Лабораторна робота №4

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 ***найшвидшого спуску***. Коефіцієнти a, b, c, d, e задані в таблиці, $\varepsilon = 0.001$.

№	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

Градієнтні методи. Метод найшвидшого спуску

Розглядається задача безумовної оптимізації:

$$f(x) \to \min, \quad x \in E^n. \tag{1}$$

Функція $f(x)$ диференційована на E^n. В градієнтних методах мінімізації за напрямок спуску $p^{(k)}$ з точки $x^{(k)}$ (нагадаємо формулу: $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, k = 0,1,2,....$) вибирається антиградієнт функції $f(x)$ у точці $x^{(k)}$, тобто

$$x^{(k+1)} = x^{(k)} - \alpha_k f'(x^{(k)}), \quad \alpha_k > 0, k = 0,1,2,.... \tag{2}$$

або в координатній формі

$$x_i^{(k+1)} = x_i^{(k)} - \alpha_k \frac{\partial f(x^{(k)})}{\partial x_i}, i = \overline{1,n}.$$

На промені $X = \left\{x \in E^n \mid x = x^{(k)} - \alpha f'(x^{(k)}), \alpha > 0\right\}$, який направлений за антиградієнтом функції $f(x)$ у точці $x^{(k)}$ введемо функцію

$$g_k(\alpha) = f\left(x^{(k)} - \alpha f'(x^{(k)})\right), \ \alpha \ge 0 \tag{3}$$

і визначимо α_k з умови

$$g_k(\alpha_k) = \min_{\alpha \ge 0} g_k(\alpha), \alpha_k > 0. \tag{4 }$$

Метод (2),(3), в якому кроковий множник α_k визначається з умови мінімізації функції $f(x)$ вздовж напрямку градієнту, носить назву методу **найшвидшого спуску**.

Алгоритм методу

Початковий етап.

Вибрати довільну початкову точку $x^0 \in E^n$ і параметр точності пошуку $\varepsilon > 0$. Обчислити $\operatorname{grad} f = f'(x)$ і покласти $k = 0$.

Основний цикл.

Крок1.Обчислити $\operatorname{grad} f\left(x^{(k)}\right)$.

Крок2.Якщо $\left|\operatorname{grad} f\left(x^{(k)}\right)\right| \le \varepsilon$, то покласти $\hat{x} = x^{(k)}$ і зупинити обчислення; інакше перейти на крок 3.

Крок3.Обчислити кроковий множник α_k з умови (4) для функції (3).

Крок4.Обчислити наступне наближення

$$x^{(k+1)} = x^{(k)} - \alpha_k f'(x^{(k)}), \quad \alpha_k > 0.$$

Крок5. Покласти $k = k+1$ і перейти на крок 1.

Приклад. Знайти мінімум функції

$$f(x_1,x_2)=(2x_1-x_2)^2+2(x_2-2)^2+x_1x_2,\ x=(x_1,x_2)\in E^2$$

*Розв'язання***.** Покладемо $\varepsilon=0.01$. В якості початкового наближення візьмемо точку $x^{(0)}=\left(x_1^{(0)};x_2^{(0)}\right)=(2.00;2.00)$. В якості критерію зупинки: $\left\|f'\left(x^{(k)}\right)\right\|_2\le\varepsilon$.

Точку мінімуму будемо шукати за формулою $x^{(k+1)}=x^{(k)}-\alpha_k f'(x^{(k)})$, де множник $\alpha_k>0$ знаходимо з умови $g_k(\alpha_k)=\min\limits_{\alpha\ge0}g_k(\alpha)$, $g_k(\alpha)=f\left(x^{(k)}-\alpha f'(x^{(k)})\right)$.

Знаходимо $\operatorname{grad} f(x)$: $\operatorname{grad} f(x)=f'(x)=\left(\dfrac{\partial f}{\partial x_1};\dfrac{\partial f}{\partial x_2}\right)=\left(8x_1-3x_2;\ -3x_1+6x_2-8\right)$.

Знаходимо градієнт в початковій точці

$$\operatorname{grad} f(x)\Big|_{x=x^{(0)}}=f'\left(x^{(0)}\right)=\left(\frac{\partial f}{\partial x_1};\frac{\partial f}{\partial x_2}\right)\Bigg|_{(2;2)}=(10;-2).$$

Перевіряємо $\left\|f'\left(x^{(0)}\right)\right\|_2=\sqrt{10^2+2^2}=\sqrt{104}=10.198>0.01$.

Ітерація 1.

$$x^{(1)}=x^{(0)}-\alpha_0 f'\left(x^{(0)}\right)=(2;2)-\alpha_0(10;-2)=(2-10\alpha_0;2+2\alpha_0).$$

Складаємо функцію $g_0(\alpha)=f\left(x^{(0)}-\alpha f'(x^{(0)})\right)$ та досліджуємо її на мінімум.

$$g_0(\alpha)=f\left(x^{(0)}-\alpha f'(x^{(0)})\right)=f(2-10\alpha;2+2\alpha)=(2-22\alpha)^2+8\alpha^2+(2-10\alpha)(2+2\alpha).$$

$$g_0'(\alpha)=0\Leftrightarrow-104+944\alpha=0\Rightarrow\alpha=0.11$$

Отже, $\alpha_0=0.11$. Тоді $x^{(1)}=x^{(0)}-\alpha_0 f'\left(x^{(0)}\right)=(2;2)-0.11\cdot(10;-2)=(0.9;2.22)$.

Обчислюємо $f\left(x^{(0)}\right)=f(2;2)=8$ та $f\left(x^{(1)}\right)=f(0.9;2.22)=2.271$. Умова монотонності $f\left(x^{(1)}\right)<f\left(x^{(0)}\right)$ виконується. Перевіряємо критерій зупинки:

Обчислюємо $f'\left(x^{(1)}\right)=\left(8x_1^{(1)}-3x_2^{(1)};\ -3x_1^{(1)}+6x_2^{(1)}-8\right)\Big|_{(0.9;\,2.22)}=(0.54;2.62)$.

Перевіряємо $\left\|f'\left(x^{(1)}\right)\right\|_2=\sqrt{0.54^2+2.62^2}=2.675>0.1$.

Ітерація 2.

$$x^{(2)}=x^{(1)}-\alpha_1 f'\left(x^{(1)}\right)=(0.9;2.22)-\alpha_1(0.54;2.62)=(0.9-0.54\alpha_1;2.22-2.62\alpha_1)$$

Складаємо функцію $g_1(\alpha)=f\left(x^{(1)}-\alpha f'(x^{(1)})\right)$ та досліджуємо її на мінімум.

$$g_1(\alpha)=f\left(x^{(1)}-\alpha f'(x^{(1)})\right)=(-.42+1.54\alpha)^2+2(.22-2.62\alpha)^2+(.9-.54\alpha)(2.22-2.62\alpha)$$

Тоді $g_1'(\alpha)=35.03\alpha-7.156\Rightarrow\alpha=0.204$. Отже $\alpha_1=0.204$.

Тоді $x^{(2)} = x^{(1)} - \alpha_1 f'\left(x^{(1)}\right) = \left(0.79; 1.686\right)$. Знаходимо $f\left(x^{(2)}\right) = f\left(0.79; 1.686\right) = 1.54$. Умова монотонності $f\left(x^{(2)}\right) < f\left(x^{(1)}\right)$ виконується.

Перевіряємо критерій зупинки. Обчислюємо $f'\left(x^{(2)}\right) = (1.262; -0.256)$. Тоді $\left\|f'\left(x^{(2)}\right)\right\|_2 = 1.288 > 0.01$.

Продовжуємо обчислення. Результати наведені в таблиці. Процедура була зупинена при виконанні умови $\left\|f'\left(x^{(k)}\right)\right\| < \varepsilon\ (\varepsilon = 0.1)$. Відзначимо, що метод найшвидшого спуску, як і всі градієнтні методи, добре працює на початковій стадії процесу мінімізації, а в околі стаціонарної точки його збіжність уповільнюється. Слід зауважити, що дана задача має точний розв'язок $\hat{x} = \left(\frac{24}{39}; \frac{64}{39}\right) \approx \left(0.6154; 1.641)\right)$ та $f_{\min} = \frac{56}{39} \approx 1.4359$.

k	$x^{(k)}$	$f\left(x^{(k)}\right)$	$-f'\left(x^{(k)}\right)$	α_k	$\left\|f'\left(x^{(k)}\right)\right\|$	$x^{(k+1)}$
0	(2.00; 2.00)	8.00	(–10; 2)	0.11	10.198	(0.9; 2.22)
1	(0.9; 2.22)	2.271	$(-0.54; -2.62)$	0.204	2.675	(0.79; 1686)
2	(0.79; 1686)	1.54	$(-1.262; 0.256)$	0.11	1.288	(0.651; 1.714)
3	(0.651; 1.714)	1.449	$(-0.067; -0.329)$	0.204	0.336	(0.637; 1.647)
4	(0.637; 1.647)	1.4376	$(-0.159; 0.033)$	0.11	0.163	(0.62; 1.65)
5	(0.62; 1.65)	1.4361	$(-0.008; -0.041)$	0.203	0.042	(0.618; 1.642)
6	(0.618; 1.642)	1.4359	$(-0.02; 0.004)$	0.11	0.02	(0.616; 1.642)
7	(0.616; 1.642)	1.4359	$(-0.001; -0.005)$		0.005	