

DOI:

УДК 519.7

Л.О. Олійник, к.ф.-м.н, доцент, ola19@i.ua

Дніпровський державний технічний університет, м. Кам'янське

ОПЕРАТОРНА МОДЕЛЬ РЕКОМБІНАЦІЇ В ГЕНЕТИЧНИХ АЛГОРИТМАХ

В роботі наведено описання класу лінійних інволютивних операторів, що діють у декартовому добутку евклідового простору. Цей клас лінійних операторів представляє алгебраїчну модель конструювання операторів рекомбінації: кросоверу та мутації.

Клас інволютивних лінійних операторів дозволяє формувати, виходячи з невеликої за потужністю множини пар батьківських векторів-хромосом, популяцію потомків векторів-хромосом, які мають найкращі властивості в сенсі значень фітнес-функції. Крім того, запропоновано один з можливих варіантів алгоритму пошуку оптимального розв'язку.

Ключові слова: кросовер; мутація; фітнес-функція; лінійний інволютивний оператор.

This article provides a description of the class of linear involutive operators functioning in a Cartesian product of Euclidean space. This class of linear operators represents an algebraic model for constructing recombination operators: crossover and mutations.

The class of involutive linear operators allows us to form a population of descendants of chromosome vectors. Basing on the small quantity of the set of pairs of parental vectors-chromosomes, these linear operators have the best properties of the values of the fitness function. Also one of the possible variants of the algorithm for finding an optimal solution is proposed.

Keywords: crossover; mutation; fitnessfunction; linearinvolutive operator.

Постановка проблеми

Робота присвячена описанню алгебраїчних властивостей операторів рекомбінації (схрещування або кросоверу та мутації [1—3]), що використовуються при застосуванні генетичних алгоритмів до задач оптимізації.

Генетичні алгоритми відносять до класу так званих евристичних алгоритмів, що імітують принципи генетичного наслідування та природного селективного відбору кращих розв'язків. Генетичні алгоритми є досить дієвим засобом розв'язання певного класу задач оптимізації. Важливим елементом структури генетичного алгоритму є визначення «області пошуку» (популяції), тобто деякої підмножини векторного простору потенційних розв'язків (хромосом). Визначення стартової множини початкових наближень розв'язків (стартової популяції) виконується випадковим добором пар початкових векторів (батьківських пар хромосом). З отриманої популяції батьківських пар хромосом, застосуванням операторів рекомбінації отримують наступне покоління хромосом, що наслідує від попереднього набору хромосом кращі властивості, з точки зору поведінки цільової функції (фітнес-функції).

Формулювання мети дослідження

Метою роботи є дослідження властивостей класу лінійних операторів, що діють у n -вимірному евклідовому просторі і забезпечують процедуру рекомбінації (схрещування або кросоверу та мутації).

Виклад основного матеріалу

Нехай R^n — евклідовий простір з ортонормованим базисом $\{e_1, \dots, e_n\} \in R^n$, у якому міститься область допустимих розв'язків (хромосом) деякої задачі оптимізації.

Розглянемо клас лінійних операторів з $L(R^n; R^n)$ вигляду:

$$\forall x = (\xi_1, \dots, \xi_n), P_j x = P_j \left(\sum_{i=1}^n \xi_i e_i \right) = \xi_j e_j, Q_j x = Q_j \left(\sum_{i=1}^n \xi_i e_i \right) = \xi_{n-j+1} e_j.$$

Будемо називати ці оператори $P_j, Q_j, j = \overline{1, n}$ елементарними. Ці оператори є проекторами на одновимірні підпростори в R^n . Розглянемо оператори вигляду: $P_j^k = \sum_{i=j}^k P_i$, $Q_j^k = \sum_{i=j}^k Q_i$. Зрозуміло, що ці оператори також є операторами проектування на k -вимірні підпростори в R^n . Дійсно,

$$\forall x = (\xi_1, \dots, \xi_n) \quad P_j^k x = P_j^k \left(\sum_{i=1}^n \xi_i e_i \right) = \sum_{i=j}^k \xi_i e_i, \quad Q_j^k x = Q_j^k \left(\sum_{i=1}^n \xi_i e_i \right) = \sum_{i=j}^k \xi_{n-j+1} e_i.$$

Сума усіх n операторів P_j є одиничним оператором $I = \sum_{i=1}^n P_i$, а сума усіх n операторів

Q_j є одиничним інволютивним оператором $J = \sum_{i=1}^n Q_i$ ($J^{-1} = J$).

Крім того, неважко переконатися, що оператори P_j^k, Q_j^k мають наступні властивості:

$$1) \quad P_j^k P_j^k = P_j^k \quad \text{— ідемпотентність.} \quad 2) \quad Q_{n-j+1}^j Q_j^j = P_j^j$$

Поряд з операторами P_j^k, Q_j^k розглянемо оператори $\tilde{P}_j^k = I - P_j^k, \tilde{Q}_j^k = J - Q_j^k$.

Так як для отримання популяції векторів-потомків необхідно мати певну множину батьківських пар векторів-розв'язків, то логічно розглядати область пошуку у декартовому добутку $R^n \times R^n$, елементами якого є вектори $\hat{X} = \begin{pmatrix} x \\ y \end{pmatrix}$, де $x = (\xi_1, \dots, \xi_n), y = (\eta_1, \dots, \eta_n) \in R^n$.

Нехай $L(R^n \times R^n; R^n \times R^n)$ простір лінійних операторів. Довільний лінійний оператор цього простору представляється у вигляді блочної матриці, елементи якої є лінійні оператори з $L(R^n; R^n)$. Розглянемо клас лінійних операторів, що визначаються наступними блочними матрицями [4]:

$$\hat{P}_j^k = \begin{pmatrix} \tilde{P}_j^k & P_j^k \\ P_j^k & \tilde{P}_j^k \end{pmatrix}, \quad \hat{Q}_j^k = \begin{pmatrix} \tilde{Q}_j^k & Q_j^k \\ Q_j^k & \tilde{Q}_j^k \end{pmatrix}.$$

Операції додавання та множення операторних матриць визначаються звичайним чином, а у разі коли вектори-хромосоми мають двійкове представлення використовується додавання за модулем 2.

Оператори, породжувані операторними матрицями \hat{P}_j^k є операторами перестановок або інволютивними операторами. Крім того ці оператори є граничними елементами, лінійного опуклого замкненого простору самоспряжених стохастичних лінійних операторів.

Неважко пересвідчитись, що інволютивні оператори \hat{P}_j^k є унітарними, тобто задовольняють умовам $\hat{P}_j^k = (\hat{P}_j^k)^{-1} = (\hat{P}_j^k)^*$, $\|\hat{P}_j^k\| = 1$ (тут використовується «кубічна норма» [4]). Оператори \hat{P}_j^k здійснюють обмін групами координат між векторами, тобто виконують однокомпонентне схрещування (кросовер). Оператори \hat{Q}_j^k здійснюють перестановку координат у кожному з векторів з наступним обміном між векторами. Ця операція об'єднує в собі і схрещування, і мутацію.

Якщо схрещування виконується кількома окремими групами координат, то необхідно застосовувати наступні оператори $P_{j_1 \dots j_m}^{k_1 \dots k_m}$, $Q_{j_1 \dots j_m}^{k_1 \dots k_m}$, які діють наступним чином:

$$P_{j_1 \dots j_m}^{k_1 \dots k_m} x = P_{j_1 \dots j_m}^{k_1 \dots k_m} \left(\sum_{i=1}^n \xi_i e_i \right) = \sum_{p=1}^m \sum_{l_i=j_p}^{k_p} \xi_i e_i, \quad Q_{j_1 \dots j_m}^{k_1 \dots k_m} x = Q_{j_1 \dots j_m}^{k_1 \dots k_m} \left(\sum_{i=1}^n \xi_i e_i \right) = \sum_{p=1}^m \sum_{l_i=j_p}^{k_p} \xi_{n-i+1} e_i.$$

Побудуємо за допомогою цих операторів блочні матриці, тобто лінійні оператори з $L(R^n \times R^n; R^n \times R^n)$:

$$\hat{P}_{j_1 \dots j_m}^{k_1 \dots k_m} = \begin{pmatrix} \tilde{P}_{j_1 \dots j_m}^{k_1 \dots k_m} & P_{j_1 \dots j_m}^{k_1 \dots k_m} \\ P_{j_1 \dots j_m}^{k_1 \dots k_m} & \tilde{P}_{j_1 \dots j_m}^{k_1 \dots k_m} \end{pmatrix}, \quad \hat{Q}_{j_1 \dots j_m}^{k_1 \dots k_m} = \begin{pmatrix} \tilde{Q}_{j_1 \dots j_m}^{k_1 \dots k_m} & Q_{j_1 \dots j_m}^{k_1 \dots k_m} \\ Q_{j_1 \dots j_m}^{k_1 \dots k_m} & \tilde{Q}_{j_1 \dots j_m}^{k_1 \dots k_m} \end{pmatrix}.$$

Оператори $P_{j_1 \dots j_m}^{k_1 \dots k_m}$ здійснюють так зване мультикомпонентне схрещування, (обмін кількома групами координат). Оператори $Q_{j_1 \dots j_m}^{k_1 \dots k_m}$ — мультикомпонентне схрещування і мутацію.

Побудовані лінійні оператори $\hat{P}_j^k, \hat{Q}_j^k, P_{j_1 \dots j_m}^{k_1 \dots k_m}, Q_{j_1 \dots j_m}^{k_1 \dots k_m}$ належать класу інволютивних лінійних операторів (блочних матриць), які в свою чергу є граничними елементами більш широкого класу лінійних стохастичних операторів (стохастичних блочних матриць).

Має місце наступне твердження[4].

Твердження 1. Сукупність інволютивних операторних матриць утворює групу унітарних операторів відносно операції множення.

Так як оператори P_j^k, \tilde{P}_j^k представляють собою діагональні матриці, то можна їм бієктивно поставити у відповідність двійкові вектори, тобто бінарні коди, які складаються за правилом: якщо діагональний елемент дорівнює 1 то, відповідна координата двійкового коду також одиниця, інакше нуль. Тобто

$$\chi_j^k \Leftrightarrow \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{pmatrix}, \quad \tilde{\chi}_j^k \Leftrightarrow \begin{pmatrix} 1 \oplus \sigma_1 \\ \vdots \\ 1 \oplus \sigma_n \end{pmatrix}, \quad \sigma_j = \begin{cases} 1, & p_{jj} = 1, \\ 0, & p_{jj} \neq 1. \end{cases}$$

Серед усіх операторів виділимо так звані оператори одновимірного кросоверу \hat{P}_j^j , які обмінюють тільки по одній координаті між векторами $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n) \in R^n$.

Має місце наступне твердження.

Теорема 1. Сукупність одновимірних операторів кросоверу утворює базис у підпросторі інволютивних операторів.

Доведення. Покажемо, що довільний оператор кросоверу \hat{P}_j^k можна отримати перемножуючи між собою оператори \hat{P}_j^j , $j = 1, \dots, n$.

Зрозуміло, що оператори з двійковими кодами χ_j^j , $j = 1, \dots, n$ утворюють мінімальну лінійно незалежну систему на множині n -вимірних бінарних кодів..

Подальше доведення проводимо за індукцією.

Легко перевірити, що $\hat{P}_1^2 = \hat{P}_1^1 \hat{P}_2^2$, $\hat{P}_2^3 = \hat{P}_2^2 \hat{P}_3^3$, враховуючи унітарність цих операторів, отримаємо $\hat{P}_1^3 = \hat{P}_1^2 \hat{P}_2^2 \hat{P}_3^3 = \hat{P}_1^1 \hat{P}_2^2 \hat{P}_2^2 \hat{P}_3^3 = \hat{P}_1^1 \hat{P}_2^2 \hat{P}_3^3$.

Припустимо, що $\hat{P}_1^k = \hat{P}_1^1 \hat{P}_2^2 \dots \hat{P}_k^k$, доведемо рівність $\hat{P}_1^{k+1} = \hat{P}_1^1 \hat{P}_2^2 \dots \hat{P}_{k+1}^{k+1}$.

Неважно бачити, що $\hat{P}_j^k = \hat{P}_j^{k-1} \hat{P}_{j+1}^k$. Враховуючи цю рівність та унітарність операторів, отримаємо: $\hat{P}_1^{k+1} = \hat{P}_1^k \hat{P}_2^{k+1} = \hat{P}_1^k \hat{P}_2^k \hat{P}_3^{k+1} = \dots = \hat{P}_1^k \hat{P}_2^k \dots \hat{P}_k^k \hat{P}_{k+1}^{k+1} = \hat{P}_1^1 \hat{P}_2^2 \dots \hat{P}_k^k \hat{P}_{k+1}^{k+1}$.

Аналогічно доводиться рівність $\hat{P}_j^k = \hat{P}_j^j \hat{P}_{j+1}^{j+1} \dots \hat{P}_k^k$. Теорему доведено.

Отже виходячи з твердження теореми та бінарного кодування операторів кросоверу маємо наступне: існує 2^n операторів кросоверу у просторі $L(R^n \times R^n; R^n \times R^n)$, які отримуються як добуток базисних операторів \hat{P}_j^j , $j=1, \dots, n$. Серед цих операторів існує $\frac{n^2+n-2}{2}$ операторів класичного кросоверу (однокомпонентного).

Застосовуючи оператор \hat{P}_j^k вектору $\hat{X} = \begin{pmatrix} x \\ y \end{pmatrix}$ (x, y пара батьківських хромосом), отримаємо два нащадки z, z^* за наступним правилом: $\hat{P}_j^k \hat{X} = \hat{P}_j^k \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (\xi_1, \dots, \xi_{j-1}, \eta_j, \dots, \eta_k, \xi_{k+1}, \dots, \xi_n) \\ (\eta_1, \dots, \eta_{j-1}, \xi_j, \dots, \xi_k, \eta_{k+1}, \dots, \eta_n) \end{pmatrix}$.

Для операторів мультикомпонентного кросоверу має місце твердження.

Теорема 2. Для операторів мультикомпонентного кросоверу виконується рівність:

$$\hat{P}_{j_1 \dots j_m}^{k_1 \dots k_m} = \hat{P}_{j_1}^{j_1} \dots \hat{P}_{k_1}^{k_1} \dots \hat{P}_{j_m}^{j_m} \dots \hat{P}_{k_m}^{k_m}.$$

Нижче наведено приклад операторів кросоверу у просторі R^4 .

Приклад 1. Розглянемо оператори \hat{P}_j^k з $L(R^4 \times R^4; R^4 \times R^4)$.

Всього існує шістнадцять двійкових кодів, тобто шістнадцять операторів, серед яких нульовий і одиничний. Отже, нетривіальних операторів є чотирнадцять, чотири з яких базисні. Шість операторів $\hat{P}_1^2, \hat{P}_2^3, \hat{P}_1^3, \hat{P}_1^4, \hat{P}_2^4, \hat{P}_3^4$, отримуються як попарний добуток базисних операторів, наприклад, $\hat{P}_1^2 = \hat{P}_1^1 \hat{P}_2^2$. Чотири оператори $\hat{P}_1^3, \hat{P}_2^4, \hat{P}_1^4, \hat{P}_2^4$ отримуються добуток трьох базисних операторів, наприклад, $\hat{P}_1^3 = \hat{P}_1^1 \hat{P}_2^2 \hat{P}_3^3$.

На рис. 1 наведено схему (на чотиривимірному гіперкубі) отримання операторів кросоверу

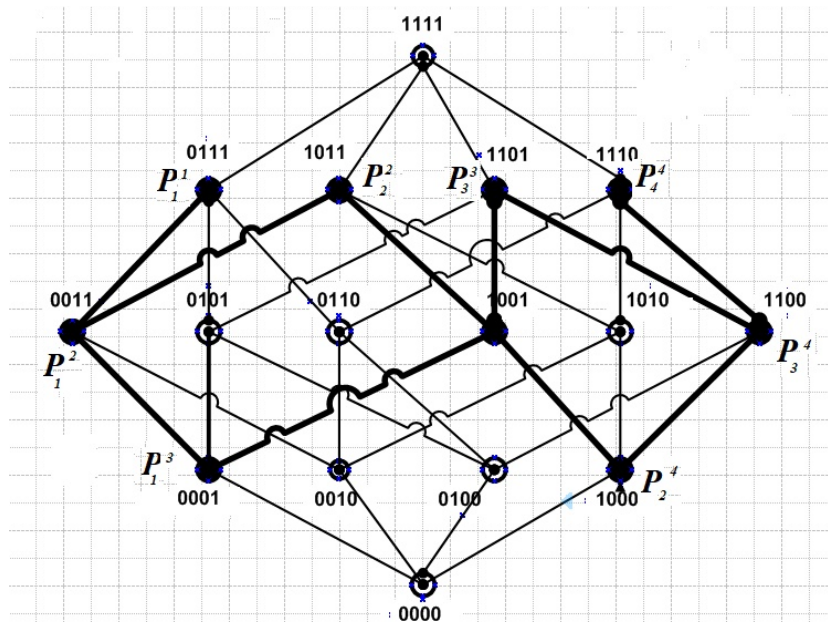


Рис. 1. Схема обчислення операторів кросоверу

Нарешті наведемо операторну інтерпретацію існуючих методів кросоверу ([2]) векторів-хромосом.

1. Дискретна рекомбінація

Фактично зводиться до випадкового вибору оператора \hat{P}_j^k (точніше пари чисел у межах $1 \dots n$). Розглядається вектор $\alpha = (\alpha_1, \dots, \alpha_n)$, координати якого є ймовірностями дій, що виконуються над координатами вектора $x = (\xi_1, \dots, \xi_n)$. Цьому випадку відповідає операторна матриця

$$\hat{P}_j^k = \begin{pmatrix} \tilde{P}_j^k(\alpha) & P_j^k(\alpha) \\ P_j^k(\alpha) & \tilde{P}_j^k(\alpha) \end{pmatrix}.$$

2. Проміжна рекомбінація

Випадково обраний вектор $\alpha = (\alpha_1, \dots, \alpha_n)$, де $0 \leq \alpha_1 \leq 1$.

Позначимо $\hat{X} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\hat{X}^* = (x, y)$, $\hat{A} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$, $\hat{A}^* = (\alpha, \alpha)$. Тоді проміжна рекомбінація ви-
значається наступним чином: $\hat{P}(\alpha) = (\hat{I} - \hat{A}^*(\hat{I} - \hat{P}_j^k))\hat{X}$

Одноточковий кросовер. Здійснюється оператором \hat{P}_j^k . Випадковий процес вибору двійкового або десяткового коду оператора.

Багатоточковий кросовер. Здійснюється оператором $P_{j_1 \dots j_m}^{k_1 \dots k_m}$. Випадковий процес вибору двійкового або десяткового коду оператора.

Використання операторів кросоверу, описання яких наведено вище, дає змогу досить ефективно знаходити розв'язки задач оптимізації. При цьому на кожному кроці роботи генетичного алгоритму використовується множина батьківських пар векторів-хромосом малої потужності.

Нехай в області з R^n , що визначається нерівностями $a_m \leq x_m \leq b_m$, $m = 1, \dots, n$, визначено функцію $F(x)$, для якої необхідно знайти мінімум з певним ступенем точності.

У запропонованій схемі обчислень на кожному кроці використовується множина з восьми пар батьківських векторів-хромосом.

На першому кроці випадковим чином визначаються дві множини по вісім точок області $x_1, \dots, x_8, y_1, \dots, y_8$, які отримують двійкове кодування за порядковим номером. Ці точки простору утворюють пари батьківських векторів-хромосом. З цих восьми пар векторів за допомогою

операторів схрещування отримуються вісім пар векторів-потомків $\hat{P}_j^k \begin{pmatrix} x_l \\ y_m \end{pmatrix} = \begin{pmatrix} x_l^* \\ y_m^* \end{pmatrix}$, для яких

обчислюються значення функції $F(x_l^*)$, $F(y_m^*)$. Серед значень функції відбираємо два найменших, і відповідні їм вектори, які позначимо x_{\min}^* , y_{\min}^* . Вектор, що відповідає найменшому значенню функції обирається центром кулі, радіус, якої дорівнює відстані між отриманими двома векторами, а саме: $R = \left\| x_{\min}^* - y_{\min}^* \right\|$, де $\|x\| = \max_k |\xi_k|$. Таким чином, отримуємо нову об-

ласть пошуку, для якої повторюємо процедуру випадкового підбору восьми пар батьківських векторів-хромосом і відбору центральної точки нової області пошук. Отже, отримаємо послідовність куль, які мають непорожній перетин. Ці перетини утворюють послідовність вкладених куль, радіуси яких прямують до нуля. Тобто послідовність центрів кожної з куль є фундаментальною, отже має границю, яка є розв'язком задачі. Процес пошуку розв'язку зупиняється, коли відстань між останньою парою центральних точок буде меншою заданої величини похибки.

Розглянемо ілюстрацію цього алгоритму на прикладі пошуку мінімуму функції [3]

$$F(x, y) = -\left(2(1-x)^2 e^{-x^2-(y+1)^2} - 7\left(\frac{x}{3} - x^3 - y^2\right) e^{-x^2-y^2} - \frac{1}{5} e^{-(x+1)^2-y^2}\right)$$

на множині $-3 \leq x \leq 3$, $-3 \leq y \leq 3$ з точністю 10^{-3} . Графік функції наведено на рис. 2.

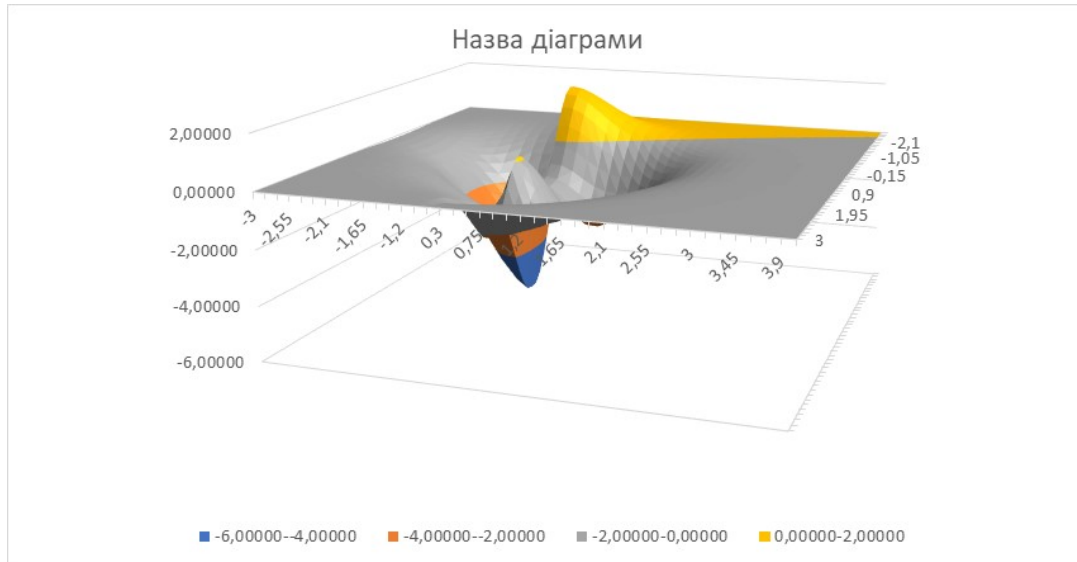


Рис. 2. Поверхня, що визначається функцією $F(x, y)$

В результаті застосування алгоритму викладеного вище із застосуванням операторів кросоверу, отримано на 10 кроці, необхідну точність розв'язку. Наближеним розв'язком є точка $(-0,3947; -0,9762)$. Значення функції в цій точці $-5,637$. На рис. 3 наведено послідовність кроків пошуку наближеного розв'язку.

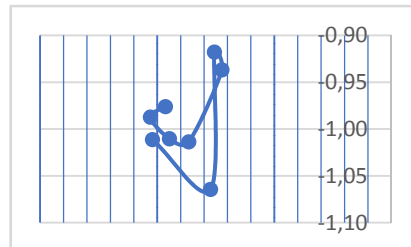


Рис. 3. схема покрокового пошуку розв'язку

На рис. 4. зображено залежності даної функції від кожної з координат і рух наближених значень функції до мінімального значення



Рис. 4. Залежності $F(x, y)$ від кожної з змінних

Нарешті, на рис. 5—7 наведено формування вкладених куль в області пошуку. Рис. 5 перших шість ітерацій. На рис. 6 — останні три ітерації.

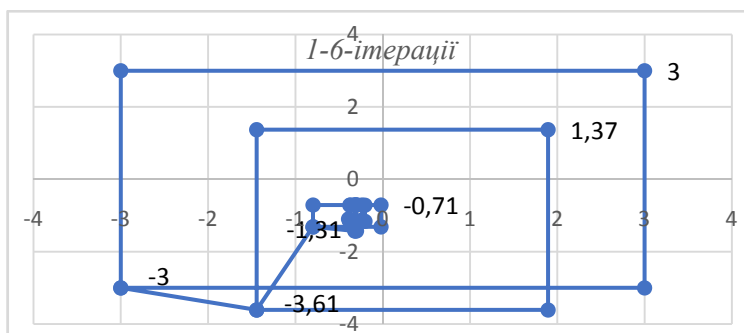


Рис. 5. Формування вкладених куль в області пошуку ітерації 1—6

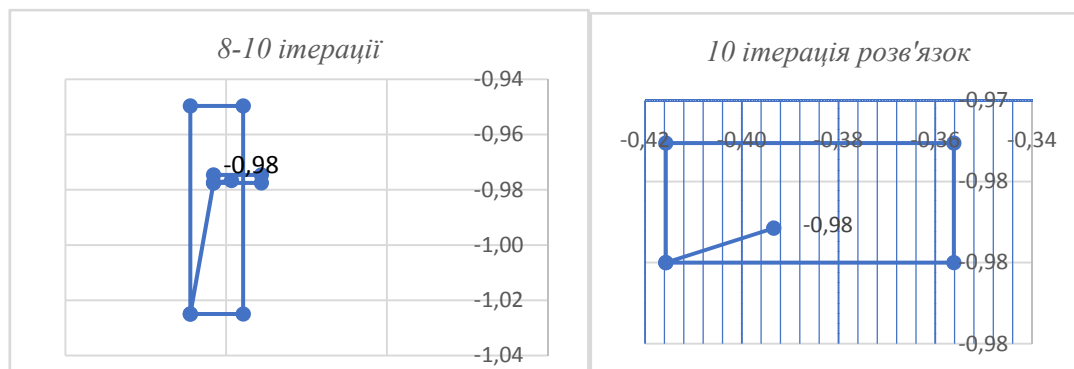


Рис. 6. Формування останніх трьох куль ітерації 7—9

Рис. 7. Зображення розв'язку

Висновки та перспективи подальших досліджень

В роботі наведено результати, що стосуються теоретичних основ генетичних алгоритмів. Представлено операторну модель формування популяції векторів-хромосом за рахунок застосування лінійних операторів кросоверу. Запропоновано один з алгоритмів пошуку розв'язків екстремальних задач.

Подальше дослідження даного напрямку полягає у розширенні класу задач, де буде ефективним застосування генетичних алгоритмів, а також удосконалення представленої операторної моделі.

Список використаної літератури

1. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности / Г. К. Вороновский, К. В. Махотило, С. Н. Петрашев, С. А. Сергеев. – Харьков: ОСНОВА, 1997. – 112 с.
2. Панченко, Т. В. Генетические алгоритмы: учебно-методическое пособие / Панченко, Т. В. // под ред. Ю. Ю. Тарасевича. – Астрахань: Издательский дом «Астраханский университет», 2007. – 87с.
3. Кононюк А. Е. Дискретно-непрерывная математика в 12-и кн./ Кононюк А. Е./ Алгоритмы. — Киев: «Освіта України», 2017. – Кн. 10, Ч.3 – 444 с.
4. Гантмахер Ф. Р. Теория матриц / Ф. Р. Гантмахер. – Москва: Наука, 1967. – 567 с.

OPERATOR MODEL OF RECOMBINATION IN GENETIC ALGORITHMS**Oleynik L.O.****Abstract**

In the process of applying genetic algorithms with chromosome vectors (approximations of solutions), certain actions are performed with the use of crossover operators and mutations ([1], [2], [3]). Investigation of the properties of these operators is an actual problem in describing the theoretical foundations of genetic algorithms.

The purpose of this work is to describe the class of linear operators functioning in the Cartesian product of the Euclidean space. This class of linear operators represents an algebraic model for constructing recombination operators (crossover and mutations).

The described class of involutive linear operators ([4]) allows us to form a population of descendants of chromosome vectors, basing on the small power of a set of pairs of parental vectors-chromosomes, while these descendants having the best properties in terms of values of the fitness function.

This approach allows you to find solutions to extreme problems quickly. In this work, one of the possible variants of the algorithm for finding an optimal solution is proposed.

The development of an algebraic approach to the theory of genetic algorithms involves several directions, namely:

- a study of the class of problems where the application of the algebraic model is efficient;
- a design of an effective algorithm for finding an optimal solution on the basis of an algebraic operator model.

References

- [1] Voronovsky G., Mahotilo K., Petrashev S., Sergeev S., «et al.».(1997) *Geneticheskiye algoritmy, iskusstvennyye neyronnyye seti i problemy virtual'noy real'nosti* [Genetic algorithms, artificial neural networks and virtual reality problems], Kharkov: OSNOVA [in Russian].
- [2] Panchenko, T., (2007). *Geneticheskiye algoritmy: uchebno-metodicheskoye posobiye* [Genetic Algorithms: teaching guide], Yu. Tarasevich, (Ed). Astrakhan: «Astrakhan University Publishing House» [in Russian].
- [3] Kononyuk A., *Dyskretno-neperervna matematyka v 12ti knigakh.* [Discrete-continuous mathematic in 12th book.] *Algorithms*, (Book 10, Part 3), Kiev: «Osvita of Ukraine», [in Russian].
- [4] Gantmacher F.R, (1967) *Teoriya matrits* [*The Theory of Matrices*]. Moscow: Science [in Russian].