УДК 519.85

[1] **В.О. Стовба**
Доктор філософії (PhD), молодший науковий співробітник
[2] **О.О. Жмуд**
Аспірант
[1,2] *Інститут кібернетики імені В.М. Глушкова НАН України, Київ*

ГРАДІЄНТНИЙ МЕТОД З КРОКОМ ПОЛЯКА ДЛЯ МІНІМІЗАЦІЇ КВАДРАТИЧНИХ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

Розглядається градієнтний метод з кроком Поляка в перетвореному просторі змінних (метод B)

$$x_{k+1} = x_k - h_k B \frac{B^T g_f(x_k)}{\left\| B^T g_f(x_k) \right\|}, \quad h_k = \frac{m\left(f(x_k) - f^* \right)}{\left\| B^T g_f(x_k) \right\|},\ k = 0,1,2,..., \tag{1}$$

який є частинним випадком субградієнтного методу з кроком Поляка [1,2]. Метод (1) можна використати для знаходження точки $x_\varepsilon^* \in R^n$ такої, що $f\left(x_\varepsilon^*\right) \le f^* + \varepsilon$, де $f(x)$, $x \in R^n$ – гладка опукла функція, $f^* = f\left(x^*\right)$, $g_f(x)$ – її градієнт, h_k – крок Поляка (Аґмона-Моцкіна-Шонберга) в перетвореному просторі змінних $y = Ax$, $A = B^{-1}$.

Якщо матриця $B = I$, де I – одинична матриця, описаний метод збігається з класичним градієнтним методом з кроком Поляка (метод A):

$$x_{k+1} = x_k - h_k \frac{g_f(x_k)}{\left\| g_f(x_k) \right\|}, \quad h_k = \frac{m\left(f(x_k) - f^* \right)}{\left\| g_f(x_k) \right\|}. \tag{2}$$

Розглянемо результати роботи градієнтного методу з кроком Поляка та його модифікацій з використанням операції перетворення простору та параметра $m > 1$ для мінімізації квадратичних функцій багатьох змінних різної структури, які задаються матрицями з великим числом обумовленості.

Приклад 1 (метод A при $m = 1, 2$). Мінімізується квадратична функція $f_1(x) = (x-e)^T D(x-e)$, де $x, e \in R^n$, $n = 10\,000\,000$, $e = (1,...,1)^T$, D – діагональна матриця, діагональні елементи якої заповнюються так: $d_i = 1 + \alpha \cdot rand[0,1]$, $\alpha \in \{2.0; 1.0; 0.5; 0.1; 0.01\}$. Для функції $f_1(x)$ точка мінімуму $x^* = (1,...,1)^T$ та $f^* = 0$. Кількості ітерацій методу для різних відношень максимального $\lambda_{\max}$ та мінімального $\lambda_{\min}$ власних чисел матриці D наведено в таблиці 1.

Таблиця 1

№	$\lambda_{\max}$	$\lambda_{\min}$	$m = 1$		$m = 2$		q
			itr	$f\left(x_k^*\right)$	itr	$f\left(x_k^*\right)$	
1	3	1	45	9.6065e-21	42	2.7832e-21	1.07
2	2	1	45	6.0550e-21	27	3.5420e-21	1.67

3	1.5	1	45	4.5327e-21	19	2.1762e-21	2.37
4	1.1	1	45	3.4582e-21	11	1.7910e-22	4.09
5	1.01	1	45	3.2485e-21	7	1.7799e-23	6.43

Результати таблиці 1 показують, що використання параметра $m=2$ для квадратичної функції $f_1(x)$ дозволяє скоротити кількість ітерацій класичного методу більше ніж у 6 разів, про що свідчить показник $q = itr_{m=1}/itr_{m=2}$.

Приклад 2 (порівняння методів А та В при $m=1,2$). Мінімізується функція $f_2(x) = x^T Hx + b^T x + c$, де $x, b \in R^n$, $n = 1\,000$; $H = A^T A$ – симетрична матриця, що задається матрицею $A = \left\{a_{ij}\right\}_{i,j}^{k,n}$, елементи якої рівні $a_{ij} = 1 + \alpha \cdot rand\left[0,1\right]$, $\alpha \in \left\{0.1;1;10;50;200\right\}$, $k = 50\,000$; $x^* = rand\left[0,10\right]$. Вектор b та скаляр c обираються так, щоб $f^* = 0$. Кількості ітерацій методів А та В при $m=1,2$ для різних відношень $\lambda = \lambda_{\max}/\lambda_{\min}$ наведено в таблиці 2.

Таблиця 2

α	$\lambda_{\max}$	$\lambda_{\min}$	λ	Метод А		Метод В	
				$m=1$	$m=2$	$m=1$	$m=2$
0.1	1.1e-03	6.2e-10	1.8e+06	7399	14	20	5
1	2.3e-03	6.2e-08	3.7e+04	1892	18	21	6
10	3.6e-02	6.2e-06	5.8e+03	932	22	23	8
50	6.8e-01	1.5e-04	4.4e+03	879	26	25	9
200	1.0e+01	2.5e-03	4.1e+03	1449	28	27	10

З таблиці 2 бачимо, що кількість ітерацій методу В при $m=1,2$ є значно меншою порівняно з методом А при $m=1$. Детальні результати обчислювальних експериментів з методами А та В для різних типів гладких та негладких функцій представлено в роботі [3].

Висновки. Градієнтний метод з кроком Поляка можна значно прискорити за допомогою операції перетворення простору та використання параметра $m>1$ для мінімізації квадратичних яружних функцій багатьох змінних.

Робота підтримана грантом Volkswagen Foundation, № 97 775 (перший співавтор).

Список використаних джерел

1. Стецюк П.И. Ускорение субградиентного метода Поляка. Теорія оптимальних рішень. Ін-т кібернетики імені В.М. Глушкова НАН України. 2012. №11. С. 151 – 60.

2. Stetsyuk P., Stovba V., Chernousova Z. Subgradient Method with Polyak's Step in Transformed Space. Optimization and Applications. OPTIMA 2018. Communications in Computer and Information Science. 2019. Vol. 974. Springer. P. 49-63.

3. Стовба В.О. Субградієнтний метод з кроком Поляка у перетвореному просторі. Рукопис дисертації на здобуття наукового ступеня доктора філософії. Інститут кібернетики імені В.М. Глушкова НАН України. 2021. 184 с.