7500	1,9142	1,7917	1,500
1,0000	2	1,8392	1,618

4. Ітераційний алгоритм із надквадратичною збіжністю в умовах неперервності за Гьольдером поділених різниць

Нехай F — нелінійний оператор, визначений на випуклій множині D банахового простору X зі значеннями в банаховому просторі Y. Для розв'язування рівняння (1) розглянемо алгоритм

$$\begin{aligned} x_{n+1} &= x_n - [F(x_n, y_n)]^{-1} F(x_n), \\ y_{n+1} &= x_{n+1} - [F(x_n, y_n)]^{-1} F(x_{n+1}),\ n = 0,1,\dots, \end{aligned} \tag{37}$$

де для кожних $x_n, y_n \in D$, $F(x_n, y_n)$ — обмежений лінійний оператор з X в Y, x_0, y_0 — задані. Кількість обчислень на кожній ітерації методу (37) є практично така сама, як у класичному методі хорд.

При дослідженні методу (37) вимагалось обмеженість норми поділеної різниці другого порядку оператора F або неперервності за Ліпшицем (Гьольдером) поділеної різниці другого порядку [3, 28].

Ми проводимо дослідження локальної збіжності методу (37) (умови типу Коші) і напівлокальної збіжності (умови типу Канторовича) при слабших умовах, ніж в усіх інших відомих працях. Зокрема, ми вимагаємо лише неперервність за Гьольдером поділених різниць першого порядку, так як і для класичного методу хорд [10, 18]. Отримана формула залежності порядку збіжності від сталої Гьольдера. Крім того, ми вперше використовуємо методику мажорант для дослідження напівлокальної збіжності методу (37). Також доведено теорему про єдиність розв'язку.

Розглядаємо відкриту опуклу множину D простору X і припускаємо, що F неперервно диференційована за Фреше в D. Вважаємо також, що поділена різниця $F(x, y)$ задовольняє умову Гьольдера, якщо існує невід'ємна стала k така, що

$$\|F(x,y)-F(u,v)\| \le k(\|x-u\|^\alpha + \|y-v\|^\alpha),\ \alpha \in (0,1] \tag{38}$$

для $x, y, u, v \in D$ з $x \neq y$ і $u \neq v$. У цьому випадку ми кажемо, що F має на D неперервну за Гьольдером поділену різницю. При цьому відомо [18], що існує похідна Фреше від F в D і вона задовольняє $F(x,x) = F'(x),\ x \in D$.

Сформулюємо теорему про локальну збіжність ітераційного процесу (37).

Теорема 6. *Нехай F — нелінійний оператор, визначений на відкритій опуклій множині D банахового простору X зі значеннями в банаховому просторі Y. Припустимо, що рівняння $F(x) = 0$ має розв'язок $x^* \in D$ і існує оборотна похідна Фреше $F'(x^*)$. Нехай F має в D поділені різниці першого порядку, які задовольняють умову Гьольдера*

$$\begin{aligned}&\|F'(x^*)^{-1}(F(x,y)-F(u,v))\| \le p_*(\|x-u\|^\alpha + \|y-v\|^\alpha),\\ &\alpha \in (0;1],\ p_* \ge 0.\end{aligned} \tag{39}$$

Припустимо, що відкрита куля $U = U(x^*, r_*)$ з центром x^* і радіусом

$$r_*^\alpha = \frac{1}{(2+\sqrt{1+2^\alpha})p_*} \tag{40}$$

така, що $U \subset D$.

Тоді для $x_0, y_0 \in U$ ітераційний процес (37) коректно визначений і генеровані ним послідовності $\{x_n\}_{n\ge0}$ і $\{y_n\}_{n\ge0}$, які належать U, збігаються до x^* і задовольняють нерівностям

$$\|x_{n+1}-x^*\| \le \frac{p_*\|y_n - x^*\|^\alpha}{1-p_*(\|x_n-x^*\|^\alpha + \|y_n-x^*\|^\alpha)}\|x_n-x^*\|, \tag{41}$$

$$\|y_{n+1}-x^*\| \le \frac{p_*(\|x_{n+1}-x_n\|^\alpha + \|y_n-x^*\|^\alpha)}{1-p_*(\|x_n-x^*\|^\alpha + \|y_n-x^*\|^\alpha)}\|x_{n+1}-x^*\|. \tag{42}$$

Зауваження 1. Радіус збіжності r^*, отриманий у цій теоремі, є дещо менший, ніж радіус збіжності методу хорд [18] $r_*^\alpha = \dfrac{1}{3p_*}$.

Наслідок 4. *Ітераційний процес (37) збігається локально до нуля функції F з порядком збіжності рівним принаймні* $\dfrac{1+\alpha}{2} + \sqrt{\left(\dfrac{1+\alpha}{2}\right)^2 + \alpha^2}$.

Зауваження 2. Порядок збіжності методу (37) у випадку $\alpha = 1$ дорівнює $1+\sqrt{2}$.

Теорема про напівлокальну збіжність методу (37) має такий вигляд.

Теорема 7. *Нехай F — нелінійний оператор, визначений на відкритій опуклій множині D банахового простору X зі значеннями в банаховому просторі Y. Припустимо, що:*

1) лінійний оператор $A_0 = F(x_0, y_0)$, де $x_0,\ y_0$ $(x_0 \neq y_0)$ є дві точки з D, має обернений,

2) a, c і p_0 — три невід'ємні числа такі, що

$$\|x_0 - y_0\| \le a,\ \|A_0^{-1}F(x_0)\| \le c,\ c > a \tag{43}$$

і в D виконується така умова Гьольдера ($\alpha \in (0,1]$)

$$\|A_0^{-1}(F(x,y) - F(u,v))\| \le p_0(\|x-u\|^\alpha + \|y-v\|^\alpha). \tag{44}$$

3) r_0 — невід'ємне число таке, що

$$r_0 \ge c/(1-\gamma), \ \ p_0(r_0^\alpha + 2(r_0-a)^\alpha) < 1, \tag{45}$$

де $\gamma = p_0(r_0-a)^\alpha/(1-[r_0^\alpha + (r_0-a)^\alpha])$ (зауважимо, що $0 \le \gamma < 1$),

4) замкнена куля $U_0 = U(x_0, r_0)$ міститься в D.

Тоді:

(i) дійсні послідовності $\{t_n\}_{n\ge0}, \{s_n\}_{n\ge0}$, визначені як

$$t_0 = r_0, \ s_0 = r_0 - a, \ t_1 = r_0 - c, \tag{46}$$

і для $k \ge 0$

$$t_{k+1} - t_{k+2} = \frac{p_0(s_k - t_{k+1})^\alpha}{1 - p_0[(t_0 - t_{k+1})^\alpha + (s_0 - s_{k+1})^\alpha]}(t_k - t_{k+1}) = B_{k+2}(t_k - t_{k+1}), \tag{47}$$

і

$$t_{k+1} - s_{k+1} = \frac{p_0(s_k - t_{k+1})^\alpha}{1 - p_0[(t_0 - t_k)^\alpha + (s_0 - s_k)^\alpha]}(t_k - t_{k+1}) = C_{k+1}(t_k - t_{k+1}) \tag{48}$$

невід'ємні, спадні і збігаються до деякого $t^* \in R$ такого, що

$$r_0 - c/(1-\gamma) \le t^* < t_0.$$

(ii) Ітераційний процес (37) коректно визначений і утворені ним послідовності $\{x_n\}_{n\ge0}$ і $\{y_n\}_{n\ge0}$ збігаються до розв'язку $x^* \in U(x_0, r_0)$ рівняння $F(x) = 0$. Крім того, справедливі нерівності

$$\|x_n - x^*\| \le t_n - t^*, \ \ \|y_n - x^*\| \le s_n - t^*, \qquad n = 0,1,2,\ldots, \tag{49}$$

і для $n \ge 1$

$$\|x_n - x^*\| \le \frac{p_0(s_{n-1} - t_n)^\alpha}{1 - p_0[(t_0 - s_0)^\alpha + (t_0 - t_n)^\alpha + (t_0 - t^*)^\alpha]}(t_{n-1} - t_n). \tag{50}$$

Тепер можемо отримати результат єдиності розв'язку.

Теорема 8. *Нехай F — нелінійний оператор, визначений на відкритій опуклій множині D банахового простору X зі значеннями в банаховому просторі Y. Припустимо, що:*

1) виконуються умови теореми 7;

2) r_0 з теореми 7 додатково задовольняє умову

$$2p_0 r_0^\alpha + (r_0 - a)^\alpha + p_0 a^\alpha < 1. \tag{51}$$

Тоді ітераційний алгоритм (37) коректно визначений і генерована ним послідовність $\{x_n\}$ належить $U(x_0, r_0)$ і збігається до єдиного розв'язку x^* рівняння $F(x) = 0$ в $U(x_0, r_0)$.

На основі досліджень, проведених за єдиним підходом, можна встановити співвідношення між порядками та радіусами збіжності розглянутих методів. В табл. 2 наведено порядки збіжності та радіуси збіжності досліджуваних методів.

З таблиці 2 бачимо, що найвищий порядок збіжності має метод (37), а найнижчий — метод хорд. Одночасно метод хорд має найбільший радіус збіжності, який збігається з радіусом збіжності методу Ньютона [20]. Методи, запропоновані Потра та Курчатовим, мають середні показники як за швидкістю збіжності, так і за радіусом області збіжності. Зауважимо, що отримані оцінки розраховані на найгірший випадок. Вони гарантують збіжність методів зі встановленими порядками збіжності. Реальні області збіжності розглянутих методів можуть бути ширшими ніж указані в теоремах.

Табл. 2. Зв'язок між порядком збіжності і радіусом збіжності методів (11), (36), (24), (37) та Ньютона

Метод	Порядок m	Радіус r
хорд	1,618	$\dfrac{1}{3p_*}$
Потра	1,939	$\dfrac{2}{3p_* + \sqrt{9p_*^2 + 16q_*}}$
Курчатова	2,000	$\dfrac{2}{3p_* + \sqrt{9p_*^2 + 32q_*}}$
метод (37)	2,414	$\dfrac{1}{(2+\sqrt{3})p_*}$
Ньютона	2	$\dfrac{1}{3p_*}$

5. Збіжність різницевих методів за узагальнених умов Ліпшиця

У праці [27] при дослідженні методу Ньютона запропоновано узагальнені умови Ліпшиця для оператора похідної, в яких замість константи L використано деяку додатну інтегровну функцію.

У статті [14] ми ввели аналогічні узагальнені умови Ліпшиця для оператора поділеної різниці першого порядку і за цих умов обґрунтували збіжність різницевих методів, які досліджувались вище за умов Ліпшиця (Гьольдера) зі сталою L. Проте L в умовах Ліпшиця не обов'язково має бути константою, а може бути додатною інтегровною функцією.

Позначимо $B(x_0, r) = \{x : \|x - x_0\| < r\}$ — кулю радіуса r з центром в точці x_0. Умови

$$\|F(x,y) - F(u,v)\| \le \int_0^{\|x-u\|+\|y-v\|} L(u)du \quad \forall x,y,u,v \in B(x_0,r) \tag{52}$$

та

$$\|F(x,y)-F'(x_0)\| \le \int_0^{\|x-x_0\|+\|y-x_0\|} L(u)du \quad \forall x,y \in B(x_0,r). \qquad (53)$$

називатимемо узагальненими умовами Ліпшиця або такими, що містять L у середньому.

У цьому пункті ми розглядаємо двопараметричний клас методів типу хорд [7]

$$x_{n+1} = x_n - F(u_n,v_n)^{-1}F(x_n),\ n = 0,\ 1,\ 2,\ldots, \qquad (54)$$

де $u_n = x_n + a_n(x_{n-1}-x_n)$, $v_n = x_n + b_n(x_{n-1}-x_n)$, $a_n \in [-1;1]$, $b_n \in [0;1]$. Поклавши $a_n = b_n = 0$, отримаємо класичний метод Ньютона, поклавши $a_n = 0$, $b_n = 1$, матимемо метод хорд (11), за $a_n = -1$, $b_n = 1$ — метод лінійної інтерполяції Курчатова [4]. Як бачимо, запропонований метод (54) містить як методи Ньютона та Курчатова, для яких доведена квадратична збіжність, так і методи типу хорд з дещо нижчою збіжністю.

За узагальнених умов Ліпшиця досліджено збіжність методу типу хорд (54). Показано надлінійну швидкість збіжності ітераційного процесу (з порядком $(1+\sqrt{5})/2$) для постійних параметрів, для незростаючої послідовності $\{b_n\}$ і постійного a_n, а при певному виборі параметрів одержано і квадратичну збіжність.

Теорема 9. ***Н****ехай F — нелінійний оператор, визначений у відкритій опуклій області D банахового простору X зі значеннями в банаховому просторі Y. Припустимо, що:*

1) рівняння (1) має розв'язок x^* в D, існує похідна за Фреше $F'(x^*)$ і вона є оборотна;

2) в кулі $B(x^*,r) = \{x \in D : \|x-x^*\| \le r\}$ функція $F(x)$ має поділені різниці першого порядку $F(x,y)$, які задовольняють умову Ліпшиця з усередненим L:

$$\| F'(x^*)^{-1}F(x,y) - I \| \le \int_0^{\|x-x^*\|+\|y-x^*\|} L(z)dz, \qquad (55)$$

де $x,\ y \in B(x^*,r)$ і L — неспадна;

3) $r > 0$ задовольняє нерівність

$$\frac{\int_0^{(1+2|a_0|)r} L(z)dz}{1-\int_0^{(2-a_0+|a_0|)r} L(z)dz} \le 1. \qquad (56)$$

Тоді послідовність (54) збігається для всіх $x_{-1},\ x_0 \in B(x^*,r)$ і

$$\| x_{n+1}-x^* \| \le q_1 \| x_{n-1}-x^* \| \| x_n - x^* \|, \qquad (57)$$

де

$$\rho_{max} = \max\{\rho(x_{-1}),\rho(x_0)\},$$

$$q_1 = \frac{\int_0^{(1+2|a_0|)\rho_{\max}} L(z)dz}{\left(1-\int_0^{(2-a_0+|a_0|)\rho_{\max}} L(z)dz\right)\rho_{\max}}.$$

Наслідок 5. *Порядок збіжності ітераційного процесу (54) дорівнює* $\left(1+\sqrt{5}\right)/2$.

Зауваження 3. Вибравши $a_n = 0$ та $b_n = O(\| x_n - x^* \|)$, можна довести квадратичну збіжність процесу (54).

Результати дослідження умов та швидкості збіжності двокрокового методу (37) наведені в такій теоремі.

Теорема 10. *Нехай F– нелінійний оператор, визначений на відкритій опуклій множині D банахового простору X зі значеннями у банаховому просторі Y. Припустимо, що: 1) рівняння $F(x) = 0$ має розв'язок $x^* \in D$, існує похідна Фреше $F'(x^*)$ і вона є оборотна; 2) F має поділені різниці $F(x, y)$ в $B(x^*, r) \subset D$, які задовольняють центральну умову Ліпшиця з L в середньому*

$$\left\|F'(x^*)^{-1}(F(x^*, x^*) - F(x, y))\right\| \le \int_0^{\rho(x)+\rho(y)} L(u)du \tag{58}$$

де $x, y \in B(x^*, r)$, $\rho(x) = \|x - x^*\|$ і L — неспадна функція; 3) r задовольняє нерівність

$$\int_0^{3r} L(u)du \Big/ \left(1 - \int_0^{2r} L(u)du\right) \le 1. \tag{59}$$

Тоді метод (37) збігається для всіх $x_0, y_0 \in B(x^*, r)$ таких, що $\rho(y_0) < \rho(x_0)$, і

$$\|x_{n+1} - x^*\| \le \frac{\int_0^{\rho(y_n)} L(u)du\rho(x_n)}{1 - \int_0^{\rho(x_n)+\rho(y_n)} L(u)du} \le \frac{q_1}{\rho(x_0)}\rho(x_n)\rho(y_n); \tag{60}$$

$$\|y_{n+1} - x^*\| \le \frac{\int_0^{\|x_{n+1}-x_n\|+\rho(y_n)} L(u)du\rho(x_{n+1})}{1 - \int_0^{\rho(x_n)+\rho(y_n)} L(u)du} \le \frac{q_2}{\rho(x_1)}\rho(x_{n+1})\rho(x_n), \tag{61}$$

де величини

$$q_1 = \frac{\int_0^{\rho(x_0)} L(u)du}{1 - \int_0^{\rho(x_0)+\rho(y_0)} L(u)du}, \quad q_2 = \frac{\int_0^{\rho(x_1)+\rho(x_0)+\rho(y_0)} L(u)du}{1 - \int_0^{\rho(x_0)+\rho(y_0)} L(u)du} \tag{62}$$

менші за 1, $n = 0, 1, \dots$.

Область єдиності розв'язку встановлено в теоремі.

Теорема 11. *Нехай $F(x^*) = 0$, існує $F'(x^*)^{-1}$, F має поділені різниці $F(x, x^*)$ в $B(x^*, r)$, які задовольняють радіальну умову Ліпшиця з L в середньому*

$$\left\|F'(x^*)^{-1}F(x, x^*) - I\right\| \le \int_0^{\rho(x)} L(u)du \qquad \forall x \in B(x^*, r), \tag{63}$$

де $\rho(x) = \|x - x^*\|$ і L — додатна інтегровна функція. Нехай r задовольняє нерівність $\int_0^r L(u)du \le 1$.

Тоді рівняння (1) має єдиний розв'язок x^* в $B(x^*, r)$.

7. Висновки

У цій статті наведено результати дослідження низки відомих ітераційно-різницевих методів з порядками збіжності від $(1+\sqrt{5})/2$ до $1+\sqrt{2}$ для розв'язування нелінійних операторних рівнянь: методу хорд, методу лінійної інтерполяції Курчатова, методу Потра та методу з порядком збіжності $1+\sqrt{2}$. Дослідження методів проведено за єдиною схемою, при досить слабких умовах, що дозволило порівняти методи між собою. Зокрема, методи вивчались при класичних умовах Ліпшиця, при умовах Гьольдера, при узагальнених умовах Ліпшиця, які накладались на поділені різниці першого або другого порядку. Отримано залежність порядку збіжності від сталих Гьольдера. Запропоновано та обгрунтовано параметричний метод типу хорд, який містить як часткові випадки низку відомих методів.

Література

1. Бартиш М. Я. Исследование параметрических итерационних процессов для решения нелинейных уравнений / М. Я.Бартіш, С. М. Шахно // Проблемы управления и информатики. — 1997. — №2. — С. 22–30.
2. Бартіш М. Я. Про один ітераційний метод розв'язування функціональних рівнянь / М. Я. Бартіш. — Допов. АН УРСР. Сер. А. — 1968. — № 5. — С. 387–391.
3. Бартіш М. Я. Про один різницевий метод розв'язування нелінійних операторних рівнянь / М. Я. Бартіш, Ю. М. Щербина // Доп. АН УРСР. Сер. А. — 1972. — №7. — С. 579–582.
4. Курчатов В. А. Об одном методе линейной интерполяции решения функциональных уравнений / В. А. Курчатов // Докл. АН СССР. Сер. Математика. Физика. — 1971. — Т. 198, № 3. — С. 524–526.
5. Трауб Дж. Итерационные методы решения уравнений / Дж. Трауб: Пер. с англ. — М.: Мир, 1985. — 264с.
6. Ульм С. Об обобщенных разделенных разностях / С. Ульм // Известия АН ЭССР. Физика. Математика. — 1967. — Т. 16. — С. 13–26.
7. Шахно С. Двопараметричні методи типу хорд для розв'язування нелінійних рівнянь / С. Шахно, С. Граб, Г. Ярмола // Вісник Львів. ун-ту. Сер. прикл. мат. та інформ. — 2009. — Вип. 15. — С. 117–127.
8. Шахно С. М. Дво- і трикрокові ітераційні процеси для розв'язування нелінійних рівнянь / С. М. Шахно, О. М. Макух // Вісник Львів. ун-ту. Сер. прикл. мат. та інформ. — 2006. — Вип. 11. — С. 99–106.
9. Шахно С. М. Дослідження ньютонівських ітераційних процесів, що використовують внутрішні ітерації / С. М. Шахно // Математ. методи і фіз.-мех. поля. — 1999. — Т. 42, №1. — С. 39–44. — Англ. переклад: Journal of Mathematical Sciences. — 2002.– V. 109. — №1. — P. 1203–1208.
10. Шахно С. М. Застосування нелінійних мажорант для дослідження методу хорд розв'язування нелінійних рівнянь / С. М. Шахно // Математичні студії. — 2004. — Т.22. — № 1. — С. 79–86.
11. Шахно С. М. Збіжність неточних різницевих методів при узагальнених умовах Ліпшиця / С. М. Шахно // Математ. методи і фіз.-мех. поля. — 2009. — Т. 52.– № 3. — С. 30–40. — Англ. переклад: Journal of Mathematical Sciences. — 2010.– V. 171, №4. — P. 453–465.

12. Шахно С. Локальна збіжність ітераційно-різницевих методів розв'язування нелінійних операторних рівнянь / С. Шахно, О. Макух // Вісник Львів. ун-ту. Сер. прикл. мат. та інф. — 2003. — Вип. 7. — С. 124–131.
13. Шахно С. М. Метод Стефенсена за узагальнених умов Ліпшиця для поділених різниць першого порядку / С. М. Шахно // Математичні студії. — 2009. — Т. 31, № 2. — С. 90–95.
14. Шахно С. М. Метод хорд при узагальнених умовах Ліпшиця для поділених різниць першого порядку / С. М. Шахно // Матем. вісник НТШ — 2007. — Т. 4. — С. 296–305.
15. Шахно С. М. Про двокроковий ітераційний процес в узагальнених умовах Ліпшиця для поділених різниць першого порядку / С. М. Шахно // Математ. методи і фіз.-мех. поля — 2009. — Т. 52, № 1. — С. 59–66. — Англ. переклад: Journal of Mathematical Sciences — 2010. — V. 168, №4. — P. 576–584.
16. Шахно С. М. Про ітераційні методи в умовах неперервності за Гельдером поділених різниць другого порядку / С. М. Шахно, О. М. Макух // Математ. методи і фіз.-мех. поля — 2006. — Т. 49, № 2. — С. 90–98.
17. Шахно С. М. Про різницевий метод з квадратичною збіжністю для розв'язування нелінійних операторних рівнянь / С. М. Шахно // Математичні студії. — 2006. — Т. 26, № 1. — С. 105–110.
18. Argyros I. K. On an iterative algorithm for solving nonlinear operator equations / I. K. Argyros // Zeitschrift für Analysis und ihre Anwendungen. — 1991. — V. 10, №1. — P. 83–92.
19. Argyros I.K. A Kantorovich-type analysis for a fast iterative method for solving nonlinear equations / I.K. Argyros // J. Math. Anal. Appl. — 2007. — V. 332. — P. 97–108.
20. Potra F.A. On an iterative algorithm of order 1.839…for solving nonlinear operator equation / F.A. Potra // Numer. Funct. Anal. And Optimiz. — 1984–85. — 7(1). — P. 75–106.
21. Shakhno S.M. Method of order $1+\sqrt{2}$ for the solution of nonlinear equations with Hölder continuous divided differences / S.M. Shakhno // Proc. Appl. Math. Mech. — 2005. — V. 5. — P. 779–780.
22. Shakhno S. M. On a Kurchatov's method of linear interpolation for solving nonlinear equations / S. M. Shakhno // Proc. Appl. Math. Mech. — 2004. — V. 4. — P. 650–651.
23. Shakhno S. M. On an iterative algorithm with superquadratic convergence for solving nonlinear operator equations / S. M. Shakhno // Journal of Computational and Applied Mathematics. — 2009. — V. 231. — P. 222–235.
24. Shakhno S. M. On an iterative algorithm of order 1.839... for solving nonlinear least squares problems / S. M. Shakhno, O. P. Gnatyshyn // Applied Mathematics and Computation. — 2005. — V. 161, №1. — P. 253–264.
25. Shakhno S. M. On the Secant method under generalized Lipschitz conditions for the divided difference operator / S. M. Shakhno // Proc. Appl. Math. Mech. — 2007. — V. 7. — P. 2060083–2060084.
26. Shakhno S. M. On the Steffensen method under the generalized Lipschitz conditions for the divided difference operator / S. M. Shakhno // Proc. Appl. Math. Mech. — 2008. — V. 8. — P. 10855-10856.
27. Wang X. Convergence of Newton's method and uniqueness of the solution of equations in Banach space / X. Wang // IMA Journal of Numerical Analysis. — 2000. — V. 20. — P. 123–134.

28. Werner W. Über ein Verfahren der Ordnung $1+\sqrt{2}$ zur Nullstellenbestimmung / W. Werner // Numer. Math. — 1979 — V. 32. — P. 333–342.

Факультет прикладної математики та інформатики, Львівський національний університет імені Івана Франка, вул. Університетська, 1, Львів, 79000, Україна.

Надійшла 05.03.2012

Журнал обчислювальної та прикладної математики | 2012, №1(107) | Journal of Computational & Applied Mathematics

УДК 519.71

СТАТИСТИКА ВКЛЮЧЕННЯ

В. В. Алєксєєнко

Резюме. Розглядається нова міра близькості між вибірками, яка базується на порівнянні наближених за допомогою Гіпотези Хілла функцій розподілу еталонної вибірки і еталонної доповненої елементами тестової. Запропонована міра близькості по суті схожа на p-статистику і при цьому має більш низький рівень складності алгоритму обчислення.

Вступ

Нехай $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_m)$ — вибірки, отримані простим випадковим вибором із генеральних сукупностей G та H відповідно. Для перевірки гіпотези про те, що генеральні сукупності G та H збігаються (функції розподілу $F_G = F_H$) можна застосувати p-статистику. В реальних задачах маємо виміри якоїсь характеристики об'єктів дослідження і певні еталонні значення для визначених об'єктів, що є задачею класифікації.

Мета роботи — побудувати нову статистику для визначення міри близькості двох вибірок. Основна ідея побудови статистики — доповнення еталонної вибірки елементами тестової, побудова наближеної функції розподілу для еталонної і доповненої вибірки та побудова нормованої міри близькості між функціями розподілу.

Для порівняння наводиться p-статистика, оскільки нова міра близькості так само базується на гіпотезі Хілла. Однак складність алгоритму обчислення нової статистики менша, ніж у p-статистиці.

Статистика включення.

Нехай $x_1, x_2, ..., x_n$ — вибірка, отримана простим випадковим вибором із генеральної сукупності G, породженої неперервною випадковою величиною із функцією розподілу F_G, $x_{(1)} < x_{(2)} < ... < x_{(n)}$ — відповідний варіаційний ряд.

Можемо розглянути наближення F_G у значеннях вибірки згідно гіпотези Хілла:

$$F_G^{(1)}\left(x_{(i)}\right) = \frac{i}{n+1}. \tag{1}$$

Нехай $y_1, y_2, ..., y_m$ — вибірка, отримана простим випадковим вибором із генеральної сукупності H, породженої неперервною випадковою величиною із функцією розподілу F_H. Припустимо, що $F_G = F_H$. Тоді можемо припустити, що $z_1, z_2, ..., z_{m+n} = x_1, x_2, ..., x_n, y_1, y_2, ...y_m$ — вибірка, отримана простим випадковим вибором із генеральної сукупності G, $z_{(1)} < z_{(2)} <$ $< ... < z_{(m+n)}$ — відповідний варіаційний ряд.

Можемо розглянути наближення F_G у значеннях вибірки згідно гіпотези Хілла:

$$F_G^{(2)}\left(z_{(i)}\right) = \frac{i}{n+m+1}. \tag{2}$$

Оскільки $\left\{x_i, i = \overline{1,n}\right\} \subset \left\{z_i, i = \overline{1,m+n}\right\}$, то таким чином визначено і $F_G^{(2)}\left(x_{(i)}\right)$.

Позначимо

$$I_0 = \left(-\infty; x_{(1)}\right), I_n = \left(x_{(n)}, \infty\right), I_j = \left(x_{(j)}, x_{(j+1)}\right) | j = \overline{1,n-1}. \tag{3}$$

Нехай ξ_1 — неперервна випадкова величина з функцією розподілу $F_G^{(1)}$, ξ_2 — неперервна випадкова величина з функцією розподілу $F_G^{(2)}$.

Позначимо

$$p_j^{(1)} = P\left(\xi_1 \in I_j\right), p_j^{(2)} = P\left(\xi_2 \in I_j\right). \tag{4}$$

Побудуємо статистику, що характеризуватиме близькість цих випадкових величин:

$$d = \sum_{i=0}^{n} \left|p_i^{(1)} - p_i^{(2)}\right|. \tag{5}$$

Позначимо l_i — кількість $y_j | y_j \in I_j$.

Лема 1. *В рамках припущень справедлива рівність* $p_j^{(2)} = \frac{1+l_j}{m+n+1}$.

Доведення. $p_0^{(2)} = F_G^{(2)}\left(x_{(0)}\right) = F_G^{(2)}\left(z_{(l_0+1)}\right) = \frac{l_0+1}{m+n+1}$.

Нехай $x_{(j)} = z_{(a)} | j = \overline{1,n-1}$, тоді в рамках позначень $x_{(j+1)} = z_{(a+l_j+1)}$,

$$p_j^{(2)} = F_G^{(2)}\left(z_{(a+l_j+1)}\right) - F_G^{(2)}\left(z_{(a)}\right) = \frac{l_j+1}{m+n+1},$$

$$p_n^{(2)} = 1 - F_G^{(2)}\left(z_{(m+n-l_n)}\right) = \frac{l_n+1}{m+n+1}.$$

□

Наслідок 1. *Вектор* $\bar{l} = (l_0, l_1, ..., l_n)$ *однозначно визначає значення* d.

Враховуючи сутність позначення l_i, справедливо

$$l_i \geqslant 0; \sum_{i=0}^{n} l_i = m, \tag{6}$$

$$\left|p_a^{(2)} - p_a^{(1)}\right| = \left|\frac{1+l_a}{m+n+1} - \frac{1}{n+1}\right| = \left|\frac{(n+1)(1+l_a) - (m+n+1)}{(n+1)(m+n+1)}\right| =$$

$$= \left|\frac{n + nl_a + l_a + 1 - m - n - 1}{(n+1)(m+n+1)}\right| = \left|\frac{(n+1)l_a - m}{(n+1)(m+n+1)}\right| = \frac{\left|l_a - \frac{m}{n+1}\right|}{m+n+1}.$$

Позначимо $r\left(l_a\right)=\left|p_a^{(2)}-p_a^{(1)}\right|=\frac{\left|l_a-\frac{m}{n+1}\right|}{m+n+1}$. Тоді

$$d\left(\bar{l}\right)=\sum_{i=0}^{n} r\left(l_i\right). \tag{7}$$

Теорема 1. *Якщо статистика d досягає мінімального значення, то виконується умова* $\left|l_i-\frac{m}{n+1}\right|<1\,\forall i=\overline{0,n}$.

Доведення. Доведемо від супротивного. Покажемо, що для довільного вектора $\overline{l_1}=(l_0,l_1,...,l_n)$, для якого не виконується умова теореми, існує вектор $\overline{l_2}$, для якого значення d менше.

Для вектора $\overline{l^1}$ $\exists\, a:\left|l_a-\frac{m}{n+1}\right|\geqslant 1$:

1) $l_a-\frac{m}{n+1}\geqslant 1\Rightarrow l_a\geqslant 1+\frac{m}{n+1}$.

Тоді знайдеться $b:l_b<\frac{m}{n+1}$.

Нехай це не так, тоді $\sum\limits_{i=0}^{n} l_i\geqslant n\frac{m}{n+1}+\frac{m}{n+1}+1=1+m>m$ — протиріччя.

Отже, $\exists b:l_b<\frac{m}{n+1}$.

Розглянемо вектор $\overline{l_2}$, в якому $l_a\to l_a-1$, $l_b\to l_b+1$, а інші елементи збігаються. Тоді

$$\begin{aligned}
&d\left(\overline{l_1}\right)-d\left(\overline{l_2}\right)=r\left(l_a\right)+r\left(l_b\right)-r\left(l_a-1\right)-r\left(l_b+1\right)=\\
&=\frac{\left|l_a-\frac{m}{n+1}\right|}{m+n+1}+\frac{\left|l_b-\frac{m}{n+1}\right|}{m+n+1}-\frac{\left|l_a-1-\frac{m}{n+1}\right|}{m+n+1}-\frac{\left|l_b+1-\frac{m}{n+1}\right|}{m+n+1}=\\
&\qquad=\frac{l_a-\frac{m}{n+1}}{m+n+1}--\frac{l_a-1-\frac{m}{n+1}}{m+n+1}+\frac{\frac{m}{n+1}-l_b}{m+n+1}-\frac{\left|l_b+1-\frac{m}{n+1}\right|}{m+n+1}=\\
&\qquad\quad=\frac{\frac{m}{n+1}+1-l_b-\left|l_b+1-\frac{m}{n+1}\right|}{m+n+1}=\begin{cases}\frac{2\left(\frac{m}{n+1}-l_b\right)}{m+n+1}>0, l_b+1>\frac{m}{n+1},\\ \frac{2}{m+n+1}>0, l_b+1\leqslant\frac{m}{n+1}.\end{cases}
\end{aligned}$$

Отже, $d\left(\overline{l_2}\right)<d\left(\overline{l_1}\right)$.

2) $l_a-\frac{m}{n+1}\leqslant -1\Rightarrow l_a\leqslant\frac{m}{n+1}-1$.

Тоді знайдеться $b:l_b>\frac{m}{n+1}$.

Нехай це не так, тоді $\sum\limits_{i=0}^{n} l_i\leqslant n\frac{m}{n+1}+\frac{m}{n+1}-1=m-1<m$ — протиріччя.

Отже, $\exists b:l_b>\frac{m}{n+1}$.

Розглянемо вектор $\overline{l_2}$, в якому $l_a\to l_a+1$, $l_b\to l_b-1$, а інші елементи збігаються. Тоді

$$d\left(\overline{l_1}\right) - d\left(\overline{l_2}\right) = r\left(l_a\right) + r\left(l_b\right) - r\left(l_a+1\right) - r\left(l_b-1\right) =$$
$$= \frac{\left|l_a - \frac{m}{n+1}\right|}{m+n+1} + \frac{\left|l_b - \frac{m}{n+1}\right|}{m+n+1} - \frac{\left|l_a + 1 - \frac{m}{n+1}\right|}{m+n+1} - \frac{\left|l_b - 1 - \frac{m}{n+1}\right|}{m+n+1} =$$
$$= \frac{\frac{m}{n+1} - l_a}{m+n+1} - \frac{\frac{m}{n+1} - l_a - 1}{m+n+1} + \frac{l_b - \frac{m}{n+1}}{m+n+1} - \frac{\left|l_b - 1 - \frac{m}{n+1}\right|}{m+n+1} =$$
$$= \frac{l_b + 1 - \frac{m}{n+1} - \left|l_b - 1 - \frac{m}{n+1}\right|}{m+n+1} = \begin{cases} \frac{2}{m+n+1} > 0, l_b - 1 > \frac{m}{n+1}, \\ \frac{2\left(l_b - \frac{m}{n+1}\right)}{m+n+1} > 0, l_b - 1 \leqslant \frac{m}{n+1} - 1. \end{cases}$$

Отже, $d\left(\overline{l_2}\right) < d\left(\overline{l_1}\right)$.

□

Теорема 2.

$$\min(d) = \frac{2(n+1)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)}{n+m+1}. \tag{8}$$

Доведення. 1) $\left\{\frac{m}{n+1}\right\} = 0$.

Покажемо, що тоді мінімальне значення статистики $= 0$.

За теоремою 1 $\left|l_i - \frac{m}{n+1}\right| < 1\,\forall i = \overline{0,n}$. Оскільки $\frac{m}{n+1}$ — ціле, то якщо $l_i \neq \frac{m}{n+1}$, то $\left|l_i - \frac{m}{n+1}\right| \geqslant 1$, що суперечить теоремі 1. Тому мінімальне значення $\min(d) = \sum\limits_{i=0}^{n} r\left(l_i\right) = \sum\limits_{i=0}^{n} r\left(\frac{m}{n+1}\right) = 0$.

2) $\left\{\frac{m}{n+1}\right\} > 0$.

За теоремою 1 l_i може дорівнювати або $\left[\frac{m}{n+1}\right]$, або $\left[\frac{m}{n+1}\right] + 1$. Серед всіх l_i : $(n+1)\left\{\frac{m}{n+1}\right\}$ елементів дорівнюють $\left[\frac{m}{n+1}\right] + 1$ і $(n+1)\left(1 - \left\{\frac{m}{n+1}\right\}\right)$ елементів дорівнюють $\left[\frac{m}{n+1}\right]$, інакше не виконується $\sum\limits_{i=0}^{n} l_i = m$.

Тоді

$$\min(d) = \sum_{i=0}^{n} r\left(l_i\right) = (n+1)\left\{\frac{m}{n+1}\right\}\frac{\left|\left[\frac{m}{n+1}\right] + 1 - \frac{m}{n+1}\right|}{m+n+1} +$$
$$+ (n+1)\left(1 - \left\{\frac{m}{n+1}\right\}\right)\frac{\left|\left[\frac{m}{n+1}\right] - \frac{m}{n+1}\right|}{m+n+1} = \frac{2(n+1)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)}{m+n+1}.$$

□

Теорема 3. *Статистика d досягає максимального значення, якщо $\exists\, a : l_a = m$ та максимальне значення*

$$\max\left(d\right) = \frac{2m}{m+n+1}. \tag{9}$$

Доведення. 1) Покажемо, що для довільного вектора, для якого не виконується умова теореми існує інший вектор, значення статистики на якому буде більшим.

Якщо не виконується умова теореми, то $\exists\,\overline{l_1}$, для якого $\exists\, a : l_a \geqslant \frac{m}{n+1}$ і $\exists b \neq a : l_b > 0$.

Розглянемо вектор $\overline{l_2} : l_a \to l_a + l_b, l_b \to 0$.

$$d\left(\overline{l^2}\right) - d\left(\overline{l^1}\right) = r\left(l_a + l_b\right) + r\left(0\right) - r\left(l_a\right) - r\left(l_b\right) =$$

$$= \frac{\left|l_a + l_b - \frac{m}{n+1}\right| + \left|0 - \frac{m}{n+1}\right|}{m+n+1} - \frac{\left|l_a - \frac{m}{n+1}\right| + \left|l_b - \frac{m}{n+1}\right|}{m+n+1} =$$

$$= \frac{l_b + \frac{m}{n+1} - \left|l_b - \frac{m}{n+1}\right|}{m+n+1} = \begin{cases} \frac{2\frac{m}{n+1}}{m+n+1} > 0, l_b \geqslant \frac{m}{n+1}, \\ \frac{2l_b}{m+n+1} > 0, l_b < \frac{m}{n+1}. \end{cases}$$

Отже, $d\left(\overline{l_2}\right) > d\left(\overline{l_1}\right)$.

2)Серед всіх l_i: n елементів дорівнюють 0 і 1 елемент дорівнює m, інакше не виконується $\sum\limits_{i=0}^{n} l_i = m$.

Тоді

$$\max\left(d\right) = \sum_{i=0}^{n} r\left(l_i\right) = nr\left(0\right) + r\left(m\right) = \frac{\frac{mn}{n+1}}{m+n+1} + \frac{m + \frac{m}{n+1}}{m+n+1} = \frac{2m}{m+n+1}.$$

□

Наслідок 2. *За теоремами 2 і 3*

$$\frac{2\left(n+1\right)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)}{m+n+1} \leqslant d \leqslant \frac{2m}{m+n+1},$$

$$0 \leqslant \frac{(m+n+1)d - 2(n+1)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)}{2\left(m - \left(n+1\right)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)\right)} \leqslant 1.$$

Означення 1. Позначимо

$$D = \frac{\left(m+n+1\right)d - 2\left(n+1\right)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)}{2\left(m - \left(n+1\right)\left\{\frac{m}{n+1}\right\}\left(1 - \left\{\frac{m}{n+1}\right\}\right)\right)}. \tag{10}$$

Назвемо D — **статистикою включення.**

Зауваження 1. За наслідком 2 статистика включення приймає значення з з відрізка $[0,1]$.

Складність алгоритму обчислення

Як було показано, статистика d однозначно визначається вектором $\bar{l}$:

$$d\left(\bar{l}\right) = \sum_{i=0}^{n} r\left(l_i\right).$$

Тому для обчислення статистики включення достатньо знайти елементи цього вектора.

Для цього достатньо побудувати варіаційні ряди $x_{(1)} < x_{(2)} < ... < x_{(n)}$ і $z_{(1)} < z_{(2)} < ... < z_{(m+n)}$, тобто фактично відсортувати масиви довжиною n і $m+n$.

Результуюча складність алгоритму — $(m+n)\log(m+n)$.

Для порівняння розглянемо складність алгоритму обчислення p-статистики, для якої обчислюються довірчі інтервали для кожного з під інтервалів еталонної вибірки, що потребують перебору кожного елемента тестової вибірки.

В результаті отримаємо мінімально можливу складність алгоритму n^2m.

Таким чином, складність алгоритму нової запропонованої статистики значно менша ніж складність обчислення p-статистики.

Практичне застосування

В реальних задачах маємо виміри якоїсь характеристики об'єктів дослідження і певні еталонні значення для визначених об'єктів, що є задачею класифікації.

Розглянемо набір еталонних вибірок, кожна з яких характеризує свій клас. Визначимо такий критерій класифікації — тестова вибірка належить класу, статистика включення між еталонною вибіркою класу і тестовою вибіркою мінімальна серед усіх еталонних вибірок.

Для апробації статистики побудуємо 4 еталонні вибірки, породжені випадковими генераторами з нормальними функціями розподілу і різними характеристиками.

Далі згенеруємо набір тестових вибірок і обчислимо для них статистику включення з еталонними вибірками.

Були згенеровані еталонні вибірки об'ємом 1000 нормальними генераторами випадкових величинами з параметрами μ σ: $(1,2)$, $(1,3)$, $(4,2)$, $(4,3)$.

Далі такими самими генераторами кожного типу було згенеровано по 100 тестових вибірок об'ємом 1000 кожного класу і була виконана класифікація згідно критерію. Серед 400 тестових вибірок помилково було класифіковано одну.

Було повторено експеримент, зменшивши об'єм тестових вибірок до 200. Правильно було класифіковано по класам відповідно $(73, 97, 72, 96)$.

Наступний експеримент — об'єми еталонних вибірок 200, тестових — 100. Правильно було класифіковано по класам відповідно $(86, 71, 85, 71)$.

Таким чином, навіть на невеликих об'ємах вибірок отримали непоганий результат класифікації. Велика кількість елементів вибірок забезпечує високу точність класифікації.

Висновки

В результаті роботи було побудовано нову статистику, що дає змогу обчислювати міру близькості між двома вибірками. Побудова статистики дозволяє реалізувати алгоритм обчислення не високої складності.

Статистика апробована на задачі класифікації тестової вибірки.

Література

1. Клюшин Д. Непараметрический критерий эквивалентности генеральных совокупностей, основанный на мере близости между выборками / Д. Клюшин, Ю. Петунин // Укр. матем. журн. — 2003. — 5(2). — С. 147–163.
2. Алексєєнко В. Узагальнення p-статистики для вибірок з повтореннями / В. Алексєєнко, Д. Клюшин // Вісник Київського університету. Серія: фізико-математичні науки. — 2011. — 4. — С. 113–116.
3. Hill B. Posteriori distribution of percentiles: Bayes' theorem for sampling from a population / B. Hill // J. Amer. Statist. Assoc. — 1968 — 63. — P. 677–691.

Факультет кібернетики, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64, Київ, 01601, Україна.

Надійшла 23.03.12

Журнал обчислювальної та прикладної математики | 2012, №1(107) | Journal of Computational & Applied Mathematics

УДК 004.93

РОЗПІЗНАВАННЯ НОМЕРНИХ ЗНАКІВ З ВИКОРИСТАННЯМ ПЕРЕТВОРЕННЯ ГОКА

Т. П. Зінько

Резюме. Розглядається задача виокремлення номерного знаку автомобіля (плейту), нормалізація "плейту" та виявлення символів на зображенні номера для проведення їх подальшого розпізнавання за допомогою алгоритму перетворення Гока.

Вступ

Задача виділення номерного знаку автомобіля ("плейту") пов'язана із задачею ідентифікації рухомого автомобіля. Виділення самого "плейту" полягає у фіксації на рухомому зображені певного автомобіля, а саме області або частини, яка його ідентифікує, для подальшого аналізу тієї інформації, яку містить в собі виділене зображення для подальшого аналізу (символи: цифри, букви). Ця прикладна задача зводиться до задачі захвату певного автомобіля, виокремлення певної частини "плейту" та аналізу самого зображення на "плейті". Задача нормалізації та сегментації номерного знака розглядалися в роботі [7], а виокремлення або пошук самого "плейту" використовували, наприклад, двовимірне хвильове перетворення Хаара, як описувалося в роботі [8]. Розпізнавання зображення проводиться за алгоритмом перетворення Гока роботи Донченка В. С. [4–6] як у класичному вигляді, так і у вигляді схеми Гок-пар.

У роботі, яка наведена нижче, розглядається задача у повній постановці із використанням стандартних елементів та їх розвитком для аналізу зображення і виділення інформації про зображені символи. Зокрема, виділення символу на "плейті" пов'язане із виділенням місця, де може бути розташований аналізований символ, що називають сегментацією; стандартизацію відповідних зображень у виділених сегментах та аналіз того, що зображено з метою ідентифікації з певного набору символів. Ці набори символів та спосіб їх розташування можуть бути різними в різних системах формування "плейтів". Відповідна задача виділення місць, на яких зображено символи (сегментація) і аналізу відповідних сегментів відбувається по-різному. Розпізнавання символів, які зображені на тому чи іншому сегменті відбувається з врахуванням тої апріорної інформації про те, якою може бути множина символів. Сегментація відбувається для стандартизованого зображення за інформацією, яка відома про місце і спосіб розташування символів на знаку, а розпізнавання самих символів, які розташовані на тому чи іншому сегменті, можуть здійснюватися різними способами.

У роботі пропонується здійснювати розпізнавання з використання перетворення Гока у двох варіантах: класичному та в рамках загальної схеми Гок-пар.

Зміст задачі розбивається на етапи: ідентифікації рухомого авто; виділення області, де розташований номерний знак; нормалізації виділеного “плейту”; його сегментації та розпізнавання виділених об’єктів. Отже, блок-схема алгоритму виглядатиме як показано на рис. 1:

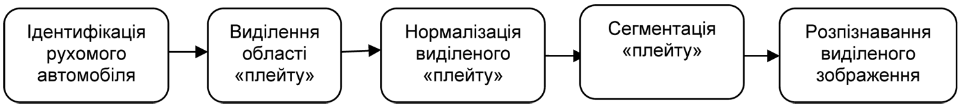

Рис. 1. Блок-схема алгоритму розпізнавання номерних знаків автомобіля

1. Загальна постановка задачі

У роботі розглядається задача виокремлення номерного знаку автомобіля (“плейту”), нормалізація “плейту” та виявлення символів на зображенні номера для проведення їх подальшого розпізнавання за допомогою алгоритму перетворення Гока. Нормалізація зображення необхідне для повороту “плейту”, а саме сукупність символів, які зображені на ньому, горизонтально. Сегментація символів ґрунтується на використанні моделей розташування символів на номері. Запропонований алгоритм нормалізації і сегментації символів дозволяє використовувати його в системах розпізнавання зображення автомобільних номерів (рис. 2). Стандартизоване зображення: горизонтальний розмір — 520, вертикальний — 112, тло — біле, символи — чорні, згідно державного стандарту ДСТУ 4278:2004.

Рис. 2. а) стандартизоване зображення номера; б) приклади “плейтів” для нормалізації і сегментації

Для вдалого розпізнавання “плейту” на вхідних зображеннях номерів, приклади яких наведені на рис. 2, пропонується розбити задачу на три етапи. На першому етапі виконується нормалізація зображення, яке полягає в повороті зображення таким чином, щоб символи на ньому розташовувалися горизонтально. На другому етапі проводиться обробка нормалізованого зображення з метою виділення символів. На третьому — відбувається розпізнавання символів за допомогою перетворення Гока.

2. Нормалізація зображення автомобільного номера

Суть нормалізації полягає в автоматичному обчисленні невідомих параметрів перетворень, якими наділяються вхідні зображення, і подальшим приведенням до стандартизованого вигляду. Нормалізація зображення номерного знака проводиться за два кроки. На першому кроці визначається кут повороту номера в площині зображення. На другому — виконується алгоритм отримання нормалізованого зображення номера з результуючого зображення з урахуванням кута його повороту (рис. 3)

Рис. 3. Схема отримання нормалізованого зображення номера

Визначення кута повороту зображення номерного знака виконується із використанням користуванням декількох етапів обробки та аналізу зображень. На першому етапі виконується операція підкреслення контурів на зображенні, при цьому використовується лінійний оператор Собеля для контурів, який має маску згортки або ядро оператора:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Оператор Собеля заснований на згортці зображення невеликими цілочисельними фільтрами у вертикальному і горизонтальному напрямках, тому його відносно легко обчислювати. Оператор використовує ядра 3×3, з якими згортають вихідне зображення для обчислення наближених значень похідних по горизонталі і по вертикалі. Наведений оператор більш чутливий до напрямів контурів, близьким до горизонтального, тому дозволяє добре виділити на зображенні верхню і нижню частину номерного знака (рис. 4)

а)

б)

Рис. 4. а) фрагмент вихідного зображення зі знайденим положенням номера; б) результат підкреслення контурів, застосовуючи оператор Собеля

На другому етапі виконується розрахунок карти щільності знайдених точок контурів у просторі коефіцієнтів лінійних залежностей просторових координат згідно перетворенню Гока (ПГ).

3. Перетворення Гока в обробці плейтів

Як відомо (див., наприклад, [1]) ПГ веде початок від патента американського інженера P. V. C. Hough [2] і використовувався як інженерний інструмент виділення прямих на бінарізованому зображенні. В розширеному варіанті, в якому цей інструмент існує наразі, він може використовуватися для виділення кривих другого порядку на числовій площині та відповідних поверхонь в тривимірному просторі. З певним модифікаціями, які стосуються реалізації та виділення кривих на контурному зображенні, цей засіб існує і наразі (див, наприклад, [3]). Вичерпну математичну теорію ПГ можна знайти в роботі [4] (див. також [5]).

3.1. Математична модель класичного ПГ

За загальним змістом ПГ як засіб обробки зображення передбачає наявність на зображенні (бінарізованому) кривих певного заданого параметричного сімейства $y = g_\theta(x), \theta \in \Theta \subseteq R^K$, що зв'язують дві координати зображення x, y.

Точки контурів зображення вважаються представленими послідовністю спостережуваних значень $(x_i, y_i), x_i \in X, y_i \in Y, i = \overline{1,n}$, в парах яких елементи x, y зв'язані співвідношеннями

$$y_i = g_{\theta_i}(x_i), \quad i = \overline{1,n}, \tag{1}$$

що відповідають кривим заданого параметричного сімейства. В моделі спостережень (1) може бути представлена і адитивна помилка спостережень ε:

$$y_i = g_{\theta_i}(x_i) + \varepsilon_i,\ i = \overline{1,n} \tag{2}$$

Специфікою моделі спостережень (1) чи (2), що відрізняє її від відповідника регресійної моделі спостережень є те, що в кожному із спостережень параметр θ є своїм, можливо, відмінним від значень параметрів кривих в інших спостереженнях. Така модель спостережень означає, що кожна із точок (x, y) вибірки, загалом, належить своїй кривій, що визначається значенням параметру $\theta_i \in \Theta, i = \overline{1,n}$. Сутністю ПГ є визначення кількості точок зображення: елементів послідовності, представлених послідовністю (x_i, y_i), $i = \overline{1,n}$, які належать кривій $y = g_\theta(x)$ для кожного із можливих значень параметру $\theta \in \Theta \subseteq R^K$. Представленими на зображенні вважаються ті криві, що відповідають тим параметрам кривих, які є локальними максимумами акумуляторної функції. Складність організації відповідної процедури полягає в тому, що у класичному варіанті перетворення множина можливих значень параметру $\theta \in \Theta$ є континуальною. Множина можливих значень параметру Θ піддається "дискретизації" через розбиття множини параметрів на скінченну кількість підмножин: $\Theta = \bigcup_{k \in K} \Theta_k$. Саме для кожного з цих елементів розбиття підраховується кількість точок послідовності спостережень – точок контурів зображення, що належать кривим

$$y = g_\theta(x), \theta \in \Theta_k, k \in K.$$

Ця кількість точок елементів послідовності (x_i, y_i), $x_i \in X$, $y_i \in Y$, $i = \overline{1, n}$ на елементах розбиття Θ_k, $k \in K$ називається ***акумуляторною функцією***.

Саме зображення при цьому називаються ***простором зображення***, а розбиття $\{\Theta_k,\ k \in K\}$ множини можливих значень параметру Θ називають ***простором Гока***.

Заслуга Гока полягає у пропозиції спеціальної процедури підрахунку значень акумуляторної функції на $\{\Theta_k,\ k \in K\}$ у випадку параметричного сімейства прямих в нормальній параметризації $\rho = x\cos\phi + y\sin\phi$, $\theta = (\rho, \phi) \in \Theta = \Pi$, $\Pi-$ прямокутник в R^2, а простір Гока – сукупність прямокутників побудованих за дискретизацією вісей $O\rho, O\phi$.

Зауважимо, що у цьому випадку розбиття природно параметризується двома індексами $i, j : \Theta_{ij}-$ номерами дискретів розбиття за кожною із вісей. Гок запропонував перевіряти чи належить точка (x, y) множині кривих $y = g_\theta(x)$, $\theta \in \Theta_k$, $k \in K$ через перевірку еквівалентної умови: наявності не порожнього перетину Θ_k з множиною параметрів тих кривих параметричного сімейства, яким належить досліджувана точка (x, y). Саме цю множину пропонується в загальній теорії перетворення Гока [4] називати ***перетворенням Гока точки*** (x, y). За позначення $L_{(x,y)}$ перетворення Гока точки (x, y) зображення та позначення $A(\Theta_k)$, $k = \overline{1, K}$, маємо формалізацію процедури підрахунку акумуляторної функції [4] у вигляді:

$$A(\Theta_k) = \sum_{i=1}^{n} \delta(\Theta_k \cap L_{(x_i,y_i)}), \quad k = \overline{1, K}, \tag{3}$$

$$L_{(x_i,y_i)} = \{\theta \in \Theta : y_i = g_\theta(x_i)\}, \quad i = \overline{1, n}.$$

3.2. Загальна схема ПГ

В роботі [6] запропонована загальна схема перетворення Гока: для абстрактних просторів: зображень та Гока. Основою такого перенесення є виконання для графіків кривих $G_\theta = \{(x, y) : y = g_\theta(x)\}, \theta \in \Theta$ і перетворень Гока $L_{(x,y)}, (x, y) \in X \times Y$ принципової властивості.

Лема 1 (фундаментальна в теорії ПГ)**.** *Для класичної схеми ПГ та в позначеннях, описаних вище,*

1. *для довільного* $\theta \in \Theta$, $(x, y) \in G_\theta$ *випливає, що* $\theta \in L_{(x,y)}$
$$s = (x, y) \in G_\theta \Rightarrow \theta \in L_{(x,y)};$$
2. *для довільного* $s = (x, y) \in S_0$ *з того, що* $\theta \in L_{(x,y)}$ *випливає, що*
$$s \in G_\theta : \theta \in L_s \Rightarrow s \in G_\theta.$$

За згаданого підходу Гок-парою просторів будемо називати четвірку $(S_0, S_p, G_\theta, \theta \in S_p, L_s,\ s \in S_0)$, в якій:

• S_0, S_p- абстрактні множини;

•$G_\theta, \theta \in S_p\ , L_s, s \in S_0-$ параметричні системи підмножин в S_0, S_p відповідно: $G_\theta \subseteq S_o$, $\theta \in S_p$, $L_s \subseteq S_p$, $s \in S_p$;

• параметричні сімейства G, L задовольняють умові узгодженості:

1. для довільного $\theta \in S_p$ з того, що $s \in G_\theta$ випливає, що $\theta \in L_s$:
$$s \in G_\theta \Rightarrow \theta \in L_s; \tag{4}$$
2. для довільного $s \in S_0$ з того, що $\theta \in L_s$ випливає, що $s \in G_\theta$:
$$\theta \in L_s \Rightarrow s \in G_\theta;$$

• відповідності між параметрами та підмножинами в параметричних сімействах є взаємно однозначними:

$$\theta \leftrightarrow G_\theta, \theta \in S_p;$$

$$s \leftrightarrow L_s, s \in S_0.$$

Акумуляторна функція для підмножин простору параметрів $\pi \subseteq S_p$ за послідовністю спостережень $s_i \in S_0,\ \ i = \overline{1,n}$ визначається аналогічно (3)

$$A(\pi) = \sum_{i=1}^{n} \delta(\pi \cap L_{s)}), \quad \pi \subseteq \Pi. \tag{5}$$

Множина можливих аргументів акумуляторної функції мусить бути такою, щоб можна було говорити про локальні максимуми. В іншому разі, оцінювати параметр, представлений у послідовності спостережень, треба за глобальним максимумом.

3.3. Загальна схема ПГ в розпізнаванні номерів

Загальна схема ПГ може бути з успіхом застосована у побудові ефективних алгоритмів розпізнавання номерів. Дійсно, позначимо:

1. S_0 — стандартну ділянку зображення символу в прямокутній системі координат, прив'язану до лівого кута зображення:$(x, y) \in S_0 \Leftrightarrow x, y \in [0, M]$;
2. S_p — множину можливих символів зображення в певній нумерації:
$$S_p = \{\theta_1, ..., \theta_M\};$$
3. $G_\theta, \theta \in S_p$ — множина координат простору зображення, що відповідають символу$\theta \in S_p$(зображення символу $\theta \in S_p$) ;
4. $L_{(x,y)} \subseteq S_p$ — множина символів, якій може належати точка (x, y) зображення символу.

Теорема 1. *У введених вище позначеннях четвірка представляє собою абстрактну Гок -пару просторів.*

Доведення. Доведення прямо випливає із того, що для введених вище об'єктів виконується умова узгодженості. □

Зауваження 1. *В рамках наведеної вище теореми зображення символу $\theta \in S_p$ може розумітися по-різному:*

1. *як сукупна множина координат точок символу $\theta \in S_p$;*

2. *як частина сукупної множини координат точок символу $\theta \in S_p$, що відповідають, наприклад, координатам точки перетину символу із стандартним набором променів, достатніх для розрізнення наявного набору символів.*

Визначення символу, представленого на аналізованій ділянці плейту здійснюється за максимумом акумуляторної функції, визначеній у відповідності до співвідношення (5) за множини П допустимих значень аргументу акумуляторної функції, що співпадає із сукупністю одноелементних множин простору параметрів, тобто, із простором параметрів S_p.

4. Розпізнавання символів плейта в загальній схемі ПГ та сегментація символів

Алгоритм розпізнавання плейтів на основі застосування загальної схеми ПГ полягає у тому, що за елементами бінарізованого зображення аналізованого символу обчислюється акумуляторна функція. Символ, який представлений на зображенні вважається той, для якого акумуляторна функція є максимальною.

Сегментація є одним з найбільш важливих етапів обробки і розпізнавання зображення, мета якої полягає у виділенні на зображеннях зв'язкових областей (об'єктів) з приблизно однаковими характеристиками яскравості або забарвленості. Причому останнім часом у зв'язку із необхідністю автоматичної обробки і розпізнавання все зростаючого потоку як статичних так і динамічних відео даних, найбільш важливе значення надається синтезу ефективних процедур сегментації, що працюють в режимі реального часу і дозволяють отримувати надійні результати, які будуть стійкими щодо перешкод.

Застосуємо один із багатьох методів та алгоритмів сегментації зображень, який коротко опишемо нижче. Детальніше з алгоритмом сегментації можна ознайомитися в роботі [8].

Алгоритм сегментації символів номерного знаку ґрунтується на тому, що середня яскравість в міжсимвольних інтервалах, хоча б, нижча середньої яскравості в зображеннях символів. Загальна схема алгоритму сегментації складається з двох основних частин:

1. знаходимо всі індекси стовпців, які відповідають локальним мінімумам середньої яскравості стовпців $c_i,\ \ i = 1, ..., n$.
$$c_i = c_i(B) = \frac{1}{m}\sum_{j=1}^{m} b_{ij},$$
$$c(B) = \frac{1}{n}\sum_{i=1}^{n} c_i(B),$$
де $m-$висота плейту в точках, $n-$ довжина плейту в точках, $B = \{b_{ij}\}$ — матриця яскравостей точок в тонах сірого, $0 \le b_{ij} \le b^{\max},\ i = 1, ..., n,\ j = 1, ..., m$.
2. знаходимо і видаляємо з цього списку індексів помилкові контури символів.

В результаті необхідно знайти індекси стовпців-границь між символами.

Для сегментації символів на автомобільному номері пропонується використати підхід, заснований на підгонці під реальне зображення різних моделей розташування символів на номері. Кожна з моделей відповідає визначеному стандарту розташування символів. Розглядається модель розташування символів в однорядкових номерних знаках: {ББ ЦЦЦЦ ББ}, де Б — буква, Ц — цифра.

Кожну модель можна представити у вигляді зображення темних прямокутників, які відповідають символам, на світлому фоні, як показано на рис. 5.

Рис. 5. Приклад моделі розташування символів

Якщо номери мають інший тип, наприклад, білі символи на червоному тлі, або на чорному, то перед зіставленням моделі з зображенням останнє можна інвертувати по яскравості. Як критерій відповідності моделі зображенню номера використовується величина:

$$K(x, y, W, H) = \frac{S_w - S_b}{\sigma_w}.$$

Тут S_w — середня яскравість зображення під світлою областю, S_b — середня яскравість зображення під чорною областю, σ_w — середньоквадратичний розкид яскравості зображення під світлою областю, x і y — координати моделі всередині зображення номера, W і H — довжина і висота моделі. Чим більше значення критерію $K(x, y, W, H)$, тим більше модель відповідає зображенню номера. Підгонка моделі під зображення полягає у виборі згідно прийнятим критерієм найкращого положення і розмірів моделі. Після цього, згідно того ж критерію, приймається рішення про найкращий тип моделі для поточного зображення номера. Якщо враховувати можливі похибки при нормалізації номера, на основі запропонованого критерію можна незалежно визначити більш точне положення для кожного символу окремо поблизу знайденого його положення з використанням всієї моделі. Приклади спільної роботи послідовних етапів нормалізації зображені номери і сегментації символів наведено на рис. 6. У верхньому рядку рис. 6 показані вихідні зображення номерів, в другому рядку — результат їх нормалізації, в нижньому рядку — результат сегментації символів на номері (кожен символ являє собою окреме зображення)

Рис. 6. Приклади роботи алгоритмів нормалізації зображення номерів і сегментації символів

Отримані після сегментації зображення символів дозволяють використовувати їх для розв'язання задачі розпізнавання. Запропонований алгоритм сегментації дозволяє також визначити тип номерного знака, а отже, з'ясувати, є чи є кожен із символів буквою або цифрою, що дозволить полегшити аналіз зображень на наступному етапі розпізнавання. Запропоновані досить прості моделі дозволяють виконувати операцію сегментації з високою ефективністю за якістю і швидкості, а також дозволяють у разі необхідності без значенням значних додаткових зусиль розширити число використовуваних моделей номерного знака шляхом введення в розгляд інших можливих розташувань символів.

Висновки

Розглянуті у статті алгоритми нормалізації зображень номерних знаків за кутом повороту і сегментації символів на них за якістю і швидкості роботи дозволяють використовувати їх в системі розпізнавання автомобільних номерів в якості попередніх етапів підготовки зображень до розпізнавання. Для виконання одного з етапів алгоритму: для розпізнавання окремого символу, теоретично обґрунтована, побудована і застосована схема загального ПГ. Запропоновані моделі зображень номерів можуть бути легко доповнені або замінені іншими моделями, відповідними іншим стандартам розташування символів на номері. Подальші дослідження в напрямку створення системи автоматичного розпізнавання можуть бути пов'язані з розробкою алгоритмів розпізнавання символів номера, одержуваних на виході розробленого комплексу алгоритмів нормалізації і сегментації зображення номерного знака.

Література

1. Hart P. E. How the Hough Transform was Invented. / P. E. Hart // IEEE Signal Processing Magazine, Issue 6. — 2009 (November). — V. 26 — P. 18–22.
2. Hough P. V. C. Method and Means for Recognizing Complex Patterns. / P. V. C. Hough — U.S. Patent 3069354, 1962.
3. Wikipedia: Hough transform. `http://en.wikipedia.org/wiki/Hough_transform`

4. Донченко В. С. Множинні моделі невизначеності. / В. С. Донченко // автореф. дис. на здобуття наук. ступеня докт. фіз.-мат. наук. — К. 2007. — 26 с.
5. Donchenko V. S. Hough Transform and Uncertainty. / V. S. Donchenko // Proceedings: X International Conference "Knowledge — Dialog — Solution". — V. 1. — June 16–23, 2003. — Varna (Bulgaria). — P. 391–395.
6. Donchenko V. S. General Scheme of the Hough Transform and the Properties of the Hough estimation in the special case of Discrete Spaces. / Donchenko V. S. // Statistical research report S-901 87. — University of Umea, Sweden. — June, 1994. — 10 p.
7. Мурыгин К. В. Нормализация изображения автомобильного номера и сегментация символов для последующего распознавания / Мурыгин К. В. // Искусственный интеллект. 2010. — № 3. — С. 367–369.
8. Антощук С. Г. Архитектура системы распознавания автомобильных номеров [Электронный ресурс] / С. Г. Антощук, В. О. Давыдов, А. А. Нутович. // Труды Одесского политехнического университета. — Вып. 1, 2002.

Факультет кібернетики, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 64, Київ, 01601, Україна.

Надійшла до редколегії 11.03.2012 р.

Журнал обчислювальної та прикладної математики | 2012, №1(107) | Journal of Computational & Applied Mathematics

УДК 519.6

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ ОРБИТАЛЬНОГО ДВИЖЕНИЯ МАГНИТНО ВЗАИМОДЕЙСТВУЮЩИХ ТЕЛ МЕТОДОМ ЧИСЛЕННОГО ЭКСПЕРИМЕНТА

С. С. Зуб, С. И. Ляшко, В. С. Ляшко

Резюме. Проведен численный эксперимент по моделированию квазипериодического движения двух тонких и длинных магнитно взаимодействующих цилиндрических тел. Изложена методика исследования устойчивости магнитной системы, включающая в себя: численное моделирование гамильтоновых уравнений движения, метод Монте-Карло (ММК) заполнения фазового объема, процедуры статистической обработки данных, полученных в результате численного эксперимента. Результаты численного эксперимента свидетельствуют об устойчивом характере найденных квазипериодических движений в окрестности относительного равновесия.

Введение

На заре ядерной физики, когда еще не была осознана природа устойчивости ядра, предлагались различные модели взаимодействия, в том числе искались объяснения устойчивости ядра с позиций классической теории. Среди прочих объяснений нуклонного взаимодействия на роль ядерной силы примерялась магнитная сила. Однако, из работ Тамма и Гинзбурга стало понятно, что в поле с потенциальной энергией постоянного магнитного диполя (“закон обратных кубов”) “как в классической, так и в квантовой теории движение является лимитационным, т. е. имеет место падение частицы на центр” [3]. Позднее, как известно, для квантовых систем устойчивость ядра нашла объяснение, основанное на принципе неопределенности Гейзенберга. На этом был закончен этап поиска магнитных моделей ядра. Но именно тогда возникла новая задача классической механики о поиске устойчивых конфигураций в классических магнитных системах. Несмотря на то, что “проблема $-1/r^3$” породила некоторый стереотип “о глобальной неустойчивости магнитных систем”, она же стала отправной точкой для поиска исключений. Именно Гейзенбергу принадлежит идея об учете пространственной протяженности частицы или “учете реакции собственного поля частицы” [3]. Связанная с этим идея “тесных” магнитных конфигураций привела к появлению модели взаимодействия двух “магнитных гантелей” (см. ниже), которая, пожалуй, является наиболее простой магнитной конфигурацией с магнитной потенциальной энергией, не подчиняющейся закону обратных кубов.

Возможность устойчивого орбитального движения в системе двух тонких и длинных цилиндрических магнитов впервые была предсказана и исследована Козорезом в 1974 году [1].

Математическое описание динамики системы он проводил в рамках лагранжева формализма [2], а для обоснования устойчивости ссылался на теорему Румянцева [4].

Сформулированные им условия устойчивости орбитального движения, как мы сегодня знаем [12], нельзя признать достаточными, но именно его пионерские работы дали толчок для более глубокого и всестороннего исследования "тесных" магнитных конфигураций.

Вообще говоря, исследование орбитального движения является одной из основных задач небесной механики со времен Тихо-Браге, Кеплера и Ньютона.

Однако только во второй половине XIX века была поставлена и решена задача о виде потенциальной энергии в задаче двух тел, приводящих к замкнутым орбитам при любых начальных условиях. Решение этой проблемы связано с именами французских математиков Бертрана, Дарбу и Альфена и подробно описано в учебной литературе по механике [5, стр.38]: "Все ограниченные орбиты в центральном поле замкнуты только в двух случаях: $U = ar^2, a \geq 0$ и $U = -k/r, k \geq 0$".

Однако, если снять некоторые ограничения, например, требование замкнутости, произвольности начальных условий, то можно говорить об устойчивых квазипериодических движениях в смысле ограниченности траектории объемом некоторого тора с заданным поперечным сечением.

В настоящей работе рассматривается консервативная нелинейная система с потенциальной энергией, зависящей как от взаимного положения тел, так и от углов, определяющих их взаимную ориентацию.

Исследованию устойчивости в нелинейных динамических системах посвящен ряд современных работ. Большинство подходов, относящихся к анализу устойчивости таких систем, имеют качественный характер и являются развитием метода Ляпунова [6, 7, 8]. Среди них существуют также методы, которые формулируют некоторые достаточно простые индикаторы устойчивости — "индексы", что позволяет численно исследовать систему и дать количественные оценки устойчивости траектории [9].

Особый класс систем, требующий специального подхода при исследовании устойчивости, представляют гамильтоновы системы. Изучение этих систем связано с такими именами, как Б. Констант, Дж.-М. Сурьо, В. И. Арнольд, А. А. Кириллов и, особено, с Дж. Е. Марсденом. Проблеме устойчивости полностью посвящены лекции Марсдена, прочитанные в Королевском математическом обществе и вышедшие отдельной книгой [10]. Его научная школа уже несколько десятилетий разрабатывает теорию устойчивости гамильтоновых систем с симметрией и применяет ее к исследованию течений жидкости, устойчивости плазмы, эластичных тел, в ОТО и квантовой теории поля.

Основная масса результатов по устойчивости гамильтоновых систем относится к случаю симплектических многообразий. Исключением стала статья Дж.-П. Ортеги и Т. С. Ратью [11], теорема которых относится к более сложному для анализа случаю пуассоновых структур. Именно эта теорема была использована нами в работе [12] для аналитического доказательства устойчивости так называемого относительного равновесия (в данном случае это круговая орбита) в системе двух магнитных гантелей.

Аналитическое доказательство устойчивости обладает бесспорным преимуществом в силу своей строгости и является, по сути, единственным способом установления устойчивости определенных видов траекторий. Возникающие в виде необходимых и достаточных условий функциональные связи между параметрами системы позволяют определить зоны устойчивости.

Однако, аналитический подход для таких сложных систем имеет свои ограничения:

1) исследуется устойчивость узкого класса траекторий, а именно, так называемые относительные равновесия;

2) устойчивость понимается в смысле малых (в пределе бесконечно малых) отклонений от относительного равновесия;

3) не дается информация о запасе устойчивости.

Эта статья посвящена описанию подхода к исследованию устойчивости магнитных систем, основанного на численном моделировании. Методика исследования тестируется на выше упомянутом примере магнитного взаимодействия в системе двух магнитных гантелей. Благодаря достоверной информации об имеющихся устойчивых орбитах эта система является отличным полигоном для отработки предлагаемого подхода к исследованию устойчивости.

Для заполнения фазового объема начальных данных применен метод Монте-Карло (ММК). При исследовании признаков стационарности в поведении системы используется модель линейной регрессии (при этом вычисляются не только параметры регрессии, но и доверительные интервалы), основанная на методе наименьших квадратов (МНК).

Численное моделирование не является альтернативой аналитическому доказательству, но позволяет проводить поисковые исследования, когда аналитические исследования представляют трудность, а также позволяет извлекать дополнительную информацию о системе, если даже имеется аналитическое доказательство устойчивости.

С философской точки зрения никакой эксперимент ни численный, ни реальный не может доказать устойчивость. Он может только дать определенные доводы в пользу устойчивости. Задача предлагаемого подхода состоит в том, чтобы предоставить убедительные доводы.

1. Уравнения движения магнитно взаимодействующих симметричных волчков

Гамильтонов формализм на основе пуассоновых структур [13] дает алгебраическое бескоординатное описание динамики, что особенно важно, когда составными частями системы являются твердые тела.

При построении модели в качестве твердого тела будем рассматривать исключительно *симметричный волчок*, который полностью описывается векотрами: $\vec{\nu}, \vec{m}$ — ось симметрии и момент импульса тела. Причем, его магнитная и массовая оси симметрии совпадают.

Для симметричного волчка переменные ν_i, m_i имеют ясный физический смысл и любые другие динамические переменные (д.п.), имеющие физический смысл, в частности, гамильтониан, выражаются через эти д.п.

В случае симметричного волчка удобно пользоваться компонентами момента относительно *неподвижной* системы координат и, соответственно, направляющими косинусами оси симметрии $\vec{\nu}$ тоже относительно *неподвижной* системы координат.

Ограничение рассмотрением симметричных волчков позволяет использовать единую декартову систему отсчета, причем компоненты векторных физических величин становятся удобными образующими для пуассоновой структуры. В этом случае уравнения движения записывать в векторном виде [13]:

$$\begin{cases} \dot{\vec{r}} = \frac{1}{m}\vec{p}; \\ \dot{\vec{p}} = -\partial_r U\vec{e} - \frac{1}{r}(\partial_{c'} U P_{\perp}^{e}(\vec{\nu}\,') + \partial_{c''} U P_{\perp}^{e}(\vec{\nu}\,'')); \\ \dot{\vec{\nu}}\,' = \alpha'(\vec{m}\,' \times \vec{\nu}\,'); \\ \dot{\vec{m}}\,' = \partial_{c'} U(\vec{e} \times \vec{\nu}\,') - \partial_{c'''} U(\vec{\nu}\,' \times \vec{\nu}\,''); \\ \dot{\vec{\nu}}\,'' = \alpha''(\vec{m}\,'' \times \vec{\nu}\,''); \\ \dot{\vec{m}}\,'' = \partial_{c''} U(\vec{e} \times \vec{\nu}\,'') + \partial_{c'''} U(\vec{\nu}\,' \times \vec{\nu}\,''), \end{cases} \tag{1}$$

где $\vec{r}$, $\vec{p}$ — орбитальные кординаты и импульсы; $\vec{\nu}'$, $\vec{m}'$ — ось симметрии и момент импульса 1-го тела; $\vec{\nu}''$, $\vec{m}''$ — ось симметрии и момент импульса 2-го тела; $r = |\vec{r}|$; $c' = (\vec{e}_r, \vec{\nu}')$; $c'' = (\vec{e}_r, \vec{\nu}'')$; $c''' = (\vec{\nu}', \vec{\nu}'')$; $\vec{e}_r = \vec{r}/r$; $P_{\perp}^{e}(\vec{\nu}') = (\vec{\nu}' - c'\vec{e})$ — проектор на плоскость перпендикулярную вектору $\vec{e}$.

2. Задача об устойчивости орбитального движения двух длинных и тонких цилиндрических магнитов

Для моделирования длинных тонких постоянных магнитов (п. м.) цилиндрической формы воспользуемся моделью магнитных гантелей, где гантеля представляет собой два фиктивных “магнитных заряда”, разнесенных на фиксированное расстояние (см. рис.1).

Потенциальная энергия такой системы может быть записана в виде [15]

$$U = \frac{\mu_0 \kappa' \kappa''}{4\pi} \sum_{\varepsilon', \varepsilon'' = \pm 1} \frac{\varepsilon' \varepsilon''}{R_{\varepsilon' \varepsilon''}}, \tag{2}$$

где $R_{\varepsilon'\varepsilon''}$ — расстояния между полюсами гантелей.

В векторной форме имеем $\vec{R}_{\varepsilon'\varepsilon''} = \vec{r} - \varepsilon' l' \vec{\nu}' + \varepsilon'' l'' \vec{\nu}''$. тогда квадрат вектора в наших обозначениях записывается так

$$\vec{R}^2_{\varepsilon'\varepsilon''}(r, c', c'', c''') = r^2 + l'^2 + l''^2 + 2r(\varepsilon'' l'' c'' - \varepsilon' l' c') - 2\varepsilon' \varepsilon'' l' l'' c'''. \tag{3}$$

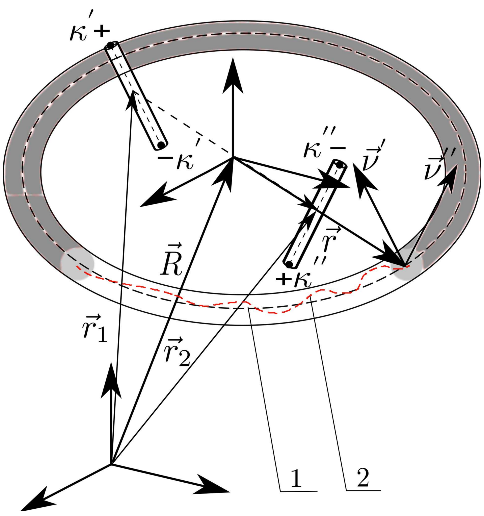

Рис. 1. Магнитные гантели:
1, 2 — траектории устойчивой и пробной (близкой к устойчивой) орбит;
$\vec{\nu}^{'}$, $\vec{\nu}^{''}$ — направляющие орты осей симметрии магнитных гантелей;
$\kappa^{'}$, $\kappa^{''}$ — фиктивные магнитные заряды в вершинах магнитных гантелей.

Перепишем функцию потенциальной энергии системы через μ — магнитный момент, зная что $\kappa = \mu/(2l)$ и учитывая, что $\varepsilon^2 = 1$, получаем

$$U = \frac{\mu_0 \mu^{'} \mu^{''}}{16\pi l^{'} l^{''}} \sum_{\varepsilon^{'}, \varepsilon^{''} = \pm 1} \frac{\varepsilon^{'} \varepsilon^{''}}{R_{\varepsilon^{'} \varepsilon^{''}}}. \tag{4}$$

Из (3) видно, что потенциальная энергия имеет вид $U(r, c^{'}, c^{''}, c^{'''})$.

3. Методика исследования устойчивости магнитных систем

Система 18-ти нелинейных ОДУ первого порядка (1) решается классическим методом Рунге-Кутты 4 порядка.

Вектор начальных значений для каждой траектории формируется из базового вектора начальных значений устойчивой круговой орбиты, удовлетворяющей условиям теоремы Ратью-Ортега, и вектора отклонений, генерируемого в пределах заданного процента от величины базового вектора.

Значения динамических переменных в конце каждого витка траектории становятся вектором начальных значений для следующего витка и поступают на вход решателя (solver) системы ОДУ.

Процесс повторяется требуемое число раз (т. е. с заданным числом оборотов). Таким образом, для каждого вектора начальных условий мы получаем траекторию с заданным числом витков, которую называем независимым испытанием.

Если траектория не выходит за границы заданного тора, то считается, что тело устойчиво движется по “орбите”. Обработчик событий фиксирует прохождение каждого полного витка траектории, останавливает решатель и передает весь массив насчитанных д.п. в процедуру статистической обработки и анализа данных, затем аккумулирует их в массив. Из значений д.п. на конце витка формируется вектор начальных данных для нового витка и подается на вход решателя системы ОДУ.

Генерируя вектор отклонений начальных условий, мы случайным образом заполняем фазовый объем начальными значениями для всех динамических переменных в окрестности устойчивой круговой орбиты.

Моделируя движение в течение заданного времени и ограничиваясь некоторым числом успешно пройденных витков для квазислучайно выбранных начальных условий, мы проверяем гипотезу об устойчивости системы, т. е. испытываем траекторию на удержание в заданном торе. Такой способ испытания на устойчивость системы, по сути, является методом Монте-Карло.

Такой подход оправдан, т. к. многочисленные эксперименты показывают, что в случае нарушения условий теоремы, неустойчивость развивается очень быстро, а именно: уход с квазиорбиты наблюдается между первым и вторым витком, тогда как в испытаниях проверяется 10 и более витков.

Программа моделирования динамики разбивается на ряд независимых блоков, реализованных в виде отдельных процедур, функций и скриптов:

- расчет физических параметров системы (MotionParam);
- инициализация вектора начальных условий базовой орбиты и параметры системы (NTestOrbit);
- генератор вектора начальных условий (InitPertGen);
- вычисление сил, импульсов и периода (dU_r, dU_c1, dU_c2, dU_c3, getP, getT);
- решение ОДУ моделирующих динамику системы (sys17, IsVarOrbit).
- обработчик событий (Orbit_events):
  - — полный виток;
  - — выход за границы тора (падение или уход на бесконечность);
- накопление и обработка данных эксперимента (accum, analysis);
- графическое представление данных (tqplots);
- формирование отчета (stbreport).

Программа моделирования дает полную информацию о динамике системы, т. е. массив всех динамических переменных системы как функцию дискретного времени на квазиорбите. Но хранение всей информации о состоянии динамической системы в случае успешного завершения витка представляется избыточным. Поэтому, аккумуляция данных осуществляется именно на этапе моделирования движения на полном витке, когда имеется вся информация о состоянии системы.

Для дальнейшего анализа на каждом витке мы записываем вектор начальных условий, максимальное отклонение каждой из динамических величин и квазипериод. Для анализа нам могут понадобиться выборочные статистические характеристики квазислучайных распределений динамических переменных, а именно выборочные средние (среднее арифметическое, медиана, дисперсия, среднее квадратическое отклонение и размах выборки). Все эти величины вычисляются и аккумулируются в конце каждого полного витка.

Программа хранит динамические величины, сгруппированные в виде векторов, компоненты которых заданы в декартовой системе координат, связанной с центром масс системы.

Для наблюдения орбитального движения удобнее использовать цилиндрическую систему координат (ЦСК). Запишем наши динамические переменные в ЦСК. Компоненты векторов по z совпадает, а x, y надо выразить через ρ, α.

Имеем $\rho(x, y) = \sqrt{x^2 + y^2}$, $\alpha(x, y) = \arctan(y/x)$ – компоненты вектора положения в ЦСК. Аналогично, находим компоненты векторов $\vec{n}'$, $\vec{m}'$, $\vec{n}''$, $\vec{m}''$.

Нас интересует отклонения динамических переменных системы на квазиорбите от случая относительного равновесия, т. е. в нашем случае от круговой орбиты, т. е. в каждой искомой точке траектории (при фиксированом α) мы ищем разность по ρ и z между значением д.п. на тракктории и соответствующим значением на круговой орбите (см. рис.2).

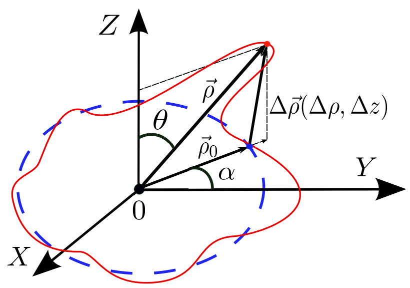

Рис. 2. Вектор положения центра масс и его разность между реальной траекторией и круговой орбитой в ЦСК.

Если $\vec{\rho}$ это радиус-вектор точки на траектории, то соответствующий ему радиус-вектор точки на круговой орбите $\vec{\rho}_0 = \frac{\vec{e}}{\rho}|\vec{\rho}_0|$.

В начальный момент времени имеем $\vec{\rho}_0 \div \vec{e}_1$, $\vec{v}_0 \div \vec{e}_2$, $\vec{\omega}_0 \div \vec{e}_3$. Из $\vec{e}_2 = \vec{e}_3 \times \vec{e}_1$ следует $\vec{v}_0 = \vec{\omega}_0 \times \vec{\rho}_0$ и, соответственно, $\vec{e}_3 = \vec{e}_1 \times \vec{e}_2$ т. е. $\vec{\omega}_0 = \vec{\rho}_0 \times \frac{\vec{v}_0}{\rho_0^2}$.

Импульс $\vec{p}_0 = m\vec{\omega}_0 \times \vec{\rho}_0$ и угловая скорость $\vec{\omega}_0 = \vec{\rho}_0 \times \frac{\vec{p}_0}{m\rho_0^2}$.

Мы предполагаем, что в случае устойчивого поведения все траектории испытаний находятся в границах некоторого тора, охватывающего базовую круговую орбиту. Такое поведение, учитывая стремительный характер развития неустойчивости для магнитных сил, само по себе является указанием на устойчивость движения по данным траекториям.

Однако, мы можем уточнить и количественно оценить “стационарность” поведения системы, выявляя тренды, которые в будущем могут развиться в неустойчивость, используя модель линейной регрессии и вычисляя доверительные интервалы для коэффициентов регрессии.

Если коэффициент наклона линейной регрессии P_1 статистически неотличим от нуля на большом числе оборотов, это является дополнительным аргументом в пользу предположения об устойчивости данной траектории.

Таким образом, исследование устойчивости моделируемого движения в каждом испытании можно разбить на следующие шаги:

1. Для каждого испытания строим линейно регрессионную модель (ЛРМ). Находим также доверительные интервалы для коэффициентов регрессии.

2. Проверяем отличие от нуля коэффициента P_1. Если доверительный интервал для коэффициента P_1 содержит нуль, следовательно, статистически значимого отклонения не выявлено, т. е. с вероятностью менее $100(1-\alpha)\%$ считаем $P_1 = 0$. Параметр α позволяет задать величину уровня значимости ЛРМ.

3. Вычисляем некоторый показатель “качества тренда” т. е. индекс вида $TQindex = \frac{|P_1|}{|P_{1max}-P_{1min}|}$ для каждого испытания. Находим его максимум среди всех испытаний для каждой динамической переменной. Соответствующие испытания для каждой переменной считаем “наихудшими”.

4. Для каждого такого испытания повторно проводим анализ индекса, увеличивая число витков в испытании. Строим зависимость P_1 от числа витков и проверяем попадание нуля в доверительный интервал.

Реализация алгоритма требует многократного выполнения большого числа однотипных действий для каждого испытания. Так как испытания полностью независимы, имеется возможность распараллелить эти процессы, т. е. проводить испытания на различных вычислительных узлах.

При распараллеливании по испытаниям обычный цикл, в котором насчитывается траектория, заменяется его параллельным аналогом. В этом случае все сложности реализации параллельных вычислений скрыты от пользователя MATLAB и решаются на уровне пакета PCT.

Проведение параллельных вычислений с PCT позволяет однотипно, используя средства пакета запускать задачи как на многоядерных настольных компьютерах, обеспечивая до двенадцати виртуальных рабочих узлов с возможностью использования преимуществ GPUs, так и на кластерах через Distributed Computing Server, а теперь и в gLite.

Предлагаемая методика исследования и программа, основанные на прямом моделировании динамики системы и варьировании начальных условий по методу Монте-Карло, является важным инструментом исследования устойчивости. Программа позволяет:

- наблюдать фазовые траектории динамической системы;
- быстро проверять параметры системы на возможность удержания в фазовом объеме для заданного числа витков;
- проводить исследование динамической системы на предмет проявления стационарного поведения системы, которое заключается в том, что отсутствуют тренды максимальных (на витке) отклонений динамических переменных, прежде всего пространственных, но и не только.

Основными данными численного эксперимента являются максимальные отклонения динамических переменных от базовой круговой орбиты на каждом витке траектории.

Разумеется, программа может накапливать все переменные на каждом шаге интегрирования, но эта информация явно избыточна, ее хранение сильно перегружает память вычислительного узла. Поэтому, лучше проводить обработку динамических переменных, полученных в результате моделирования, по завершению каждого полного витка траектории и хранить только результаты обработки (максимальные отклонения и статистические характеристики).

Таким образом, формирование и анализ трендов для каждой динамической переменной начинается после аккумуляции и обработки результатов численного эксперимента по всей траектории из N витков. Сначала по накопленным данным находим регрессионную модель и доверительные интервалы, а затем строим тренды для каждой переменной. Наконец, проверяется достоверность равенства нулю P_1.

Критерием оценки стационарности процесса является TQindex.

Тогда поиск наихудших с точки зрения стационарности трендов сводится к поиску испытания с максимальным показателем. Для каждой динамической переменной будет свой показатель и, следовательно, соответствующее ему испытание.

Для найденных испытаний мы устраиваем дополнительную серию испытаний, фиксируя вектор начальных условий, соответствующий данной траектории, но увеличивая время испытания, т. е. моделируя большее число витков. При этом мы анализируем зависимость показателя от числа витков и ищем номер витка, на котором интересующая нас динамическая переменная достигает своего максимума.

Если с увеличением числа витков индекс имеет тенденцию к уменьшению или хотя бы не растет, то такое поведение свидетельствует в пользу стационарности процесса и, соответственно, указывает на предпосылки к устойчивости.

4. Анализ результатов численного эксперимента

На основе предсказанной аналитически устойчивой круговой орбиты с заданными начальными условиями при заданном отклонении (в 1%) заполняется фазовый объем начальных условий. Затем проводится численное моделирование 10 полных витков для всех вариаций по 100 испытаний.

При этом максимальное отклонение динамических переменных для всех испытаний не превышало 8% от начальных условий.

Из всех испытаний наибольший интерес представляют испытания с наихудшими трендами отклонений по пространственным переменным и по квазипериоду. В конечном счете, любая неустойчивость должна проявиться в поведении трендов квазипериода траектории и пространственных переменных.

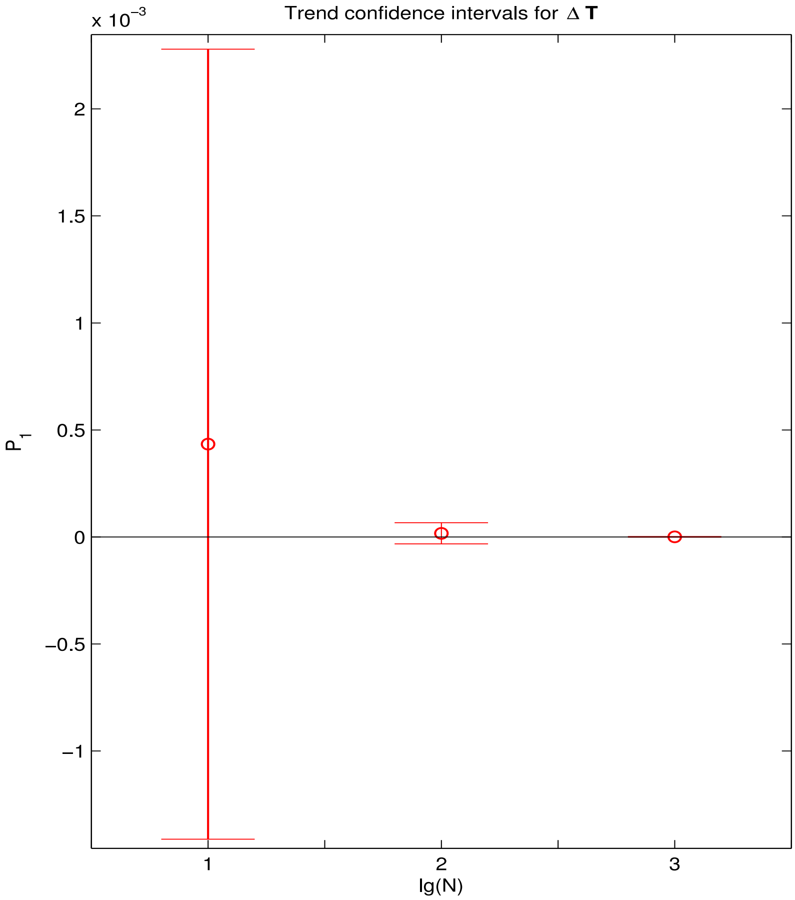

Рис. 3. Доверительные интервалы тренда отклонений по периоду (ΔT):
P_1 — коэффициентов линейной регрессии;
N — число полных витков на орбите.

Оказалось, что с достоверностью 95% полученные тренды коэффициентов линейной регрессии P_1 величин $\Delta\rho_i = \frac{\rho(i)-\rho_0}{R0}$, $\Delta z_i = \frac{z(i)}{R0}$ и $\Delta T_i = \frac{T(i)-T0}{T0}$ имеют вид (см. рис. 3–5), т. е. в пределах траектории они не имеют статистически значимого тренда. На графиках мы представляем наихудшие тренды и их доверительные интервалы. В данном случае можно говорить о стационарном поведении, что по нашему мнению, является убедительным доводом в пользу устойчивости исследуемой конфигурации с запасом устойчивости не менее 1% от начальных данных.

Выводы

Основой для получения гамильтониана магнитно взаимодействующих тел для достаточно широкого класса систем, включающих как п. м., так и сверхпроводящие элементы, послужили работы [16, 17].

Гамильтонов формализм на основе пуассоновых структур позволил дать алгебраическое бескоординатное описание динамики магнитно взаимодействующих осесимметричных твердых тел. В этом случае можно дать описание вращающихся твердых тел в фиксированной инерциальной системе, а не во вращающейся с телом, как это обычно делается. Все динамические переменные группируются в 6 векторных динамических переменных и удается записать как скобки Пуассона, так и уравнения движения в векторном виде.

Создана программа моделирования динамики системы и исследования устойчивости. Для заполнения фазового объема начальных данных применен метод Монте-Карло. Для исследования стационарности поведения орбитального движения используется линейная регрессия, основанная на МНК (при этом вычисляются не только параметры регрессии, но и доверительные интервалы для них).

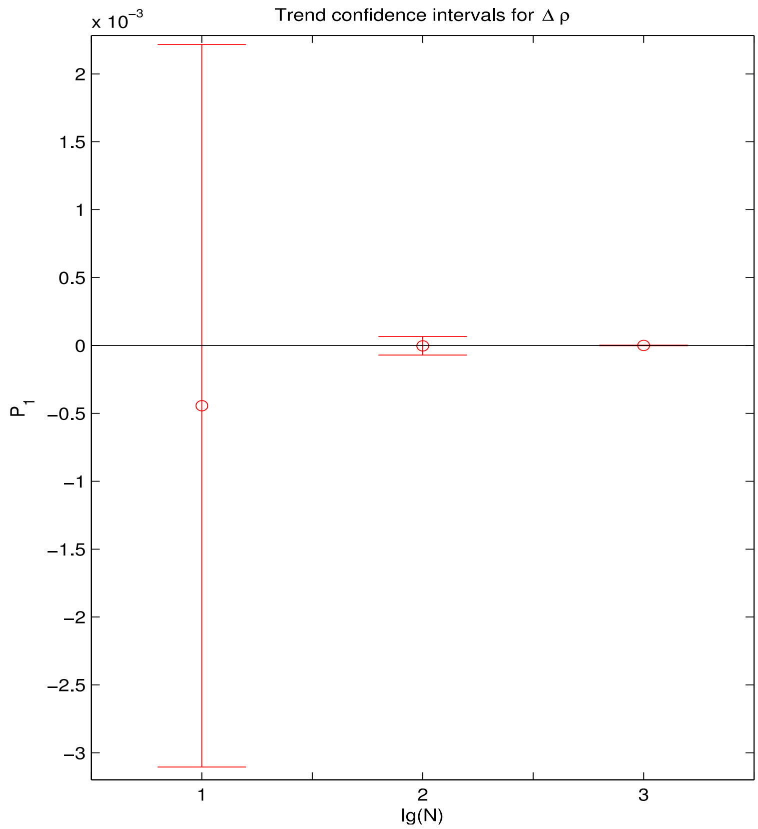

Рис. 4. Доверительные интервалы тренда отклонений в радиальном направлении ($\Delta\rho$):
P_1 — коэффициентов линейной регрессии;
N — число полных витков на орбите.

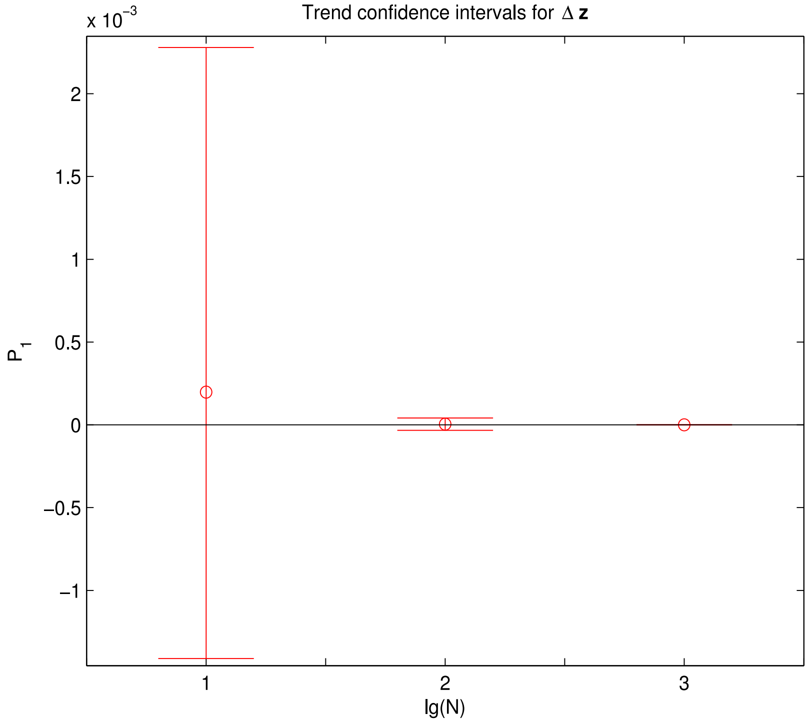

Рис. 5. Доверительные интервалы тренда отклонений в вертикальном направлении (Δz):
P_1 — коэффициентов линейной регрессии;
N — число полных витков на орбите.

Результаты аналитического исследования устойчивости орбитального движения согласуются с результатами численного моделирования. Численное моделирование позволяет получить дополнительную информацию о характере устойчивых орбит, например, о запасе устойчивости, об изменении периода орбит (на самом деле движение не периодическое, а квазипериодическое).

Создана основа для построения методики исследования динамики и устойчивости систем для широкого класса магнитно взаимодействующих тел. Численное моделирование не является альтернативой аналитическим доказательствам, но позволяет проводить поисковые исследования, когда аналитические исследования представляют трудность, а также извлекать дополнительную информацию о системе, если даже имеется аналитическое доказательство устойчивости.

Литература

1. Козорез В. В. О задаче двух магнитов / В. В. Козорез // Изв. АН СССР, Механика твердого тела. — 1974. — № 3. — С. 29–34.
2. Григорьева Л. В. Про динамическую задачу двух свободных цилиндрических магнитов и ее Maple-моделирование / Л. В. Григорьева, В. В. Козориз,

А. В. Козориз, С. И. Ляшко // Докл. Нац. акад. наук Украины. — 2007. — Т. 11. — С. 41–47.

3. Гинзбург В. Л. Теория мезона и ядерные силы / В. Л. Гинзбург // Усп. физ. наук. — 1947. — Т. 31, вып. 2. — С. 174–209.
4. Румянцев В. В., Об устойчивости движения по отношению к части переменных / В. В. Румянцев // Вестн. Моск. ун-та. Математика. — 1957. — № 4. — С. 9–16.
5. Арнольд В. И., Мат. методы класс. мех. / В. И. Арнольд. — М. : Наука, 1971. — 636 с.
6. Шильников Л. П. Методы качественной теории в нелинейной динамике. / Л. П. Шильников, А. Л. Шильников, Д. В. Тураев, Л. Чуа ; пер. с анг. — Москва-Ижевск : Институт компьютерных исследований, 2003. — 221 с.
7. Самойленко Ю. И. Проблемы и методы физической кибернетики / Ю. И. Самойленко. — К. : научное издание, 2006. — 642 с.
8. Комаров Ю. А. Некоторые замечания об экстремальной функции Ляпунова для линейных систем / Ю. А. Комаров, Д. Я. Хусаинов // Укр. мат. журн. — 1983. — 35, — № 6. — С. 750–753.
9. Skokos Ch. Alignment indices: A new, simple method for determining the ordered or chaotic nature of orbits / Ch. Skokos // J. Phys. A. — Mathematical and Theoretical. — 2001. — V. 34. — P. 129–143.
10. Marsden J. E. Lectures on Mechanics / J. E. Marsden — London : Cambridge University Press, 1992. — 254 p.
11. Ortega J-P. Non-linear stability of singular relative periodic orbits in Hamiltonian systems with symmetry / J-P. Ortega, T. S. Ratiu // J. Geom. Phys. — 1999. — V. 32. — P. 160–188.
12. Зуб, С. С. Дослідження стійкості орбітального руху в системі двох магнітно взаємодіючих тіл / С. С. Зуб // Вісник Київського нац. унів. — 2011, вип. 2. — P. 176–184.
13. Зуб, С. С. Гамильтонов формализм для магнитного взаимодействия свободных тел / С. С. Зуб // Журн. обчисл. та прикл. матем. — 2010, № 3(102). — С. 49–62.
14. Ландау, Л. Д. Теоретическая физика. Т. I. Механика. : учеб. / Л. Д. Ландау, Е. М. Лифшиц ; 5-е изд. стереот. — М. : ФИЗМАТЛИТ, 2004. — 224 с. — ISBN 5-9221-0055-6
15. Zub S. Mathematical model of magnetically interacting rigid bodies [Электронный ресурс] / S. Zub // PoS(ACAT08)116. — 2009.
16. Zub S. S. Contact-free Static Stable Equilibrium in the Ground and Space Systems [Электронный ресурс] / S. S. Zub // International scientific conference "Int. Conference on Magnetically Levitated Systems and Linear Drivers (MAGLEV'2002)", September 3–5 2002: proceedings / Lausanne, Switzerland. — 2002. — PP02105.
17. Зуб, С. С. Вплив топології надпровідних елементів на стійкість рівноваги вільного тіла: автореф. дис. канд. тех. наук / С. С. Зуб. — К., 2005. — 15 с.

Факультет кибернетики, Киевский национальный университет имени Тараса Шевченко, ул. Владимирская, 64, Киев, 01601, Украина.

Поступила 29.12.2011

Журнал обчислювальної та прикладної математики | **2012, №1(107)** | Journal of Computational & Applied Mathematics

УДК 519.217; 519.718

ПРО АСИМПТОТИЧНУ ПОВЕДІНКУ РОЗВ'ЯЗКІВ СИСТЕМ СТОХАСТИЧНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ З ПУАССОНОВИМИ ЗБУРЕННЯМИ І МАРКОВСЬКИМИ ПАРАМЕТРАМИ

Т. О. Лукашів, А. В. Нікітін

Резюме. Використано апарат функцій Ляпунова для дослідження асимптотичної стохастичної стійкості в цілому, асимптотичної стійкості в середньому квадратичному в цілому сильного розв'язку стохастичних диференціальних рівнянь з пуассоновими збуреннями з урахуванням внутрішніх марковських параметрів і зовнішніх перемикань типу ланцюга Маркова.

Вступ

Початкові дослідження стійкості ймовірнісних систем у різних постановках проводились в працях [1, 2, 3, 4]. Однак оригінальний підхід до задач стійкості систем з випадковими параметрами вперше був запропонований у працях [5, 6]. Цей підхід дозволив використати для стохастичних диференціальних рівнянь (СДР) основні результати класичної теорії стійкості. В результаті вдалося встановити ряд характерних властивостей ймовірнісних систем. В цьому напрямку працювали видатні математики XX–XXI ст. Р. З. Хасьмінський, В. С. Королюк, Й. І. Гіхман, А. В. Скороход, Г. Дж. Кушнер, В. М. Константинов, В. В. Калашников, Є. Ф. Царков, В. К. Ясинський та інші. У всіх згаданих роботах розглядалася в основному стійкість випадкових процесів з неперервними фазовими траєкторіями, як розв'язок СДР Іто, або випадкових процесів, які мають скінченні стрибки і описуються узагальненими стохастичними рівняннями Іто-Скорохода. У монографії І. Я. Каца [6] розглянута модель стохастичних рівнянь з марковськими параметрами, так званих рівнянь з випадковою структурою, які дозволяють розглядати стійкість систем з розривними фазовими траєкторіями. Відмітимо, що така ситуація природно виникає при описанні численних фізико-технічних процесів. У монографії Є. Ф. Царкова, М. Л. Свердана [7] розглянута стійкість детермінованих різницевих і динамічних систем з урахуванням марковських параметрів та імпульсних марковських перемикань. Праці [8, 9] узагальнюють результати монографій [6] і [7], а саме: розглянуто стохастичне дифузійне рівняння з урахуванням марковських параметрів, які зумовлюють внутрішню зміну структури системи із збереженням властивості стохастичної неперервності реалізації за І. Я. Кацом, а також враховуються зовнішні марковські перемикання у випадкові моменти часу за Є. Ф. Царковим. У роботі узагальнено результати В. К. Ясинського, І. В. Юрченка, Т. О. Лукашіва [8] на випадок систем випадкової структури зі скінченною післядією з урахуванням зовнішніх марковських перемикань, а у працях [10, 11] — на випадок систем стохастичних диференціально-функціональних та диференціально-різницевих рівнянь з марковськими параметрами.

1. Постановка задачі

На ймовірнісному базисі [12] $(\Omega, \mathfrak{F}, \mathbf{F} \equiv \{\mathfrak{F}_t \subset \mathfrak{F}, t \geq 0\}, \mathbf{P})$ розглянемо дифузійне стохастичне диференціально-функціональне рівняння, яке будемо трактувати як динамічну систему випадкової структури [6]

$$dx(t) = a(t, \xi(t), x(t))dt + b(t, \xi(t), x(t))dw(t)+$$
$$+\int_U c(t, \xi(t), u, x(t))\tilde{\nu}(du, dt) \tag{1}$$

із зовнішніми марковськими перемиканнями

$$\Delta x(t)|_{t=t_k} = g(t_k-, \xi(t_k-), \eta_k, x(t_k-)),$$
$$t_k \in S \equiv \{t_n\} \Uparrow,\ n = 0, 1, 2, ...,\ \lim_{n\to\infty} t_n = +\infty \tag{2}$$

і початковими умовами

$$\xi(t_0) = y \in \mathbf{Y}, x(t_0) = x_0 \in \mathbf{R}^m,\ \eta_{k_0} = h \in \mathbf{H}. \tag{3}$$

Тут $\xi(t)$ — марковський процес із значеннями в метричному просторі $\mathbf{Y}$ з перехідною ймовірністю $\mathbf{P}(s, y, t)$; $(\eta_k, k \geq 0)$ — ланцюг Маркова із значеннями в метричному просторі $\mathbf{H}$ з перехідною ймовірністю на k-ому кроці $\mathbf{P}_k(h, G)$; $x(t) = x(t, \omega)$: $[0, +\infty) \times \Omega \to \mathbf{R}^m$, $w(t) \equiv w(t, \omega)$ — одновимірний стандартний вінерів процес, $\tilde{\nu}(du, dt)$ — центрована пуассонова міра [12, 13].

Припустимо, що вимірні за сукупністю змінних відображення a: $\mathbf{R}_+ \times \times \mathbf{Y} \times \mathbf{R}^m \to \mathbf{R}^m$, b: $\mathbf{R}_+ \times \mathbf{Y} \times \mathbf{R}^m \to \mathbf{R}^m$, c: $\mathbf{R} \times \mathbf{Y} \times U \times \mathbf{R}^m \to \mathbf{R}^m$ та g: $\mathbf{R}_+ \times \mathbf{Y} \times \mathbf{H} \times \mathbf{R}^m \to \mathbf{R}^m$ задовольняють за останнім аргументом умову Ліпшиця

$$|a(t, y, x^{(1)}) - a(t, y, x^{(2)})| + |b(t, y, x^{(1)}) - b(t, y, x^{(2)})|+$$
$$+\int_U |c(t, y, u, x^{(1)}) - c(t, y, u, x^{(2)})|\Pi(du) + |g(t, y, x^{(1)}) - g(t, y, x^{(2)})| \leq$$
$$\leq L|x^{(1)} - x^{(2)}|,\ L > 0, \tag{4}$$

при $\forall t \geq 0,\ y \in \mathbf{Y},\ h \in \mathbf{H}$, і умову

$$|a(t, y, 0)| + |b(t, y, 0)| + \int_U |c(t, y, u, 0)|\Pi(du)+$$
$$+|g(t, y, h, 0)| = c < \infty. \tag{5}$$

Вказані умови щодо a, b, c і g гарантують існування сильного розв'язку задачі (1)–(3) з точністю до стохастичної еквівалентності при будь-яких $t_0 \geq 0$, $x \in \mathbf{R}^m$ і заданих реалізаціях марковського процесу $\{\xi(t), t \geq t_0\} \in \mathbf{Y}$ і ланцюга Маркова $\{\eta_k, k \geq k_0\}$ [14].

2. Основні означення

Випадкові зміни структури параметра $\xi(t) \in \mathbf{Y}$ в дифузійному стохастичному диференціальному рівнянні (1), як правило, будемо враховувати одним із таких способів [6, 7].

I. Нехай $\xi(t) \in \mathbf{Y}$ — суто розривний скалярний марковський процес, умовна ймовірність якого допускає розклад [13]

$$\mathbf{P}\{\xi(t+\Delta t) \in [\beta, \beta+\Delta\beta]/\xi(t) = \alpha \neq \beta\} = p(t, \alpha, \beta)\Delta\beta\Delta t + o(\Delta t),$$

$$\mathbf{P}\{\xi(\tau) = \alpha, t < \tau < t + \Delta t/\xi(t) = \alpha\} = 1 - p(t, \alpha)\Delta t + o(\Delta t),$$

де $P\{\cdot/\cdot\})$ — умовна ймовірність; $o(\Delta t)$ — нескінченно мала величина вищого порядку малості відносно Δt. Зазначимо, що за умови регулярності майже всі реалізації є кусково-сталими неперервними справа функціями.

II. Скалярний процес $\xi(t)$ — однорідний марковський ланцюг зі скінченним числом станів $\mathbf{Y} \equiv \{y_1, y_2, \dots, y_k\}$ і відомими параметрами q_{ij} за умови $q_i = \sum\limits_{j \neq i} q_{ij}, i, j = \overline{1, k}$. При цьому умовні ймовірності допускають розклад

$$\mathbf{P}\{\xi(t+\Delta t) = y_i/\xi(t) = y_i\} = q_{ij}\Delta t + o(\Delta t),$$

$$\mathbf{P}\{\xi(\tau) = y_i, t < \tau < t + \Delta t/\xi(t) = y_i\} = 1 - q_i\Delta t + o(\Delta t).$$

III. У момент τ зміни структури системи $y_i \to y_j$ відбувається випадкова стрибкоподібна зміна фазового вектора $x(\tau - 0) = x$, $x(\tau) = z$, для якого задана умовна щільність $p_{ij}(\tau, z)$, а саме:

$$\mathbf{P}\{x(\tau) \in [z, z+dz]/\xi(t-0) = x\} = p_{ij}(\tau, z/x)dz + o(dz).$$

Позначимо через $\mathbf{P}_k((y, h), \Gamma \times G)$ перехідну ймовірність ланцюга Маркова $(\xi(t_k), \eta_k)$ на k-ому кроці. Відповідно до прийнятих в теорії ймовірностей позначень [13, 14, 15] (зв'язаних із цим ланцюгом) введемо індекси так, щоб виконувались рівності

$$\mathbf{P}_{y,h}^{t_k}(\xi(t_{k+1}) \in \Gamma, \eta_{k+1} \in G) \equiv \mathbf{P}_k((y, h), \Gamma \times G)$$

при всіх $t_k \geq t_0$, $y \in \mathbf{Y}$, $h \in \mathbf{H}$ і борелевих $\Gamma \in \mathbf{Y}$ та $G \in \mathbf{H}$.

Тепер введемо функцію

$$\mathbf{P}_k((y, h, \varphi), \Gamma \times G \times C) \equiv$$

$$\equiv \mathbf{P}_{y,h}^{t_k}(x(t_{k+1}, t_k, y, h) \in C, \xi(t_{k+1}) \in \Gamma, \eta_{k+1} \in G)$$

при всіх $t_k \in S \cup \{t_0\}$, $k \in \mathbf{N} \cup \{0\}$, $x \in \mathbf{R}^m$, $y \in \mathbf{Y}$, $h \in \mathbf{H}$ і борелевих $C \subset \mathbf{R}^m$, $\Gamma \subset \mathbf{Y}$, $G \in \mathbf{H}$.

Означення 1. Дискретний оператор Ляпунова $(l\nu_k)(y, h, x)$ на послідовності вимірних скалярних функцій $\nu_k(y, h, x)$: $\mathbf{Y} \times \mathbf{H} \times \mathbf{R}^m \to \mathbf{R}^1, k \in \mathbf{N} \cup \{0\}$, для СДР (1) із зовнішніми марковськими перемиканнями (2) визначаємо рівністю [13]

$$l\nu_k(y,h,x) \equiv \int\limits_{\mathbf{Y}\times\mathbf{H}\times\mathbf{R^m}} \mathbf{P}_k(y,h,x)(du\times dz\times dl)\nu_{k+1}(u,z,l) - \nu_k(y,h,x). \quad (6)$$

Означення 2. Функцією Ляпунова для системи випадкової структури (1)–(3) назвемо послідовність невід'ємних функцій $\{\nu_k(y,h,x), k\geq 0\}$ таких, що виконуються умови:

1) при всіх $k\geq 0,\ y\in\mathbf{Y},\ h\in\mathbf{H},\ x\in\mathbf{R}^m$ визначено дискретний оператор Ляпунова $(l\nu_k)(y,h,x)$ (6);

2) при $r\to+\infty$

$$\underline{\nu}(r) \equiv \inf_{\substack{k\in\mathbf{N},y\in\mathbf{Y}\\ h\in\mathbf{H},|x|\geq r}} \nu_k(y,h,x)\to+\infty; \quad (7)$$

3) при $r\to 0$

$$\overline{\nu}(r) \equiv \sup_{\substack{k\in\mathbf{N},y\in\mathbf{Y}\\ h\in\mathbf{H},|x|\leq r}} \nu_k(y,h,x)\to 0, \quad (8)$$

причому $\bar{\nu}(r)$ і $\underline{\nu}(r)$ неперервні і монотонні.

Будемо розглядати стійкість тривіального розв'язку $x\equiv 0$ системи (1)–(3), тобто виконання (5) при $c=0$.

Оскільки сильний розв'язок $x(t)\equiv x(t,\omega)\in\mathbf{R}^m$ СДР (1) одночасно визначається за допомогою початкових даних $x(t_0)=x_0$, $\xi(t_0)=y$, $\eta_{k_0}=h$, то надалі його позначатимемо $x(t,t_0,y,h,x_0)$.

Означення 3. Систему випадкової структури (1)–(3) назвемо:

– стійкою за ймовірністю в цілому, якщо для $\forall\varepsilon_1>0,\varepsilon_2>0$ можна вказати таке $\delta>0$, що з нерівності $|x|<\delta$ випливає нерівність

$$\mathbf{P}\left\{\sup_{t\geq t_0}|x(t,t_0,y,h,x)|>\varepsilon_1\right\}<\varepsilon_2$$

при всіх $y\in\mathbf{Y},h\in\mathbf{H}$, $x\in\mathbf{R}^m$ і $t_0\geq 0$; — асимптотично стохастично стійкою в цілому, якщо вона стійка за ймовірністю і для довільного $\varepsilon>0$ існує $\delta_1>0$ таке, що

$$\lim_{T\to\infty}\mathbf{P}\left\{\sup_{t\geq t_0}|x(t,t_0,y,h,x)|>\varepsilon\right\}=0$$

при всіх $|x|<\delta_1$, $y\in\mathbf{Y},h\in\mathbf{H}$, $x\in\mathbf{R}^m$ і $T\geq t_0\geq 0$;

– p-стійкою (при деякому $p>0$) в цілому, якщо для $\forall\varepsilon>0$ можна вказати таке $\delta_2>0$, що з нерівності $|x|<\delta_2$ випливає нерівність

$$\mathbf{E}\{|x(t,t_0,y,h,x)|^p\}<\varepsilon$$

при всіх $t>t_0,t_0\geq 0,y\in\mathbf{Y},h\in\mathbf{H}$, $x\in\mathbf{R}^m$;

– асимптотично p-стійкою (при деякому $p>0$) в цілому, якщо вона p-стійка та існує таке $\delta_1>0$, що з нерівності $|x|<\delta_1$ випливає

$$\lim_{t\to\infty}\sup_{y\in\mathbf{Y},h\in\mathbf{H}}\mathbf{E}\{|x(t,t_0,y,h,x)|^p\}=0$$

при всіх $t_0 \geq 0$.

Зауважимо, що при $p = 2$ будемо мати стійкість у середньому квадратичному (l.i.m.) і асимптотичну стійкість у l.i.m.

3. Стійкість систем випадкової структури із зовнішніми марковськими перемиканнями

Встановимо достатні умови стійкості тривіального розв'язку динамічної системи випадкової структури зі скінченною післядією у різних ймовірнісних розуміннях.

Для подальшого викладення використовуватимемо оцінку розв'язку задачі (1)–(3) на інтервалах $[t_k, t_{k+1})$.

Лема 1. *При виконанні умов (4), (5) при всіх $k \geq 0$ для сильного розв'язку задачі Коші (1)-(3) має місце нерівність*

$$\mathbf{E}\left\{\sup_{t_k \leq t \leq t_{k+1}} |x(t)|^2\right\} \leq$$
$$\leq 21(1+2L)e^{9L^2(t_{k+1}-t_k)^2}\left(\mathbf{E}\left\{x^2(t_k)\right\} + 3c^2(t_{k+1}-t_k)\right). \tag{9}$$

Доведення. При всіх $t \in [t_k, t_{k+1}), t_k \geq 0$, легко записати нерівність

$$|x(t)| \leq |x(t_k)| + \int_{t_k}^{t} |a(\tau, y, x(\tau)) - a(\tau, y, 0)|d\tau +$$
$$+ \int_{t_k}^{t} |a(\tau, y, 0)|d\tau + \int_{t_k}^{t} |b(\tau, y, x(\tau)) - b(\tau, y, 0)|dw(\tau) +$$
$$+ \int_{t_k}^{t} |b(\tau, y, 0)|dw(\tau) + \int_{t_k}^{t}\int_{U} |c(\tau, y, u, x(\tau)) - c(\tau, y, u, 0)|\tilde{\nu}(du, d\tau) +$$
$$+ \int_{t_k}^{t} |c(\tau, y, u, 0)|\tilde{\nu}(du, d\tau).$$

Піднісши до квадрата ліву і праву частини одержаної нерівності, обчислюючи sup від одержаного виразу, використовуючи нерівність Коші-Буняковського і нерівність для оцінки умовного математичного сподівання від квадрата супремума інтеграла Вінера-Іто, враховуючи (4), (5), одержимо

$$\mathbf{E}\Big\{\sup_{t_k \leq t < t_{k+1}} |x(t)|^2\Big\} \leq 7\Bigg[\mathbf{E}\{x^2(t_k)\} + 2c^2(t_{k+1}-t_k) +$$
$$+ 9L^2(t_{k+1}-t_k)\mathbf{E}\Big\{\sup_{t_k \leq t < t_{k+1}} \int_{t_k}^{t_k} |x(\tau)|^2 d\tau\Big\}\Bigg].$$

Далі, застосовуючи нерівність Гронулла, легко побачити, що

$$\mathbf{E}\left\{\sup_{t_k\le t<t_{k+1}}|x(t)|^2/\mathfrak{F}_{t_k}\right\}\le 7(\mathbf{E}\{x^2(t_k)\}+3c^2(t_{k+1}-t_k))e^{9L^2(t_{k+1}-t_k)^2}.$$

Для $t=t_{k+1}$ сильний розв'язок системи (1)-(3), очевидно, повинен задовольняти нерівність

$$\mathbf{E}\{|x(t_{k+1})^2/\mathfrak{F}_{t_k}|\}\le 3[\mathbf{E}\{|x^2(t_{k+1}-)/\mathfrak{F}_{t_k}|\}+2\mathbf{E}\{|g(t_{k+1},\xi(t_{k+1}-),\eta_{k+1},x_{k+1})-$$
$$-g(t_{k+1},\xi(t_{k+1}-),\eta_{k+1},0)|^2/\mathfrak{F}_{t_k}\}+2\mathbf{E}\{|g(t_{k+1},\xi(t_{k+1}-),\eta_{k+1},0)|^2/\mathfrak{F}_{t_k}\}]\le$$
$$\le 3[(1+2L)\mathbf{E}\Big\{\sup_{t_k\le t\le t_{k+1}}|x(t)|^2/\mathfrak{F}_{t_k}\Big\}+2c^2].$$

Об'єднуючи дві останні нерівності, одержимо потрібну нерівність (9) леми 1. □

Зауваження 1. В подальшому вважатимемо, що $c=0$ в (9), а також

$$k_0=\begin{cases}\sup\{k\in|t_k\le t_0|\}, & t_0\ge t_1,\\ 0, & t\in[0,t_1).\end{cases}$$

Теорема 1. *Нехай:*

1) $0<|t_{k+1}-t_k|\le\Delta, k\ge 0, \Delta>0$*;*

2) виконується умова Ліпшиця (4);

3) існують послідовності функцій Ляпунова $\nu_k(y,h,x)$ *і* $a_k(y,h,x)$, $k\ge 0$ *такі, що на підставі системи виконується нерівність*

$$l\nu_k(y,h,x)\le -a_k(y,h,x). \tag{10}$$

Тоді сильний розв'язок системи випадкової структури (1), (3) із зовнішніми збуреннями типу ланцюга Маркова (2) асимптотично стохастично стійкий в цілому.

Доведення. Позначимо через $\mathfrak{F}_{t_k}$ — мінімальну σ-алгебру, відносно якої вимірні $\xi(t)$ при всіх $t\in[t_0,t_k]$ і η_n при $n\le k$. Тоді умовне математичне сподівання можна обчислити за формулою [13]

$$\mathbf{E}\{\nu_{k+1}(\xi(t_{k+1}),\eta_{k+1},x(t_{k+1}))/\mathfrak{F}_{t_k}\}=$$
$$=\int\limits_{Y\times H\times\mathbf{R}^m}\mathbf{P_k}(y,h,x)(du\times dz\times dl)\nu_{k+1}(u,z,l)\Bigg|_{\substack{y=\xi(t_k)\\ h=\eta_k\\ x=x(t_k)}}. \tag{11}$$

У цьому випадку за означенням дискретного оператора Ляпунова $(l\nu_k)(y,h,x)$ з рівності (11) одержимо, враховуючи (10), нерівність

$$\mathbf{E}\{\nu_{k+1}(\xi(t_{k+1}),\eta_{k+1},x(t_{k+1}))/\mathfrak{F}_{t_k}\}=\nu_k(\xi(t_k),\eta_k,x(t_k))+$$
$$+(l\nu_k)(\xi(t_k),\eta_k,x(t_k))\le\overline{\nu}(|x(t_k)|). \tag{12}$$

З рівності (9) (за нерівністю Ляпунова для моментів [12] — з існування другого моменту випливає існування першого моменту) і властивостей функції $\overline{\nu}$ випливає існування умовного математичного сподівання лівої частини нерівності (12).

Тепер, на основі (11), запишемо дискретний оператор Ляпунова $(l\nu_k)(\xi(t_k), \eta_k, x(t_k))$ вздовж розв'язків (1)–(3).

$$(l\nu_k)(\xi(t_k), \eta_k, x(t_k)) = \mathbf{E}\{\nu_{k+1}(\xi(t_{k+1}), \eta_{k+1}, x(t_{k+1}))/\mathfrak{F}_{t_k}\} - \\ -\nu_k(\xi(t_k), \eta_k, x(t_k)) \leq -a_k(\xi(t_k), \eta_k, x(t_k)) \leq 0. \quad (13)$$

Тоді при $k \geq 0$ виконується нерівність

$$\mathbf{E}\{\nu_{k+1}(\xi(t_{k+1}), \eta_{k+1}, x(t_{k+1}))/\mathfrak{F}_{t_k}\} \leq \nu_k(\xi(t_k), \eta_k, x(t_k)).$$

Тоді, за означенням супермартингала [12], послідовність випадкових величин $\{\nu_k(\xi(t_k), \eta_k, x(t_k))\}$ при $k \in \mathbf{N}$ утворює супермартингал відносно $\mathfrak{F}_{t_k}$ [7].

Далі, взявши математичне сподівання від обох частин нерівності (13), і, просумувавши за k від $n \geq k_0$ до N одержимо

$$\mathbf{E}\{\nu_{N+1}(\xi(t_{N+1}), \eta_{N+1}, x(t_{N+1}))\} - \mathbf{E}\{\nu_n(\xi(t_n), \eta_n, x(t_n))\} = \\ = \sum_{k=n}^{N} \mathbf{E}\{l\nu_k(\xi(t_k), \eta_k, x(t_k))\} \leq -\sum_{k=n}^{N} \mathbf{E}\{a_k(\xi(t_k), \eta_k, x(t_k))\} \leq 0. \quad (14)$$

Потім маємо

$$\mathbf{P}\Big\{\sup_{t\geq t_0} |x(t, t_0, y, h, x)| > \varepsilon_1\Big\} = \\ = \mathbf{P}\Big\{\sup_{n\in\mathbf{N}} \sup_{t_{k_0+n-1}\leq t\leq t_{k_0+n}} |x(t, t_0, y, h, x)| > \varepsilon_1\Big\} \leq \\ \leq \Big\{\sup_{n\in\mathbf{N}} |x(t_{k_0+n-1}, t_0, y, h, x)| > \varepsilon_1\Big\} \leq \quad (15) \\ \leq \Big\{\sup_{n\in\mathbf{N}} \nu_{k_0+n-1}(\xi(t_{k_0+n-1}), \eta_{k_0+n-1}, x(k_0+n-1)) \geq \overline{\nu}(\varepsilon_1)\Big\}, \ \forall\varepsilon_1 > 0.$$

Дійсно, якщо $\sup |x(t_k)| \geq r$, то на основі (7) виконується нерівність

$$\sup_{k\geq k_0} \nu_k(\xi(t_k), \eta_k, x(t_k)) \geq \inf_{\substack{k\geq k_0, y\in\mathbf{Y} \\ h\in\mathbf{H}, |x|\geq r}} \nu_k(y, h, x) = \overline{\nu}(r).$$

Тепер скористаємось відомою нерівністю для невід'ємних супермартингалів [12] для оцінки правої частини (15):

$$\mathbf{P}\Big\{\sup_{\mathbf{n}\in\mathbf{N}} \nu_{k_0+n-1}(\xi(t_{k_0+n-1}), \eta_{k_0+n-1}, x(t_{k_0+n-1})) \geq \\ \geq \overline{\nu}(\varepsilon_1)\Big\} \leq \frac{1}{\overline{\nu}(\varepsilon_1)} \nu_{k_0}(y, h, x) \frac{\overline{\nu}(||\varphi||)}{\overline{\nu}(\varepsilon_1)}. \quad (16)$$

Враховуючи нерівність (15), нерівність (16) дає можливість гарантувати виконання нерівності

$$\mathbf{P}\Big\{\sup_{t\geq t_0} |x(t, t_0, y, h, x)| > \varepsilon_1\Big\} < \varepsilon_2, \forall\varepsilon_1 > 0, \varepsilon_2 > 0,$$

а це означає, що система (1)–(3) стійка за ймовірністю в цілому.

З нерівності (14) випливає оцінка

$$\mathbf{E}\{\nu_{N+1}(\xi(t_{N+1}), \eta_{N+1}, x(t_{N+1}))\} \leq \nu_{k_0}(y, h, x) -$$

$$-\sum_{k=k_0}^{N}\mathbf{E}\{a_k(\xi(t_k),\eta_k,x(t_k))\}\leq \nu_{k_0}(y,h,x). \tag{17}$$

при всіх $N\geq k_0, y\in\mathbf{Y}, h\in\mathbf{H}$, $x\in\mathbf{R}^m$.

На підставі того, що послідовність $\{a_k\}, k\geq 0$, є функціоналами Ляпунова-Красовського, існують неперервні строго монотонні функції $\underline{a}(r)$ і $\overline{a}(r)$, такі, що

$$\overline{a}(|x|)\leq a_k(y,h,x)\leq \underline{a}(|x|),$$

для $\forall k\in\mathbf{N}, y\in\mathbf{Y}, h\in\mathbf{H}$, $x\in\mathbf{R}^m$, $\underline{a}(0)=\overline{a}(0)=0$.

Отже, із збіжності ряду у (17) випливає збіжність ряду

$$\sum_{k=k_0}^{\infty}\mathbf{E}\{a_k(|x(t_k,t_0,y,h,x)|)\}\quad \forall t_0\geq 0, y\in\mathbf{Y}, h\in\mathbf{H}, x\in\mathbf{R}^m.$$

Тоді в силу неперервності $\underline{a}(r)$ і рівності $\underline{a}(0)=0$ матимемо

$$\lim_{k\to\infty}|x(t_k,t_0,y,h,x)|=0. \tag{18}$$

З (18) випливає прямування до нуля за ймовірністю послідовності $\overline{\nu}(|x(t_k,\,t_0,\,y,\,h,\,x)|)$ при $k\to\infty$ $\forall t_0\geq 0, y\in\mathbf{Y}, h\in\mathbf{H}$, $x\in\mathbf{R}^m$.

Отже, з властивостей функціоналів Ляпунова-Красовського [7, 10] робимо висновок, що невід'ємний супермартингал $\nu_k(\xi(t_k),\eta_k,x(t_k))$ при $k\to\infty$ прямує до нуля за ймовірністю при всіх реалізаціях процесу $\xi(t)\equiv\xi(t,\omega)$ і послідовності $\{\eta_k\}, k\leq 1$.

Далі, невід'ємний обмежений зверху супермартингал має границю з імовірністю одиниця [15]. Тоді, використовуючи (9), одержимо

$$\lim_{k\to\infty}\mathbf{P}\Big\{\sup_{t\geq T}|x(t_k,t_0,y,h,x)|>\varepsilon\Big\}=0,$$

при всіх $y\in\mathbf{Y}$, $h\in\mathbf{H}$, $x\in\mathbf{R}^m$ і $T\geq t_0\geq 0$, що означає, асимптотичну стохастичну стійкість в цілому сильного розв'язку системи (1)–(3). Теорема 1 доведена. □

Як наслідок теореми 1 випливає

Теорема 2. *Нехай:*

1) виконуються умови 1), 2) теореми 1;

2) на підставі системи (1)–(3) для послідовності функцій Ляпунова $\{\nu_k,k\geq 0\}$ виконується нерівність $(l\nu_k)(y,h,x)\leq 0$ $\forall k\geq 0, y\in\mathbf{Y}$, $h\in\mathbf{H}$, $x\in\mathbf{R}^m$.

Тоді динамічна система випадкової структури (1)–(3) стійка за ймовірністю в цілому.

Теорема 3. *Нехай:*

1) виконуються умови 1) — 3) теореми 1;

2) функції Ляпунова $\{\nu_k\},\{a_k\},k\geq 0$, задовольняють нерівності $\forall x\in\mathbf{R}^m$:

$$c_1|x|^2\leq \nu_k(y,h,x)\leq c_2|x|^2, \tag{19}$$

$$c_3|x|^2\leq a_k(y,h,x)\leq c_4|x|^2, \tag{20}$$

при деяких $c_i > 0$, $i = \overline{1,4}$ *для всіх* $k \geq 0, y \in \mathbf{Y}$, $h \in \mathbf{H}$, $x \in \mathbf{R}^m$.

Тоді система випадкової структури (1)–(3) асимптотично стійка в середньому квадратичному в цілому.

Доведення. Використовуючи нерівність (13) для $n = k_0$, на основі (19) легко одержати нерівність

$$\mathbf{E}\{|x(t_{N+1})|\} \leq \frac{1}{c_1}\mathbf{E}\{\nu_{N+1}(\xi(t_{N+1})), \eta_{N+1}, x(t_{N+1})\} \leq$$

$$\leq \frac{1}{c_1}\{\nu_{k_0}(\xi(t_{k_0}), \eta_{k_0}, x)\} \leq \frac{c_2}{c_1}|x|^2.$$

для всіх $N \geq k, k_0 \in \mathbf{N}, x \in \mathbf{R}^m$ і початкових розподілів вектора $\{\xi(t_{k_0}), \eta_{k_0}\}$. Тому, можна стверджувати, що виконується нерівність

$$\mathbf{E}\{|x(t, t_0, y, h, x)|^2\} < \varepsilon, \forall \varepsilon > 0,$$

а це означає, що система (1)–(3) стійка в середньому квадратичному. Далі, використовуючи нерівності (14), (19), (20) можемо одержати нерівність

$$\sum_{\mathbf{k=k_0}}^{N} \mathbf{E}\{|x(t_{N+1})|^2\} \leq \frac{1}{c_3}\sum_{\mathbf{k=k_0}}^{N} \mathbf{E}\{a_k(\xi(t_k), \eta_k, x(t_k))\} \leq$$

$$\leq \frac{1}{c_3}\mathbf{E}\{\nu_{k_0}(\xi(t_{k_0}), \eta_{k_0}, x)\} \leq \frac{c_2}{c_3}|x|^2.$$

Ця нерівність гарантує збіжність ряду, членами якого виступають $\mathbf{E}\{|x(t_{N+1})|^2\}$ для будь-яких початкових даних $x(t_{k_0}) = x$ і початкових розподілів випадкового вектора $\{(\xi(t_{k_0}), \eta_{k_0}\}$.

Отже,

$$\lim_{k\to\infty} \sup_{y\in\mathbf{Y}, h\in\mathbf{H}} \mathbf{E}\{|x(t_k, t_0, y, h, x)|^2\} = 0$$

при всіх $t_0 \geq 0$, що і доводить теорему 3. □

Наслідок. Якщо виконуються умови асимптотичної стохастичної стійкості в цілому теореми 2 і виконується нерівність (19), тоді тривіальний розв'язок динамічної системи випадкової структури (1)–(3) стійкий в середньому квадратичному в цілому.

Висновки

Знайдено достатні умови асимптотичної стійкості в середньому квадратичному в цілому сильного розв'язку стохастичних диференціальних рівнянь з пуассоновими збуреннями з урахуванням як внутрішніх марковських параметрів, так і зовнішніх збурень типу ланцюга Маркова.

Література

1. Андронов А. А. О стохастическом рассмотрении динамических систем / А. А. Андронов, Л. С. Понтрягин, А. А. Витт. — ЖЭТТ, 1933. — Вып. 3, №3. — С. 165–180.
2. Ворович Н. Н. Об устойчивости движения при случайных возмущениях / Н. Н. Ворович // Изв. АН СССР. Сер. мат. — 1965. — Т. 20, №1. — С. 43-48.
3. Гихман И. И. Стохастические дифференциальные уравнения и их применения / И. И. Гихман, А. В. Скороход // К.: Наукова думка. — 1982. — 612 с.
4. Мильмар В. Д. Об устойчивости движения при наличии толчков / В. Д. Мильмар, А. Д. Мышкис // Сиб. мат. журн. — 1960. — Т. 1, №2. — С. 233-237.
5. Кац И. Я. Об устойчивости систем со случайными параметрами / И. Я. Кац, Н. Н. Красовский // Прикл. мат. и механ. — 1960. — Т. 24, Вып. 5. — С. 809-823.
6. Кац И. Я. Метод функций Ляпунова в задачах устойчивости и стабилизации систем случайной структуры / И. Я. Кац — Екатеринбург: УГАПС, 1998. — 222 с.
7. Свердан М. Л. Устойчивость стохастических импульсных систем / М. Л. Свердан, Е. Ф. Царьков — Рига: РТУ, 1994. — 300 с.
8. Лукашив Т. О. Метод функций Ляпунова исследования устойчивости стохастических систем Ито случайной структуры с импульсными марковскими переключениями. I. Общие теоремы об устойчивости импульсных стохастических систем / Т. О. Лукашив, И. В. Юрченко, В. К. Ясинский // Кибернетика и системный анализ. — 2009. — №2. — С. 135–145.
9. Лукашив Т. О. Метод функций Ляпунова исследования устойчивости стохастических систем Ито случайной структуры с импульсными марковскими переключениями. II. Устойчивость по первому приближению импульсных стохастических систем с марковскими параметрами / Т. О. Лукашив, И. В. Юрченко, В. К. Ясинский // Кибернетика и системный анализ. — 2009. — №3. — С. 146–158.
10. Лукашів Т. О. Вибрані питання стійкості систем випадкової структури зі скінченною післядією з урахуванням зовнішніх марковських перемикань / Т. О. Лукашів, Я. М. Чабанюк, В. К. Ясинський // Вісник національного університету "Львівська політехніка", серія "Фізико-математичні науки". 2010. — №687. — С. 132-131.
11. Лукашів Т. О. Про асимптотичну поведінку розв'язків систем стохастичних диференціально-різницевих рівнянь з марковськими параметрами / Т. О. Лукашів, Я. М. Чабанюк, В. К. Ясинський // Вісник національного університету "Львівська політехніка", серія "Фізико-математичні науки". 2011. — №696. — С. 75-80.
12. Королюк В. С. Ймовірність, статистика та випадкові процеси. Теорія та комп'ютерна практика: у 3 т. Т. 3 : Випадкові процеси. Теорія та комп'ютерна практика / В. С. Королюк, Є. Ф. Царков, В. К. Ясинський. — Чернівці: Золоті литаври, 2009. — 798 с.
13. Дынкин Е. Б. Марковские процессы / Е. Б. Дынкин. — М. : Физматгиз, 1969. — 859 с.
14. Скороход А. В. Асимптотические методы теории стохастических дифференциальных уравнений / А. В. Скороход. — К: Наук. думка, 1987. — 328 с.

15. Дуб Дж. Вероятностные процессы / Дж. Дуб. — М: Физматгиз, 1963. — 605с.

Факультет прикладної математики, Чернівецький національний університет імені Юрія Федьковича, вул. Коцюбинського, 2, Чернівці, 58000, Україна.

Надійшла 01.01.2010

Журнал обчислювальної та прикладної математики | **2012, №1(107)** | Journal of Computational & Applied Mathematics

УДК 517.9

ЛИНЕЙНЫЕ НОРМАЛЬНО-РАЗРЕШИМЫЕ УРАВНЕНИЯ В БАНАХОВЫХ ПРОСТРАНСТВАХ

А. А. Покутный

Резюме. Данная заметка посвящена исследованию линейного уравнения в банаховом пространстве с ограниченным нормально-разрешимым оператором. С использованием функционального подхода найдены необходимые и достаточные условия существования решений у такого уравнения.

1. Введение

Пусть L — линейный ограниченный оператор, действующий из банахова пространства E_1 в банахово пространство F_1. В дальнейшем будем предполагать, что оператор L индуцирует следующее разложение этих подпространств:

$$E_1 = N(L) \oplus X_1, \; F_1 = Y_1 \oplus R(L), \tag{1}$$

где $N(L)$ и $R(L)$ традиционно обозначают ядро и образ оператора L. Наряду с (1) имеем разложение единичного оператора на сумму проекторов

$$I_{E_1} = P_{N(L)} + P_{X_1}, \tag{2}$$

$$I_{F_1} = P_{Y_1} + P_{R(L)}. \tag{3}$$

Данная заметка посвящена одному функциональному подходу к изучению линейного уравнения

$$Lx = y, \tag{4}$$

а также представлению проекторов, фигурирующих в (2), (3). Обозначим через $dimN(L) = \mathcal{U}$ и $dimN(L^*) = \mathcal{V}$ мощности размерностей нуль-пространств операторов L и ему сопряженного L^*, соответственно (при этом их счетность не предполагается). Проекторы $P_{N(L)}$, P_{Y_1} в ряде случаев могут быть представлены при помощи разложения по системе базисных элементов, если последняя существует. Напомним основные факты, относительно некоторых базисов, которые будут в дальнейшем использоваться.

Определение 1 ([1]). Последовательность $\{e_n, n \in \mathbb{N}\}$ векторов банахового пространства называется базисом Шаудера или топологическим базисом этого пространства, если каждый его вектор x однозначно раскладывается в ряд $x = \sum_{n=1}^{\infty} \lambda_n e_n$ сходящийся по норме.

Пусть теперь $\mathcal{U}$ — множество произвольной мощности.

Определение 2 ([2]). Система элементов $\{e_\alpha, \alpha \in \mathcal{U}\}$ называется минимальной, если ни один элемент этой системы не принадлежит замкнутой линейной оболочке остальных элементов.

Условие минимальности выполняется далеко не всегда, даже в гильбертовом пространстве (см. например [3]). В дальнейшем будем предполагать существование базисных систем элементов $N(L)$ и Y_1.

Как известно [1], для произвольного банахового пространства существует изометрия на некоторое подпространство пространства $l_\infty(X)$, где X — некоторое множество. В силу этих соображений будут существовать следующие наборы изометрий:

1) существует изометрия $J_1 : E_1 \to E_2 \subset l_\infty(S_1(E_1^*))$;

2) существует изометрия $J_2 : F_1 \to F_2 \subset l_\infty(S_1(F_1^*))$, здесь $E_2 = J_1(E_1)$, $F_2 = J_2(F_1)$ замкнутые подпространства банаховых пространств $l_\infty(S_1(E_1^*))$ и $l_\infty(S_1(F_1^*))$ соответственно; $S_1(E^*)$ — единичная сфера в пространстве сопряженном к E.

Изометрия J_1 переводит каждый элемент x пространства E_1 в функцию $h_x : S_1(E_1^*) \to \mathbb{R}$, действующую по правилу $h_x(f) = f(x), f \in S_1(E_1^*)$. При этом

$$||x||_{E_1} = ||J_1(x)||_{l_\infty(S_1(E_1^*))} = ||h_x||_{E_2} = \sup_{f \in S_1(E_1^*)} |h_x(f)| = \sup_{f \in S_1(E_1^*)} |f(x)|.$$

Изометрия J_2 действует аналогичным образом из пространства F_1 в подпространство $l_\infty(S_1(F_1^*))$.

Введем следующий оператор $\mathcal{L} := J_2 L J_1^{-1} : E_2 \to F_2$. Оператор $\mathcal{L}$ делает коммутативной следующую диаграмму

$$\begin{array}{ccc} E_1 & \xrightarrow{L} & F_1 \\ J_1 \downarrow & & \downarrow J_2 \\ E_2 & \xrightarrow{\mathcal{L}} & F_2 \\ \bigcap & & \bigcap \\ l_\infty(S_1(E_1^*)) & & l_\infty(S_1(F_1^*)) \end{array} , \qquad (5)$$

а сами операторы L и $\mathcal{L}$ являются слабо подобными (пара (J_1, J_2) осуществляет слабое подобие) в категории Ban [1] (напомним, что объектами в этой категории служат произвольные банаховы пространства, а морфизмами — линейные ограниченные операторы).

Известно [1], что пара изометрий (J_1, J_2) осуществляет слабое подобие между L и $\mathcal{L}$ тогда и только тогда, когда она является изоморфизмом в категории $Mor(Ban)$. Классом объектов в этой категории служит произвольный линейный ограниченный оператор, действующий из одного банахового пространства в другое. Морфизмами в этой категории объявлены пары линейных ограниченных операторов (ρ_1, ρ_2) такие, что $\rho_1 \mathcal{L} = L \rho_2$.

Легко видеть, что

$$||\mathcal{L}||_{E_2 \to F_2} = ||J_2 L J_1^{-1}||_{E_2 \to F_2} \leq$$

$$\leq ||J_2||_{F_1 \to F_2} ||L||_{E_1 \to F_1} ||J_1^{-1}||_{E_2 \to E_1} = ||L||_{E_2 \to F_2}.$$

Для доказательства неравенства в другую сторону достаточно проделать аналогичную процедуру для оператора $L = J_2^{-1} \mathcal{L} J_1$. Отсюда заключаем, что $||\mathcal{L}||_{E_2\to F_2} = ||L||_{E_1\to F_1}$. Таким образом мы можем перейти от уравнения (4) к эквивалентному уравнению

$$\mathcal{L}h_x = p_y, \tag{6}$$

но уже определенному в функциональных пространствах. Оператор $\mathcal{L}$, в силу слабого подобия, будет индуцировать следующее разложение банаховых пространств E_2 и F_2

$$E_2 = N(\mathcal{L}) \oplus X_2, F_2 = Y_2 \oplus R(\mathcal{L}), \tag{7}$$

где X_2, Y_2 — некоторые подпространства пространств E_2 и F_2 соответственно. В силу того же подобия размерности ядра и образа оператора $\mathcal{L}$ будут такими же, как и у оператора L. Обозначим минимальную систему базисных функций нуль-пространства $N(\mathcal{L}) \subset E_2$ через $\{e_\alpha\}_{\alpha\in\mathcal{U}} \subset N(\mathcal{L})$, а минимальную систему базисных элементов (функционалов) нуль-пространства $N(\mathcal{L}) \subset F_2$ через $\{\varphi_\beta(\cdot)\}_{\beta\in\mathcal{V}} \subset N(\mathcal{L}^*)$ (предполагается, что последние существуют). В силу минимальности для базисных функций $\{e_\alpha\}_{\alpha\in\mathcal{U}}$ и базисных функционалов $\{\varphi_\beta(\cdot)\}_{\beta\in\mathcal{V}}$ существуют сопряженно биортогональная система $\{f_\alpha(\cdot)\}_{\alpha\in\mathcal{U}}$ линейных функционалов и система функций $\{\psi_\beta\}_{\beta\in\mathcal{V}}$ обладающих тем свойством, что

$$f_\lambda(e_\mu) = \delta_{\lambda\mu}, \lambda, \mu \in \mathcal{U}, \varphi_\nu(\psi_\gamma) = \delta_{\nu\gamma}, \nu, \gamma \in \mathcal{V},$$

где $\delta_{\alpha\beta}$ — символ Кронекера. Проектор на нуль-пространство оператора $\mathcal{L}$ построим следующим образом:

$$P_{N(\mathcal{L})}h_x = \sum_{\alpha\in\mathcal{U}} f_\alpha(h_x)e_\alpha. \tag{8}$$

Аналогичным образом строим проектор на подпространство Y_2:

$$P_{Y_2}p_y = \sum_{\beta\in\mathcal{V}} \varphi_\beta(p_y)\psi_\beta. \tag{9}$$

То, что так определенные операторы будут действительно проекторами, проверяется непосредственно.

Теорема 1. *Уравнение (6) разрешимо для тех и только тех $p_y \in F_2$, которые удовлетворяют равенству*

$$\sum_{\beta\in\mathcal{V}} \varphi_\beta(p_y)\psi_\beta = 0. \tag{10}$$

При выполнении условия (10) решения уравнения (6) будут иметь следующий вид

$$h_x = \sum_{\alpha\in\mathcal{U}} f_\alpha(r_z)e_\alpha + \mathcal{L}^- p_y, \tag{11}$$

для произвольной функции $r_z \in E_2$; $\mathcal{L}^-$ — оператор обобщенно-обратный к оператору $\mathcal{L}$.

Доказательство. (Набросок). Из разложения (7) следует, что оператор $\mathcal{L}$ является нормально-разрешимым [4] и более того обобщенно-обратимым [5, 6]. Тогда, как известно [5, 6], необходимым и достаточным условием разрешимости уравнения (6) является следующее

$$P_{N(\mathcal{L}^*)}p_y = 0. \tag{12}$$

Поскольку $\mathcal{L}$ нормально-разрешим, то $N(\mathcal{L}^*)^{\perp} = R(\mathcal{L})$ [4]. В то же время верно разложение (7). Тогда условие (12) можно заменить на

$$P_{Y_2}p_y = 0.$$

В силу представления (9) это равенство равносильно (10). При выполнении условия разрешимости, решения уравнения (6) будут иметь следующий вид [5]

$$h_x = P_{N(\mathcal{L})}r_z + \mathcal{L}^- p_y \tag{13}$$

для любой функции $r_z \in F_2$. В силу разложения (8), из (13) получаем представление (11). $\square$

2. Случай сепарабельных пространств

Рассмотрим более детально случай, когда пространства E_1 и F_1 сепарабельные. В этом случае оператор $\mathcal{L}$ можно рассматривать, как действующий не в функциональных пространствах, а пространствах последовательностей. Поясним сказанное. Известно [1], что в случае сепарабельности пространств E_1 и F_1 их можно изометрически вложить в некоторые подпространства пространства последовательностей l_∞. В этом случае диаграмму (5) можно заменить на следующую коммутативную диаграмму

$$\begin{array}{ccccc} E_1 & \underrightarrow{j_1} & E_3 \subset C(B_1(E_1^*)) & \underrightarrow{j_2} & E_4 \subset l_\infty \\ L\downarrow & & & & \downarrow \mathcal{L} \\ F_1 & \underrightarrow{i_1} & F_3 \subset C(B_1(F_1^*)) & \underrightarrow{i_2} & F_4 \subset l_\infty \end{array},$$

где $j_1 : E_1 \to E_3 \subset C(B_1(E_1^*))$ — изометрия (преобразование Гельфанда), сопоставляющая каждому вектору $x \in E_1$ функционал означивания $h_x : B_1(E_1^*) \to \mathbb{R}$, $h_x(f) = f(x)$, для всех функционалов f из единичного шара $B_1(E_1^*)$ сопряженного к E_1 пространства. В силу сепарабельности пространства E_1 можно утверждать [1], что $B_1(E_1^*)$ обладает счетным плотным в *слабой топологии подмножеством, которое обозначим $\{f_n\}_{n\in\mathbb{N}}$. Изометрия j_2 переводит каждый функционал $h_x \in E_3$ в вектор $(h_x(f_n))_{n\in\mathbb{N}} \in E_4$, где E_4 подпространство пространства l_∞. Изометрии $\{i_k, k = 1, 2\}$ определяются аналогичным образом. В этом случае пара изометрий (J_1, J_2), где $J_1 = j_2 o j_1, J_2 = i_2 o i_1$, будет изоморфизмом в категории $Mor(Ban)$ между L и $\mathcal{L}$. Предположим, что в такой ситуации подпространства $N(\mathcal{L}), N(\mathcal{L}^*)$ обладают базисами Шаудера, являющиеся одновременно и минимальными системами. В этом случае зафиксируем такие системы векторов $\{\overrightarrow{e}_i = (e_i^{(1)}, e_i^{(2)}, ..., e_i^{(n)}, ...) \in l_\infty, i \in \mathbb{N}\}$ и функционалов $\{\varphi_i(\cdot) \in l_\infty^*, i \in \mathbb{N}\}$, а соответствующие им биортогональные системы будем обозначать $\{f_i(\cdot)\}_{i\in\mathbb{N}} \subset l_\infty^*$ и $\{\overrightarrow{\psi}_i = (\psi_i^{(1)}, \psi_i^{(2)}, ..., \psi_i^{(n)}, ...)\}_{i\in\mathbb{N}} \subset l_\infty$.

Тогда равенства (8) и (9), задающие проекторы на соответствующие подпространства, перепишутся в следующем виде

$$\mathcal{P}_{N(\mathcal{L})}(x_1, x_2, ...) = \sum_{i\in\mathbb{N}} f_i((x_1, x_2, ...))(e_i^{(1)}, e_i^{(2)}, ...). \tag{14}$$

Аналогично

$$\mathcal{P}_{Y_2}(y_1, y_2, ...) = \sum_{i\in\mathbb{N}} \varphi_i((y_1, y_2, ...))(\psi_i^{(1)}, \psi_i^{(2)}, ...). \tag{15}$$

Введем в рассмотрение бесконечные матрицы

$$\mathcal{E} = (\overrightarrow{e}_i)_{i\in\mathbb{N}} = \begin{pmatrix} e_1^{(1)} & e_1^{(2)} & ... & e_1^{(n)} & ... \\ e_2^{(1)} & e_2^{(2)} & ... & e_2^{(n)} & ... \\ ... & ... & ... & ... & ... \\ e_k^{(1)} & e_k^{(2)} & ... & e_k^{(n)} & ... \\ ... & ... & ... & ... & ... \end{pmatrix},$$

$$\Psi = (\overrightarrow{\psi}_i)_{i\in\mathbb{N}} = \begin{pmatrix} \psi_1^{(1)} & \psi_1^{(2)} & ... & \psi_1^{(n)} & ... \\ \psi_2^{(1)} & \psi_2^{(2)} & ... & \psi_2^{(n)} & ... \\ ... & ... & ... & ... & ... \\ \psi_k^{(1)} & \psi_k^{(2)} & ... & \psi_k^{(n)} & ... \\ ... & ... & ... & ... & ... \end{pmatrix}$$

и бесконечные векторы

$$\mathcal{F}(\cdot) = (f_1(\cdot), f_2(\cdot), ..., f_n(\cdot), ...), \;\; \Phi(\cdot) = (\varphi_1(\cdot), \varphi_2(\cdot), ..., \varphi_n(\cdot), ...).$$

Тогда действие проекторов $\mathcal{P}_{N(\mathcal{L})}$ и $\mathcal{P}_{Y_2}$ на векторы из подпространства ограниченных последовательностей можно представить в виде

$$\mathcal{P}_{N(\mathcal{L})}\overrightarrow{x} = \mathcal{F}(\overrightarrow{x})\mathcal{E}, \mathcal{P}_{Y_2}\overrightarrow{y} = \Phi(\overrightarrow{y})\Psi.$$

Так как наборы $(\overrightarrow{e}_i)_{i\in\mathbb{N}}$ и $(\varphi_i)_{i\in\mathbb{N}}$ образуют базисы Шаудера в соответствующих подпространствах, то справедливыми будут следующие представления

$$\mathcal{P}_{N(\mathcal{L})}\overrightarrow{x} = \lim_{n\to\infty} \mathcal{P}_{N(\mathcal{L})}^{(n)}\overrightarrow{x}, \tag{16}$$

$$\mathcal{P}_{Y_2}\overrightarrow{y} = \lim_{n\to\infty} \mathcal{P}_{Y_2}^{(n)}\overrightarrow{y}, \tag{17}$$

где

$$\mathcal{P}_{N(\mathcal{L})}^{(n)}\overrightarrow{x} = \mathcal{F}^{(n)}(\overrightarrow{x})\mathcal{E}^{(n)}, \mathcal{P}_{Y_2}^{(n)}\overrightarrow{y} = \Phi^{(n)}(\overrightarrow{y})\Psi^{(n)},$$

а фигурирующие в представлениях (16), (17) матрицы и векторы — есть $n \times n$- и $1 \times n$-мерные срезки определенных выше бесконечномерных матриц и векторов. Соответствующие пределы будут существовать в силу определения базиса Шаудера (см. например [2]).

Если пространства $E_2 = \mathcal{H}_1$ и $F_2 = \mathcal{H}_2$ — гильбертовы, то в силу изоморфизма соответствующих объектов L и $\mathcal{L}$ категории $Mor(Ban)$ пространства E_1 и F_1 также будут гильбертовыми. В этом случае можно находить не только проекторы $\mathcal{P}_{N(\mathcal{L})}, \mathcal{P}_{Y_2}$, но и ортопроекторы, т.е. проекторы с дополнительным свойством $\mathcal{P}_{N(\mathcal{L})}^* = \mathcal{P}_{N(\mathcal{L})}$ и $\mathcal{P}_{Y_2}^* = \mathcal{P}_{Y_2}$. В случае сепарабельных гильбертовых пространств каждая тотальная ортонормированная система векторов является базисом Шаудера. Пусть $\{\overrightarrow{f}_i\}_{i\in\mathbb{N}}$ и $\{\overrightarrow{\varphi}_s\}_{s\in\mathbb{N}}$ — минимальные системы ортогональных векторов, которые составляют базисы нуль-пространств $N(\mathcal{L}) \subset \mathcal{H}_1$ и $N(\mathcal{L}^*) \subset \mathcal{H}_2$ соответственно. Известно [3], что условие минимальности системы векторов $\{\overrightarrow{f}_i, i \in \mathbb{N}\}$ в гильбертовом пространстве эквивалентно следующему: для любого j величина

$$\delta_j^2 = \lim_{n\to\infty} \frac{\Gamma(\overrightarrow{f}_j, \overrightarrow{f}_{j+1}, ..., \overrightarrow{f}_n)}{\Gamma(\overrightarrow{f}_{j+1}, \overrightarrow{f}_{j+2}, ..., \overrightarrow{f}_n)} > 0,$$

где $\Gamma(\overrightarrow{g}_1, \overrightarrow{g}_2, ..., \overrightarrow{g}_n)$ — определитель Грама системы векторов $\{\overrightarrow{g}_i\}_{i=1}^n$.

В этом случае ортопроектор $\mathcal{P}_{N(\mathcal{L})}$ на нуль-пространство $N(\mathcal{L})$ можно найти следующим образом:

$$\mathcal{P}_{N(\mathcal{L})}\overrightarrow{x} = \lim_{n\to\infty} \mathcal{P}_{N(\mathcal{L})}^{(n)}\overrightarrow{x},$$

а ортопроектор

$$\mathcal{P}_{Y_2}\overrightarrow{y} = \lim_{n\to\infty} \mathcal{P}_{Y_2}^{(n)}\overrightarrow{y},$$

где

$$\mathcal{P}_{N(\mathcal{L})}^{(n)}\overrightarrow{x} = \sum_{i,j=1}^{n} \alpha_{ij}^{(-1)}(\overrightarrow{f}_j, \overrightarrow{x})\overrightarrow{f}_i,$$

$$\mathcal{P}_{Y_2}^{(n)}\overrightarrow{y} = \sum_{s,k=1}^{n} \beta_{sk}^{(-1)}(\overrightarrow{\varphi}_k, \overrightarrow{y})\overrightarrow{\varphi}_s,$$

$\alpha_{ij}^{(-1)}$ и $\beta_{sk}^{(-1)}$ — элементы матриц, обратных к матрицам Грама

$$\Gamma(\overrightarrow{f}_1, \overrightarrow{f}_2, ..., \overrightarrow{f}_n).$$

Существование предела и то, что так определяемый оператор будет ортопроектором, следует из теоремы 7 [2, с.230] в силу монотонности набора ортопроекторов $\mathcal{P}_{N(\mathcal{L})}^{(n)}$ и $\mathcal{P}_{Y_2}^{(n)}$. В том случае, когда зафиксированы ортонормированные системы базисных векторов $(\overrightarrow{e}_i)_{i\in\mathbb{N}}$ и $(\varphi_s)_{s\in\mathbb{N}}$ ортопроекторы на нуль-пространство оператора $\mathcal{L}$ и подпространство Y_2 строятся особенно просто

$$\mathcal{P}_{N(\mathcal{L})}\overrightarrow{x} = \lim_{n\to\infty} \sum_{i=1}^{n} (\overrightarrow{x}, \overrightarrow{e}_i)\overrightarrow{e}_i,$$

$$\mathcal{P}_{Y_2}\overrightarrow{y} = \lim_{n\to\infty} \sum_{s=1}^{n} (\overrightarrow{y}, \overrightarrow{\varphi}_s)\overrightarrow{\varphi}_s.$$

3. Основной результат и его следствия

Используя изоморфность объектов L и $\mathcal{L}$ из теоремы 1 автоматически получаем основной результат о разрешимости уравнения (4) при выполнении условия (1). А именно:

Теорема 2. *Уравнение (4) разрешимо для тех и только тех $y \in F_1$, которые удовлетворяют равенству*

$$\sum_{\beta \in \mathcal{V}} \varphi_\beta(J_2 y)\psi_\beta = 0. \tag{18}$$

При выполнении условия (18) решения уравнения (4) будут иметь следующий вид

$$x = \sum_{\alpha \in \mathcal{U}} f_\alpha(J_2 z) J_1^{-1} e_\alpha + L^- y, (19)$$

для произвольного элемента $z \in E_1$; $L^- = J_1^{-1}\mathcal{L}^- J_2$ — обобщенно-обратный к оператору L.

Доказательство. Набросок доказательства. Условие (18) непосредственно следует из условия (10) (напомним, что $p_y = J_2 y$). Для того, чтобы убедиться в правильности представления (19) достаточно показать, что если оператор $\mathcal{L}^-$ является обобщенно-обратным к оператору $\mathcal{L}$, то оператор $L^- = J_1^{-1}\mathcal{L}J_2$ будет обобщенно-обратным к оператору L. То, что оператор L^- ограничен, очевидно. Из равенств

$$L^- L L^- = J_1^{-1}\mathcal{L}^- J_2 L J_1^{-1}\mathcal{L}^- J_2 = J_1^{-1}\mathcal{L}^- J_2 J_2^{-1}\mathcal{L} J_1 J_1^{-1}\mathcal{L}^- J_2 =$$
$$= J_1^{-1}\mathcal{L}^-\mathcal{L}\mathcal{L}^- J_2 = J_1^{-1}\mathcal{L}^- J_2 = L^-,$$

и

$$L L^- L = J_2^{-1}\mathcal{L} J_1 J_1^{-1}\mathcal{L}^- J_2 J_2^{-1}\mathcal{L} J_1 = J_2^{-1}\mathcal{L}\mathcal{L}^-\mathcal{L} J_1 == J_2^{-1}\mathcal{L} J_1 = L$$

и определения оператора, обобщенно-обратного к данному [6] следует, что L^- действительно представляет собой такой оператор. □

Отметим, что попутно мы установили следующий факт

Следствие 1. *Объекты L^- и $\mathcal{L}^-$ являются изоморфными в категории $Mor(Ban)$. Пара изометрий (J_2, J_1) делает коммутативной следующую диаграмму*

$$\begin{array}{ccc} F_1 & \xrightarrow{L^-} & E_1 \\ J_2 \downarrow & & \downarrow J_1 \\ F_2 & \xrightarrow{\mathcal{L}^-} & E_2 \end{array}.$$

При этом пространства E_1 и F_1, а также пространства E_2 и F_2 раскладываются в прямые суммы подпространств

$$F_1 = N(L^-) \oplus \overline{X}_1,\ E_1 = \overline{Y}_1 \oplus R(L^-),$$
$$F_2 = N(\mathcal{L}^-) \oplus \overline{X}_2,\ E_2 = \overline{Y}_2 \oplus R(\mathcal{L}^-).$$

Сформулируем следствие из теоремы 2 для случая сепарабельных пространств.

Следствие 2. *Уравнение (4) разрешимо для тех и только тех $y \in F_1$, которые удовлетворяют равенству*

$$\lim_{n\to\infty} \mathcal{P}^{(n)}_{Y_2} J_2 \overrightarrow{y} = \overrightarrow{0}. \tag{20}$$

При выполнении условия (20) решения уравнения (4) будут иметь следующий вид
$$x = \lim_{n\to\infty} J_1^{-1}\mathcal{P}^{(n)}_{N(\mathcal{L})} J_2 \overrightarrow{z} + L^- y,$$
для произвольного элемента $z \in E_1$; $L^- = J_1^{-1}\mathcal{L}^- J_2$ — обобщенно-обратный к оператору L.

Замечание 1. Отметим, что представления аналогичные (16) и (17) можно получить и в случае несепарабельных подпространств. Для этого в рассуждениях необходимо заменить сходимость последовательностей на сходимость соответствующих сетей [2]:
$$\mathcal{P}_{N(\mathcal{L})} h_x = \lim_{\alpha\in\mathcal{U}} \mathcal{P}^{(\alpha)}_{N(\mathcal{L})} h_x, \mathcal{P}_{Y_2} p_y = \lim_{\beta\in\mathcal{V}} \mathcal{P}^{(\beta)}_{Y_2} p_y.$$

Литература

1. Хелемский А. Я. Лекции по функциональному аналзу / А. Я. Хелемский — М: МЦНМО, 2004. — 552 с.
2. Канторович Л. В. Функциональный анализ / Л. В. Канторович, Г. П. Акилов — М.: Наука, 1984. — 750 с.
3. Ахиезер Н. И. Теория линейных операторов в гильбертовом пространстве / Н. И. Ахиезер, И. М. Глазман — Т. 1, Х.: Вища школа, 1977. — 315 с.
4. Треногин В. А. Функциональный анализ / В. А. Треногин — М.: Наука, 1980. — 495 с.
5. Boichuk A. A. Generalized Inverse Operators and Fredholm Boundary Value Problems / A. A. Boichuk, A. M. Samoilenko — VSP, Utrecht-Boston, 2004. — 317 p.
6. Гохберг И. Ц. Введение в одномерную теорию сингулярных интегральных операторов / И. Ц. Гохберг, Н. Я. Крупник — Кишинев: Штиинца, 1973. — 426 с.

Лаборатория краевых задач теории дифференциальных уравнений, Институт математики НАН Украины, ул. Терещенковская, 3, Киев, 01601, Украина.

Поступила 20.10.2011

Київський національний університет імені Тараса Шевченка
Журнал обчислювальної та прикладної
2012 **М А Т Е М А Т И К И** 1(107)
Заснований в 1965 році

З М І С Т

С е р і я "О п т и м і з а ц і я"

С е р і я "П р и к л а д н а м а т е м а т и к а"