Лабораторна робота №3

Розв’язати задачу безумовної оптимізації для квадратичної функції:

$$f(x) = ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2 \to min,$$

методом ***покоординатного спуску***. Коефіцієнти a, b, c, d, e задані в таблиці.

№	a	b	c	d	e
1.	1	2	2	–2	–3
2.	7	1	1	–16	–3
3.	2	2	1	–2	–6
4.	1	2	3	–2	–3
5.	3	2	5	–2	–3
6.	1	–1	8	2	–1
7.	4	2	5	–2	–3
8.	6	2	1	6	6
9.	1	–1	1	–2	1
10.	3	2	1	–2	–3
11.	3	2	2	–2	–3
12.	8	2	1	–3	–6
13.	3	2	3	–2	–3
14.	9	5	1	6	2
15.	2	2	4	–2	–3
16.	7	–1	1	7	–4
17.	3	1	1	1	5
18.	7	5	1	6	3
19.	4	2	3	–2	–3
20.	9	1	1	2	–1
21.	5	4	1	6	4
22.	1	2	4	–2	–3
23.	3	3	1	6	5
24.	6	–1	1	–3	5
25.	3	2	1	12	–6
26.	3	4	2	–2	4
27.	8	–2	1	–1	1
28.	2	2	3	–2	–3
29.	5	–2	1	–2	3

Метод покоординатного спуску

Постановка задачі: $f(x) \to min, \quad x \in E^n.$

Загальна схема ітераційних методів для розв'язання задачі безумовної мінімізації має вигляд

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \quad k = 0,1,2,.... \qquad (1)$$

де $p^{(k)}$ – напрямок спадання функції $f(x)$ (напрямок спуску) в точці $x^{(k)}$, вектор $p^{(k)} \in U(x^{(k)}, f)$, U – множина напрямків зменшення функції $f(x)$ у точці $x^{(k)}$, α_k – параметр, який регулює довжину кроку вздовж $p^{(k)}$.

У методі покоординатного спуску за напрямок спуску $p^{(k)}$ вибирається один з координатних векторів $e_1,...,e_n$, де $e_i = (0,0,...,\underset{i}{1},...,0)$. Унаслідок цього в точці $x^{(k)}$ на кожній ітерації змінюється лише одна з координат $x_i^{(k)}$.

Нехай $x^{(0)}$ – початкове наближення, а α_0 – деяке дійсне число, тоді за методом покоординатного спуску для

$$k = 1: \begin{cases} y_1^{(1)} = x^{(1)} = x^{(0)} + \alpha_0 e_1 \\ y_2^{(1)} = x^{(2)} = x^{(1)} + \alpha_1 e_2 \\ \dots\dots\dots\dots\dots\dots \\ y_n^{(1)} = x^{(n)} = x^{(n-1)} + \alpha_{n-1} e_n \end{cases} \qquad (2)$$

За формулами (2) буде реалізовано спуск за n **внутрішніх ітерацій** (n – розмірність простору E^n) з точки $x^{(0)}$ у точку $x^{(n)}$за ламаною, що складається з відрізків прямих, які паралельні координатним осям.

Спуск по усім n координатам за формулами (2) складають одну **зовнішню ітерацію.**

Наступна друга зовнішня ітерація виконується за формулами:

$$k = 2: \begin{cases} y_1^{(2)} = x^{(n+1)} = x^{(n)} + \alpha_n e_1 \\ y_2^{(2)} = x^{(n+2)} = x^{(n+1)} + \alpha_{n+1} e_2 \\ \dots\dots\dots\dots\dots\dots \\ y_n^{(2)} = x^{(2n)} = x^{(2n-1)} + \alpha_{2n-1} e_n \end{cases}, \quad \text{і таке інше.}$$

Нехай k – номер чергової зовнішньої ітерації ($k = 0,1,\dots$), i – номер тієї координати, по якій проводиться спуск, тобто номер внутрішньої ітерації (i =1,2,…,n). Тоді ітераційна рекурентна формула, яка визначає наступне наближення до точки мінімуму, запишеться у вигляді:

$$y_i^{(k)} = x^{(k-1)n+i} = x^{(k-1)n+i-1} + \alpha_{(k-1)n+i-1} \cdot e_i, k = 0,1,2,....; i = 1,2,...,n \qquad (3)$$

Після $i = n$ лічильник числа зовнішніх ітерацій k збільшується на 1, а значення $i = 1$.

До початку обчислень задається мале довільне число $\varepsilon > 0$. Ітераційний процес (3) закінчується, коли

$$\left\| y_n^{(k+1)} - y_n^{(k)} \right\| \le \varepsilon . \tag{4}$$

Існують різні способи вибору величини α_k на *k*-ій ітерації. Основна задача при виборі α_k в процесах мінімізації: забезпечити виконання нерівності

$$f\left(x^{(k+1)}\right) < f\left(x^{(k)}\right). \tag{5}$$

Розглянемо деякі способи вибору параметра α_k у методі покоординатного спуску.

Спосіб 1.

Вибір параметра α_k у методі покоординатного спуску здійснюємо з умови мінімізації цільової функції вздовж напрямку $p^{(k)}$. Наприклад, величини $\alpha_{(k-1)n+i-1}$ в методі (3) можна визначити з умови

$$f\left(x^{(k-1)n+i-1} + \alpha_{(k-1)n+i-1} \cdot e_i\right) = \min_{\alpha \in E^1} f\left(x^{(k-1)n+i-1} + \alpha \cdot e_i\right) \tag{6}$$

Спосіб 2.

Нехай $\alpha_k = \alpha_{k-1} > 0$. Обчислимо значення функції в точці $x = x^{(k)} + \alpha_k p^{(k)}$ та перевіримо виконання нерівності

$$f\left(x^{(k)} + \alpha_k p^{(k)}\right) < f\left(x^{(k)}\right). \tag{7}$$

Якщо нерівність (7) виконується, то приймемо $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$, $\alpha_{k+1} = \alpha_k$ і перейдемо до наступної $(k+2)$-ї ітерації , або виберемо $\alpha_k = 2\alpha_{k-1}$. Якщо наступне значення $f(x)$ менше за попереднє, то процес подвоєння можна продовжувати до тих пір, поки буде виконуватися умова (7).

У тому випадку, коли (7) не виконується, обчислимо значення функції $f(x)$ в точці $x = x^{(k)} - \alpha_k p^{(k)}$ і перевіримо виконання нерівності

$$f\left(x^{(k)} - \alpha_k p^{(k)}\right) < f\left(x^{(k)}\right). \tag{8}$$

Якщо (8) виконується або приймаємо $x^{(k+1)} = x^{(k)} - \alpha_k p^{(k)}$, $\alpha_{k+1} = \alpha_k$ і переходимо до наступної $(k+2)$-ї ітерації , або як раніше вибираємо $\alpha_k = 2\alpha_{k-1}$. Якщо значення $f(x)$ менше за попереднє то процес подвоєння можна продовжувати до тих пір, поки буде виконуватися умова (8).

Якщо за одну зовнішню ітерацію, яка складається з n внутрішніх ітерацій при переборі напрямків усіх координатних осей $e_1,...,e_n$ з кроком α_k реалізувалась хоча б одна вдала ітерація, то довжина кроку α_k не дробиться і зберігається на протязі наступного циклу з n ітерацій. Якщо серед останніх n ітерацій не було ні однієї вдалої, то крок α_k дробиться: $\alpha_k = \frac{1}{2}\alpha_{k-1}$ і здійснюється перехід до наступної зовнішньої ітерації .

Алгоритм методу покоординатного спуску

Початковий етап. Вибрати число $\varepsilon > 0$ для зупинки алгоритму, початкову точку $x^{(0)}$, покласти $t_1 = x^{(0)}, k = i = 1$ і перейти до основного етапу.

Основний етап.

Крок 1. Вибрати $\alpha_{(k-1)n+i-1}$ таке, щоб виконувалася нерівність

$$f\left(x^{(k-1)n+i-1} + \alpha_{(k-1)n+i-1} \cdot e_i\right) < f\left(x^{(k-1)n+i-1}\right)$$

Покласти $x^{(k-1)n+i} = x^{(k-1)n+i-1} + \alpha_{(k-1)n+i-1} \cdot e_i$. Якщо $i < n$, то замінити i на $i+1$ та повернутися до кроку 1, якщо $i = n$, то перейти до кроку 2.

Крок 2. Покласти $t_{k+1} = x^{(kn)}$. Якщо $\|t_{k+1} - t_k\| < \varepsilon$, то зупинитися. У протилежному випадку покласти $k = k+1$, $i = 1$ та перейти до кроку 1.

Алгоритм описаний.

Приклад. Розглянемо таку задачу: знайти мінімум функції

$$f(x) = f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2, x \in E^2$$

Розв'язання. Візьмемо $\varepsilon = 0,01$ та початкову точку $x^{(0)} = \left(x_1^{(0)}; x_2^{(0)}\right) = (0;3)$.

Величина α_k на кожній ітерації отримана внаслідок одновимірної оптимізації функції $f(x)$ уздовж одного з напрямків $e_i (i = 1,2)$, де $e_1 = (1;0), e_2 = (0;1)$

На *першому* кроці маємо $k = 1$, $t_1 = x^{(0)} = (0;3)$, $f(t_1) = 52$.

$i = 1$ $x^{(1)} = x^{(0)} + \alpha_0 e_1 = (0;3) + \alpha_0(1;0) = (\alpha_0;3)$; $f\left(x^{(1)}\right) = \left(\alpha_0 - 2\right)^4 + \left(\alpha_0 - 6\right)^2$.

Щоб знайти α_0 можемо застосувати один з методів одновимірної оптимізації функції $f\left(x^{(1)}\right)$ за змінною α_0, або використати класичний метод з використанням похідної: $f'_{\alpha_0}\left(x^{(1)}\right) = 4(\alpha_0 - 2)^3 + 2(\alpha_0 - 6) = 0$. Звідки маємо $\alpha_0 \approx 3.1282$. Тоді $x^{(1)} = (3.1282;3)$.

$i = 2$ $x^{(2)} = x^{(1)} + \alpha_1 e_2 = (3.1282;3) + \alpha_1(0;1) = (3.1282;3 + \alpha_1)$;

$$f\left(x^{(2)}\right) = \left(3.1282\right)^4 + \left(3.1282 - 2(3 + \alpha_1)\right)^2.$$

Знаходимо α_1. Маємо $f'_{\alpha_1}\left(x^{(2)}\right) = -4\left(3.1282 - 2(3 + \alpha_1)\right) = 0 \Rightarrow \alpha_1 = -1.4359$. Отже, $x^{(2)} = (3.1282;1.5641)$. Ми виконали дві внутрішні ітерації та одну зовнішню.

Нехай тепер $k=2$, $t_2 = x^{(2)} = (3.1282;1.5641)$. Тоді $f(t_2) = 1.6201$. Переконуємося в тому, що $f(t_2) < f(t_1)$. Перевіряємо критерій зупинки $\|t_2 - t_1\|_{E^2} < \varepsilon$:

$$\|t_2 - t_1\| = \sqrt{(3.1282-0)^2 + (1.5641-3)^2} = 3.442 > 0.01.$$

Критерій не виконується, переходимо до наступної ітерації.

$i=1 \quad x^{(3)} = x^{(2)} + \alpha_2 e_1 = (3.1282, 1.5641) + \alpha_2(1;0) = (3.1282 + \alpha_2, 1.5641)$;

$$f\left(x^{(3)}\right) = \left(1.1282 + \alpha_2\right)^4 + \alpha_2{}^2.$$

Знаходимо α_2. Маємо $f'_{\alpha_2}\left(x^{(3)}\right) = 4\left(1.1282 + \alpha_2\right)^3 + 2\alpha_2 = 0 \Rightarrow \alpha_2 = -0.4987$. Отже, $x^{(3)} = (2.6295, 1.5641)$.

$i=2 \quad x^{(4)} = x^{(3)} + \alpha_3 e_2 = (2.6295; 1.5641) + \alpha_3(0;1) = (2.6295; 1.5641 + \alpha_3)$

$$f\left(x^{(4)}\right) = \left(0.6295\right)^4 + \left(2\alpha_3 + 0.499\right)^2$$

По аналогії з попереднім, дістаємо, що $\alpha_3 = -0.2495$. Тоді $x^{(4)} = (2.6295; 1.3146)$.

Покладемо $k=3$. Тоді $t_3 = x^{(4)} = (2.6295; 1.3146)$ та $f(t_3) = 0.157$. Переконуємося в тому, що $f(t_3) < f(t_2)$. Перевіряємо критерій зупинки $\|t_3 - t_2\|_{E^2} < \varepsilon$:

$$\|t_3 - t_2\| = \sqrt{(3.1282-2.6295)^2 + (1.5641-1.3146)^2} = 0.5576 > 0.01.$$

Критерій не виконується. Продовжуємо обчислення далі.

$i=1 \quad x^{(5)} = x^{(4)} + \alpha_4 e_1 = (2.6295; 1.3146) + \alpha_4(1;0) = (2.6295 + \alpha_4; 1.3146)$

$$f\left(x^{(5)}\right) = \left(0.6295 + \alpha_4\right)^4 + \left(0.0003 + \alpha_4\right)^2.$$

Знаходимо $f'_{\alpha_4}\left(x^{(5)}\right) = 0 \Rightarrow \alpha_4 = -0.1809$. Тоді $x^{(5)} = (2.4486; 1{,}3146)$.

$i=2 \quad x^{(6)} = x^{(5)} + \alpha_5 e_2 = (2.4486; 1.3146) + \alpha_5(0;1) = (2.4486; 1.3146 + \alpha_5)$

$$f\left(x^{(6)}\right) = \left(0.4486\right)^4 + \left(2\alpha_5 + 0.1806\right)^2.$$

Знаходимо $f'_{\alpha_5}\left(x^{(6)}\right) = 0 \Rightarrow \alpha_5 = -0.0903$. Тоді $x^{(6)} = (2.4486; 1.2243)$

Покладемо $k=4$. Тоді $t_4 = x^{(6)} = (2.4486; 1.2243)$ та $f(t_4) = 0.0405$. Переконуємося в тому, що $f(t_4) < f(t_3)$. Перевіряємо критерій зупинки $\|t_4 - t_3\|_{E^2} < \varepsilon$:

$$\|t_4 - t_3\| = \sqrt{(2.4486-2.6295)^2 + (1.2243-1.3146)^2} = 0.2022 > 0.01 \quad \text{і таке інше.}$$

Після 13 зовнішніх ітерацій отримана точка (2.1593;1.0797), значення цільової функції в якій дорівнює 0.0006. Обчислення наведені в таблиці.

k	t_k	$f(t_k)$	i	e_i	$x^{(k-1)n+i-1}$	$\alpha_{(k-1)n+i-1}$	$\lVert t_{k+1}-t_k \rVert$
1	(0;3)	52	1	(1; 0)	(0;3)	3.1282	–
			2	(0; 1)	(3.1282; 3)	– 1.4359	
2	(3.1282;1.5641)	1.6201	1	(1; 0)	(3.1282;1.5641)	– 0.4987	3.442
			2	(0; 1)	(2.6295,1.5641)	– 0.2495	
3	(2.6295;1.3146)	0.157	1	(1; 0)	(2.6295;1.3146)	– 0.1809	0.5576
			2	(0; 1)	(2.4486;1.3146)	–0.0903	
4	(2.4486;1.2243)	0.0405	1	(1; 0)	(2.4486;1.2243)	–0.0912	0.2022
			2	(0; 1)	(2.3574;1.2243)	–0.0456	
5	(2.3574;1.1787)	0.0163	1	(1; 0)	(2.3574;1.1787)	– 0.0552	0.102
			2	(0; 1)	(2.3022; 1.1787)	– 0.0276	
6	(2.3022; 1.1511)	0.0083	1	(1; 0)	(2.3022; 1.1511)	– 0.0372	0.0617
			2	(0; 1)	(2.2650; 1.1511)	–0.0186	
7	(2.2650; 1.1325)	0.0049	1	(1; 0)	(2.2650; 1.1325)	– 0.027	0.0416
			2	(0; 1)	(2.2380;1.1325)	– 0.0135	
8	(2.2380;1.119)	0.0032	1	(1; 0)	(2.2380;1.119)	–0.0206	0.0302
			2	(0; 1)	(2.2174;1.119)	–0.0103	
9	(2.2174;1.1087)	0.0022	1	(1; 0)	(2.2174;1.1087)	–0.0163	0.023
			2	(0; 1)	(2.2011;1.1087)	–0.0082	
10	(2.2011;1.1005)	0.0016	1	(1; 0)	(2.2011;1.1005)	–0.0133	0.0182
			2	(0; 1)	(2.1878;1.1005)	–0.0066	
11	(2.1878;1.0939)	0.0012	1	(1; 0)	(2.1878;1.0939)	–0.011	0.0148
			2	(0; 1)	(2.1768;1.0939)	–0.0055	
12	(2.1768;1.0884)	0.001	1	(1; 0)	(2.1768;1.0884)	–0.0094	0.0123
			2	(0; 1)	(2.1674;1.0884)	–0.0047	
13	(2.1674;1.0837)	0.0008	1	(1; 0)	(2.1674;1.0837)	–0.0081	0.0105
			2	(0; 1)	(2.1593;1.0837)	–0.004	
14	(2.1593;1.0797)	0.0006					0.009

Зазначимо, що оптимальним розв’язком цієї задачі є точка $\hat{x} = (2;1)$, в якій значення цільової функції дорівнює нулю. Укажемо також, що помітне зменшення значень функції отримано за декілька перших ітерацій, тоді як на останніх ітераціях процес явно уповільнюється.