

Метод градієнтного спуску

Нехай маємо систему рівнянь

$$f_i(x_1, x_2, \dots, x_n) = 0, \quad (i = 1, 2, \dots, n) \quad (1)$$

чи в матричній формі

$$F(\bar{x}) = 0, \quad (2)$$

де $F = (f_1, f_2, \dots, f_n)^T$, $\bar{x} = (x_1, x_2, \dots, x_n)$.

Припустимо, що функції $f_i(x_1, x_2, \dots, x_n)$ дійсні й неперервно-диференційовані в їхній загальній області визначення. Розглянемо функцію

$$\Phi(\bar{x}) = \sum_{i=1}^n [f_i(x_1, x_2, \dots, x_n)]^2 = (F(\bar{x}), F(\bar{x})). \quad (3)$$

Очевидно, що кожний розв'язок системи (1) перетворює на нуль функцію $\Phi(\bar{x})$; навпаки, числа $x_1, x_2, \dots, x_n$, для яких функція $\Phi(\bar{x})$ дорівнює нулю, є коренями системи (1). Таким чином, задача зводиться до знаходження мінімуму скалярної функції багатьох змінних $\Phi(\bar{x})$.

Одним з методів мінімізації функцій багатьох змінних є метод градієнтного спуску. Якщо $\bar{x}^{(k)}$ – деяке наближення до розв'язку системи, то в методі градієнтного спуску ми одержуємо нове наближення $\bar{x}^{(k+1)}$, рухаючись за напрямком найбільшої миттєвої швидкості зміни функції $\Phi(\bar{x})$ в точці $\bar{x}^{(k)} (\text{grad} \Phi(\bar{x}^{(k)}))$ до точки, де значення $\Phi(\bar{x}^{(k+1)})$ мінімальне, тобто

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \lambda_k \text{grad} \Phi(\bar{x}^{(k)}), \quad (4)$$

де λ_k вибирається з умови мінімуму $\Phi(\bar{x}^{(k+1)})$.

Якщо λ – мала величина, квадратом і вищими ступенями якої можна знехтувати, то, розкладаючи функції $f_i(x_1, x_2, \dots, x_n)$ за степенями λ з точністю до лінійних членів і виражаючи $\text{grad} \Phi(\bar{x})$ через матрицю Якобі $W(\bar{x})$, одержимо таке представлення розрахункової формули методу градієнтного спуску

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \mu_k W^T(\bar{x}^{(k)}) F(\bar{x}^{(k)}), \quad (k = 0, 1, 2, \dots), \quad (5)$$

$$\text{де } W(\bar{x}) = \frac{\partial F}{\partial \bar{x}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \text{ матриця Якобі вектор-функції } F(\bar{x}).$$

$$\mu_k = \frac{(F(\bar{x}^{(k)})^T W(\bar{x}^{(k)}) W^T(\bar{x}^{(k)}) F(\bar{x}^{(k)}))}{(W(\bar{x}^{(k)}) W^T(\bar{x}^{(k)}) F(\bar{x}^{(k)}))^T W(\bar{x}^{(k)}) W^T(\bar{x}^{(k)}) F(\bar{x}^{(k)})} . \quad (6)$$

Слід зазначити, що ітераційний процес, побудований за методом градієнтного спуску, збігається до точного розв'язку, якщо початкове наближення $\bar{x}^{(0)}$ обране з досить малого околу кореня.

Приклад.

Методом градієнтного спуску приблизно обчислити корені системи

$$\begin{cases} x + x^2 - 2yz = 0,1; \\ y - y^2 + 3xz = -0,2; \\ z + z^2 + 2xy = 0,3, \end{cases}$$

розташовані в околі початку координат.

Розв'язок. Маємо $\bar{x}^{(0)} = (0,0,0)^T$.

$$\text{Тут } F = \begin{pmatrix} x + x^2 - 2yz - 0,1 \\ y - y^2 + 3xz + 0,2 \\ z + z^2 + 2xy - 0,3 \end{pmatrix} \text{ і } W = \begin{pmatrix} 1 + 2x & -2z & -2y \\ 3z & 1 - 2y & 3x \\ 2y & 2x & 1 + 2z \end{pmatrix}.$$

Підставляючи нульове наближення, будемо мати:

$$F(\bar{x}^{(0)}) = (-0,1; 0,2; -0,3)^T \text{ і } W(\bar{x}^{(0)}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E$$

За формулами (5) і (6) одержуємо перше наближення

$$\mu_0 = \frac{(F(\bar{x}^{(0)})^T F(\bar{x}^{(0)}))}{(F(\bar{x}^{(0)})^T F(\bar{x}^{(0)}))} = 1 \text{ і } \bar{x}^{(1)} = \bar{x}^{(0)} - 1 \cdot E \cdot F(\bar{x}^{(0)}) = (0,1; -0,2; 0,3)^T.$$

Аналогічно знаходимо друге наближення $\bar{x}^{(2)}$.

Маємо:

$$F(\bar{x}^{(1)}) = \begin{pmatrix} 0,13 \\ 0,05 \\ 0,05 \end{pmatrix}, \quad W(\bar{x}^{(1)}) = \begin{pmatrix} 1,2 & -0,6 & 0,4 \\ 0,9 & 1,4 & 0,3 \\ -0,4 & 0,2 & 1,6 \end{pmatrix}.$$

Звідси:

$$W^T(\bar{x}^{(1)}) F(\bar{x}^{(1)}) = \begin{pmatrix} 0,181 \\ 0,002 \\ 0,147 \end{pmatrix} \text{ та } W(\bar{x}^{(1)}) W^T(\bar{x}^{(1)}) F(\bar{x}^{(1)}) = \begin{pmatrix} 0,2748 \\ 0,2098 \\ 0,1632 \end{pmatrix}$$

Отже,

$$\mu_1 = \frac{0,13 \cdot 0,2748 + 0,05 \cdot 0,2098 + 0,05 \cdot 0,1632}{0,2748^2 + 0,2098^2 + 0,1632^2} = \frac{0,054374}{0,14619797} = 0,3719 \text{ і}$$

$$\bar{x}^{(2)} = \begin{pmatrix} 0,1 \\ -0,2 \\ 0,3 \end{pmatrix} - 0,3719 \begin{pmatrix} 0,181 \\ 0,002 \\ 0,147 \end{pmatrix} = \begin{pmatrix} 0,0327 \\ -0,2007 \\ 0,2453 \end{pmatrix}.$$

Для контролю обчислимо відхил

$$F(\bar{x}^{(2)}) = \begin{pmatrix} 0,032 \\ -0,017 \\ -0,007 \end{pmatrix}.$$

Завдання до лабораторної роботи № 7

Методом градієнтного спуску при заданому початковому наближенні $\bar{x}^{(0)}$ знайдіть наближені корені системи

$$\begin{cases} a_{11} \cdot x_1^2 + a_{12} \cdot x_2 \cdot x_3 + a_{13} \cdot x_2 \cdot x_4 + a_{14} \cdot x_1 = b_1, \\ a_{21} \cdot x_2^2 + a_{22} \cdot x_1 \cdot x_3 + a_{23} \cdot x_3 \cdot x_4 + a_{24} \cdot x_2 = b_2, \\ a_{31} \cdot x_3^2 + a_{32} \cdot x_1 \cdot x_2 + a_{33} \cdot x_2 \cdot x_4 + a_{34} \cdot x_3 = b_3, \\ a_{41} \cdot x_4^2 + a_{42} \cdot x_1 \cdot x_2 + a_{43} \cdot x_2 \cdot x_3 + a_{44} \cdot x_4 = b_4. \end{cases}$$

За умову закінчення ітераційного процесу прийміть $\|F(\bar{x}^{(i)})\|_{\infty} < 0,01$.

Варіанти завдань

№ варі- анта	Матриця коефіцієнтів системи a_{ij}				Стовпець вільних членів b_i	Початкове наближення коренів $x_i^{(0)}$
1	-4,1	2,4	1,3	-0,7	-18,12338	-1,1
	0,6	-1,2	3,1	2,9	4,89797	2,2
	-3,4	0,4	-0,1	-1,1	-30,61106	-3,3
	0,5	-2,8	1,4	3,3	-3,70831	0,0
2	2,4	3,1	-0,8	1,5	-4,06003	1,3
	-1,8	-4,2	2,6	0,7	-16,26196	-1,1
	0,4	1,9	-0,5	-2,6	-6,20337	3,0
	-3,2	0,2	1,1	0,9	-3,44529	0,3
3	1,5	0,9	-2,2	3,2	11,10436	-0,5
	-0,4	3,6	1,7	-2,8	-10,36585	1,0
	2,1	-0,3	-2,9	1,4	-0,27533	-2,5
	-1,6	2,5	0,7	-0,1	-16,43394	3,0
4	-2,8	0,6	1,9	-3,1	-2,74332	0,9
	1,4	-2,2	-0,8	1,2	6,78768	2,1
	-0,1	1,5	2,6	-0,9	-0,35866	1,6
	3,5	-0,6	-1,7	2,4	-7,91985	-0,1

5	1,1	4,0	-2,3	-0,5	7,25672	1,6
	-3,4	-0,7	1,2	2,4	-9,89640	-0,3
	2,5	1,9	-3,0	0,8	16,31988	2,5
	-0,1	-0,8	2,5	-1,0	-1,00392	-1,7
6	4,5	1,2	-0,6	2,1	19,99632	-2,5
	-0,6	-2,0	1,3	-3,2	-12,18632	1,6
	1,1	0,8	-2,1	0,2	-14,24772	-0,7
	-3,0	-2,7	0,4	-1,3	-22,48474	3,3
7	2,4	0,9	-1,8	0,2	0,53615	0,9
	-1,2	-3,0	-2,3	1,7	-10,59093	-2,3
	0,1	1,8	0,4	-2,2	-5,46305	1,5
	-2,0	-0,5	1,6	0,8	-5,47281	-0,8
8	0,8	-1,1	-0,4	2,7	17,47656	3,1
	-1,0	3,4	1,8	-0,2	-33,53480	-0,6
	2,1	-0,6	-2,0	1,9	9,65912	-2,5
	0,0	2,5	0,4	-3,1	-9,05794	2,0
9	-3,4	0,0	1,2	-2,1	-3,65902	-0,1
	2,1	-1,9	-0,4	1,6	0,76205	0,2
	0,0	3,4	1,1	-0,7	-1,16126	1,8
	-4,0	1,9	-0,2	2,4	-15,32033	-1,4
10	0,6	-2,1	3,0	-0,1	-0,57180	-0,2
	-1,9	3,4	-1,5	2,4	6,32067	0,8
	2,0	-0,8	2,4	-0,5	5,10568	-0,3
	-0,4	1,9	-0,2	2,5	1,30491	2,3
11	2,7	0,6	-1,5	-0,8	12,31117	2,6
	-1,5	-1,4	-0,8	2,0	9,43067	2,1
	0,3	2,2	0,5	-1,8	18,78237	-2,9
	1,6	-0,1	-2,4	0,5	13,59151	0,4
12	-0,5	2,9	0,2	-1,4	2,02812	0,7
	1,0	-0,8	-2,5	0,6	6,82786	-2,0
	-2,6	1,5	0,8	-2,2	-5,96430	-0,4
	0,8	-2,4	1,2	0,5	10,15452	1,9
13	-1,2	0,6	-0,5	1,8	-1,89552	-1,7
	0,5	-2,2	1,4	-0,9	-0,51975	0,9
	-2,4	1,6	-1,7	2,5	-4,79314	2,4
	1,8	-0,5	2,0	-1,2	17,82578	-2,6
14	-2,5	0,2	1,8	-0,2	-32,82288	3,6
	0,6	1,5	-0,4	1,2	-3,47712	-1,4
	-0,5	-2,4	-1,6	0,4	14,80896	-0,5
	1,4	-3,2	0,5	-2,0	18,41504	2,5
15	0,9	-1,0	2,2	-0,5	1,24605	-2,4
	-1,5	2,0	-1,4	0,4	-3,02800	0,6
	3,0	-0,5	0,8	-2,6	1,17404	1,4
	-2,8	1,2	-1,0	0,5	-8,58912	-1,6

16	1,9	-0,7	-1,5	2,0	4,99613	0,9
	-0,6	2,4	0,4	-0,5	-9,27880	-2,5
	-3,2	-0,1	2,5	1,4	-43,01664	-2,9
	1,1	0,5	1,8	-2,2	11,28333	1,8
17	0,6	-2,8	3,4	0,5	4,43632	-1,3
	-2,5	1,2	-0,6	1,8	4,15527	0,3
	1,4	-1,5	2,2	-0,4	9,58114	-2,5
	-0,5	2,0	-1,6	0,2	0,39000	0,0
18	1,2	-0,5	2,8	-3,2	-2,42752	3,2
	-2,6	1,4	-3,0	0,5	3,16474	-1,3
	0,5	-2,1	1,1	-2,6	8,81720	0,4
	-4,1	0,6	-0,5	1,2	-26,58144	-2,5
19	1,6	-0,5	3,2	2,1	-2,90940	-0,7
	-2,0	4,1	-0,8	-1,5	-3,30594	-2,0
	0,8	-1,5	2,4	0,1	-1,82784	-1,0
	-1,2	0,4	-2,5	-0,6	-6,36048	0,0
20	-0,4	-2,5	1,8	-2,0	0,42576	3,0
	1,1	3,2	-2,6	0,5	-30,21898	0,4
	-3,5	-0,1	1,4	-2,6	-16,27821	-2,6
	-4,0	0,2	-0,5	1,1	-6,42578	-1,0
21	2,4	-1,6	-2,5	0,7	-2,28288	0,6
	1,5	0,9	1,2	2,6	1,84254	-1,5
	-0,8	1,1	2,4	-0,5	-7,93506	1,1
	1,2	-2,8	-0,5	2,0	11,00036	1,9
22	0,4	1,9	-2,4	0,5	0,30678	1,2
	-1,5	0,6	1,2	2,4	-4,21839	2,5
	2,0	-4,1	0,5	1,6	-8,45331	-0,5
	0,1	3,0	1,4	-2,5	1,53789	1,6
23	1,7	0,6	2,1	-1,5	-5,66937	0,7
	-0,1	2,5	3,0	1,8	-4,99050	1,5
	2,6	-1,2	2,1	-0,5	13,15266	-2,2
	1,5	0,2	-1,7	2,2	8,54615	1,0
24	4,0	-0,5	2,1	0,2	18,84720	-2,5
	2,6	1,3	0,6	-1,5	3,40392	2,0
	-0,5	3,0	1,2	-2,4	-22,47560	2,2
	1,9	0,8	-0,5	1,2	-6,43244	-0,3
25	-2,5	1,6	2,1	0,1	-37,97268	3,5
	0,8	1,1	-3,0	1,5	2,22006	0,0
	2,1	-1,2	0,5	-2,7	6,74865	-1,0
	1,4	0,5	2,2	0,3	6,29984	2,1
26	0,5	2,7	-0,1	1,2	-0,86703	1,0
	3,0	-2,2	1,5	2,6	14,43240	-2,5
	1,9	0,5	2,4	-0,4	-12,59653	0,2
	-2,8	1,3	0,7	2,5	-9,05877	1,5

27	1,1	-0,6	2,2	-1,5	1,68870	2,2
	0,9	3,0	-1,6	2,4	10,30449	1,4
	-2,0	0,5	2,7	1,1	1,29780	1,0
	2,5	1,1	0,4	-1,9	2,89248	0,6
28	2,3	1,4	0,7	-1,1	4,26706	-0,7
	1,7	0,9	1,3	2,2	24,95169	3,5
	-0,2	2,6	1,5	-0,3	-11,60306	0,6
	1,4	-0,4	2,0	1,5	6,61972	-1,5
29	0,9	-0,9	1,7	2,0	1,28439	1,5
	4,1	1,4	-0,4	2,7	10,74725	-2,0
	-1,1	2,2	-3,3	1,1	-17,52025	-1,0
	0,4	3,0	1,5	-2,2	-8,26131	0,7
30	1,2	-0,7	3,1	2,4	1,15208	2,5
	2,5	1,4	-2,1	0,9	13,88584	1,0
	0,1	3,0	1,6	-2,2	-0,20264	1,6
	-2,4	0,7	2,5	1,8	-11,25706	-2,0