УДК 519.6

М. Я. Бартіш, Н. П. Огородник

ТРИКРОКОВИЙ АЛГОРИТМ МІНІМІЗАЦІЇ ФУНКЦІЙ

M. Ya. Bartish, N. Ph. Ogorodnyk. *Three step algorithm for function minimization*, Matematychni Studii, **29** (2008) 108–112.

A new approach for constructing algorithms for solving of minimization problem is proposed. This new algorithm is based on the gradient method and Newton's method. We proof a theorem where convergence of the proposed method is justified and rate of convergence is established. Numerical results are presented.

М. Я. Бартиш, Н. Ф. Огородник *Трёхшаговый алгоритм минимизации функций* // Математичні Студії. – 2008. – Т.29, №1. – С.108–112.

Построен метод решения задач минимизации функций многих переменных, базирующийся на методе Ньютона и градиентном методе. Иследована скорость сходимости данного метода. Проведены вычислительные эксперименты на различных типах функций. Сделано заключение о возможности применения предложенного метода.

1. Вступ. В літературі велика увага приділяється ітераційним методам розв'язування задач оптимізації ([1]–[5]). При цьому різні алгоритми ефективні для різних класів задач і не можна вказати найбільш ефективного алгоритму у загальному. При визначенні ефективності алгоритму враховується критерій точності результату, швидкість збіжності та складність алгоритму ([3]).

У цій статті запропоновано алгоритм розв'язування задач мінімізації функцій, який грунтується на використанні двох відомих методів — градієнтного та Ньютона. На кожному кроці за допомогою даних методів обчислюються два проміжні наближення до розв'язку, а наступне наближення алгоритму шукається як мінімум на прямій, що з'єднує отримані точки. Доведено, що швидкість збіжності алгоритму локально квадратична. Проведено тестування даного алгоритму на низці відомих задач, та зроблено його порівняння з градієнтним методом та методом Ньютона. Алгоритм демонструє свою ефективність в сенсі кількості обчислень як для функцій з "хорошими властивостями", так і для функцій, для яких метод Ньютона працює "погано".

2. Формулювання задачі. Розглянемо задачу мінімізації

$$f(x) \to \min, \qquad \text{де} \;\; x \in R^n, \quad f(x) \in C^2(R^n). \tag{1}$$

Відомими методами розв'язування задачі (1) є градієнтний метод (який має лінійну швидкість збіжності [2], [3], [5]):

$$x_{k+1} = x_k - \alpha_k f'(x_k), \tag{2}$$

2000 *Mathematics Subject Classification:* 90C52.

де параметр α_k, який задає довжину кроку, можна вибрати так ([2]):

$$\alpha_k = \arg\min_{\alpha>0} f(x_k - \alpha f'(x_k)) \tag{3}$$

або

$$f(x_{k+1}) - f(x_k) \le \varepsilon \alpha_k (f'(x_k), h_k), \tag{4}$$

та метод Ньютона (швидкість збіжності якого, за певних умов, є квадратичною [2]–[4]):

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k). \tag{5}$$

Перевагами методів (2) і (5) є довільне початкове наближення для градієнтного методу і висока швидкість збіжності в околі точки розв'язку для методу Ньютона. Враховуючи ці фактори, ми пропонуємо новий алгоритм розв'язку задачі (1). Маючи наближення x_k, ми виконуємо один крок за методом Ньютона

$$u_k = x_k - [f''(x_k)]^{-1} f'(x_k) \tag{6}$$

і один крок за градієнтним методом

$$v_k = x_k - \alpha_k f'(x_k). \tag{7}$$

Виконання одного кроку за градієнтним методом не вимагає істотних додаткових обчислень. Ми лише більш ефективно застосовуємо інформацію, яку отримали при обчисленні u_k, — використовується вже відоме значення $f'(x_k)$.

Маючи значення u_k і v_k, визначаємо наближення x_{k+1} за формулою

$$x_{k+1} = u_k - \beta_k (u_k - v_k), \tag{8}$$

де $f(u_k - \beta_k(u_k - v_k)) = \min\limits_{\beta} f(u_k - \beta(u_k - v_k))$.

Послідовність $\{x_k\}$, отримана за схемою (6)–(8), володіє кращими властивостями в сенсі швидкості збіжності, ніж послідовності отримані за схемами методів (2)–(3), (2)–(4) або (5), якщо їх застосовувати окремо. Водночас, обчислювальні затрати на кожній ітерації істотно не зростають у порівнянні з методом Ньютона.

3. Обґрунтування збіжності.

Теорема. *Нехай $D \subset \mathbb{R}^n$ — відкрита опукла область, а також:*
1. $f(x) \in C^2(\mathbb{R}^n)$ і для $x \in D$, $y \in \mathbb{R}^n$ сильно-опукла з константою опуклості $m > 0$ і

$$\|f''(x)\| \le M, \quad 0 < m \le M; \tag{9}$$

2. для $x, y \in D$ $f''(x)$ задовольняє умову Ліпшиця

$$\|f''(x) - f''(y)\| \le L\|x - y\|, \quad 0 < L < \infty; \tag{10}$$

3. початкове наближення x_0 вибрано так, що виконується умова

$$q = C(f(x_0) - f(x_*)) < 1, \ C = \frac{L^2 M^5}{2m^8}.$$

Тоді послідовність $\{x_k\}$, визначена алгоритмом (6)–(8), збігається до x_* — розв'язку задачі (1) і виконуються оцінки

$$f(x_{k+1}) - f(x_*) \le C^{-1} q^{2^{k+1}}, \quad \|x_{k+1} - x_*\| \le C_1 q^{2^k}, \quad C_1 = \sqrt{\tfrac{2}{mC}}\ .$$

Доведення. Існування і єдиність розв'язку задачі (1) випливають з умови сильної опуклості функції $f(x)$, при цьому маємо $\|f''(x)^{-1}\| \leq \frac{1}{m}$ ([2]). Нехай відоме деяке наближення x_k до розв'язку задачі мінімізації, тоді використовуючи розклад функції $f(x)$ у ряд Тейлора і умову (9), отримаємо

$$\begin{aligned} f(u_k) - f(x_*) &= (f'(x_*), u_k - x_*) + \frac{1}{2}(f''(\widetilde{x}_k)(u_k - x_*), u_k - x_*) = \\ &= \frac{1}{2}(f''(\widetilde{x}_k)(u_k - x_*), u_k - x_*) \leq \frac{M}{2}\|u_k - x_*\|^2 \end{aligned} \tag{11}$$

де $\widetilde{x}_k = u_k + \xi(x_* - u_k)$, $\xi \in (0,1)$.

Використовуючи властивість опуклих функцій можемо записати

$$(f'(u_k), u_k - x_*) = (f'(u_k) - f'(x_*), u_k - x_*) \geq m\|u_k - x_k\|^2.$$

Отже, $m\|u_k - x_k\|^2 \leq \|u_k - x_k\|\|f'(u_k)\|$. Звідки $\|u_k - x_*\| \leq \dfrac{1}{m}\|f'(u_k)\|$.

Враховуючи умову (10) і співвідношення $f'(u_k) = f'(u_k) - f'(x_k) - f''(x_k)(u_k - x_k)$, отримаємо

$$\begin{gathered} \|f'(u_k)\| \leq \|\textstyle\int_0^1 (f''(u_k + \tau(x_k - u_k)) - f''(x_k))d\tau\|\|(u_k - x_k)\| \leq \frac{L}{2}\|u_k - x_k\|^2 = \\ = \frac{L}{2}\|f''(x_k)^{-1}\|^2\|f'(x_k)\|^2 \leq \frac{L}{2m^2}\|f'(x_k)\|^2, \\ \|f'(x_k)\| = \|f'(x_k) - f'(x_*)\| = \|f''(\widetilde{x}_k)(x_k - x_*)\| \leq M\|x_k - x_*\|, \\ f(x_k) - f(x_*) = \frac{1}{2}(f''(\widetilde{x}_k)(x_k - x_*), x_k - x_*) \geq \frac{m}{2}\|x_k - x_*\|^2, \\ \|x_k - x_*\|^2 \leq \frac{2}{m}(f(x_k) - f(x_*)). \end{gathered}$$

Продовжимо оцінку (11). Враховуючи отримані оцінки, одержуємо

$$f(u_k) - f(x_*) \leq \frac{M}{2}(\frac{1}{m}\|f'(u_k)\|)^2 \leq \frac{M}{2}(\frac{1}{m}\frac{L}{2m^2}\|f'(x_k)\|^2)^2 = \frac{ML^2}{(2m^2)^3}\|f'(x_k)\|^4 \leq$$

$$\leq \frac{ML^2}{(2m^2)^3}(M\|x_k - x_*\|)^4 \leq \frac{M^5L^2}{(2m^2)^3}(\frac{2}{m})^2(f(x_k) - f(x_*))^2.$$

Отже,$f(u_k) - f(x_*) \leq C(f(x_k) - f(x_*))^2$, де $C = \dfrac{L^2M^5}{2m^2(m^2)^3}$.

Зрозуміло, що $f(x_{k+1}) - f(x_*) \leq f(u_k) - f(x_*) \leq C(f(x_k) - f(x_*))^2$.
Тоді
$C(f(x_{k+1}) - f(x_*)) \leq (C(f(x_k) - f(x_*)))^2 \leq (C(C(f(x_{k-1}) - f(x_*)))^2)^2 \leq \ldots \leq C^{-1}q^{2^{k+1}}$,
де $q = C(f(x_0) - f(x_*)) < 1$. Отже, $f(x_{k+1}) - f(x_*) \leq C^{-1}q^{2^{k+1}}$. Метод (6)–(8) збігається до розв'язку. Теорему доведено. □

Відзначимо, що вибір початкового наближення x_0, яке задовольняє умову 3, є доволі складною проблемою, тому доцільно використовувати алгоритм вигляду

$$\begin{gathered} u_k = x_k - \gamma_k[f''(x_k)]^{-1}f'(x_k),\ v_k = x_k - \alpha_k f'(x_k),\ x_{k+1} = u_k - \beta_k(u_k - v_k), \\ f(u_k - \beta_k(u_k - v_k)) = \min_{\beta} f(u_k - \beta(u_k - v_k)). \end{gathered} \tag{12}$$

За певних умов, квадратична збіжність в (12) буде досягатися локально.

4. Апробація методу. Нами розглянуто ряд прикладів і проведено порівняння алгоритму (12) з градієнтним методом та узагальненим методом Ньютона

$$x_{k+1} = x_k - \alpha_k[f''(x_k)]^{-1}f'(x_k), \tag{13}$$

які є базовими для алгоритму (12). Обчислення проводили до виконання умови $\|x_{k+1} - x_k\| \leq \varepsilon$.

1. Розширена функція Бейля ([5])

$$f(x) = \sum_{i=1}^{n/2}[(1.5 - x_{2i-1}(1 - x_{2i}^3))^2 + (2.25 - x_{2i-1}(1 - x_{2i}^2))^2 + (2.625 - x_{2i-1}(1 - x_{2i}^3))^2],$$

$$x^0 = (4, -0.5, \dots, 4, -0.5), \qquad x^0 = (9, -0.5, \dots, 9, -0.5), \qquad n = 2, 4, \dots$$

$$x^* = (2.125, 0 \dots, 2.125, 0), \qquad f(x^*) = 0.65625n/2.$$

2. Штрафна функція 2 ([5]) $f(x) = \sum_{i=1}^{n}(x_i - 1)^2 + 10^{-3}(\sum_{i=1}^{n} x_i^2 - 0.25)^2$,

$$x^0 = (1, 2, \dots, n), \qquad x^0 = (-10, -10, \dots, -10), \qquad n = 2, 3, \dots.$$

Точка розв'язку залежить від значення n.

3. Розширена функція Розенброка ([5]) $f(x) = \sum_{i=1}^{n/2}[100(x_{2i} - x_{2i-1}^2)^2 + (1 - x_{2i-1})^2]$,

$$x^0 = (-0.5, -0.5, \dots, -0.5, -0.5), \qquad x^0 = (-1, 1, \dots, -1, 1), \qquad n = 2, 4, \dots,$$

$$x^* = (1, 1, \dots 1), \qquad f(x^*) = 0.$$

4. $f(x) = 5x_1 + \frac{50000}{x_1} + 20x_2 + \frac{72000}{x_2} + 10x_3 + \frac{144000}{x_3} + 15x_4 + \frac{1500}{x_4}$,

$$x^0 = (1, 1, 1, 1), \qquad x^0 = (10, 10, 10, 10), \qquad x^* = (100, 60, 120, 10), \qquad f(x^*) = 6100.$$

У наступному наборі прикладів виконується умова $\det[f''(x^*)] = 0$, тобто гессіан в точці розв'язку є вироджений. Для таких функцій швидкість збіжності методу Ньютона при наближенні до точки розв'язку переходить у лінійну.

5. $f(x) = \sum_{i=1}^{n}(1 - x_i)^{2i}e^{(1-x_i)^2}$ ([6])

$$x^0 = (0.5, 0.5, \dots, 0.5), \qquad x^0 = (2, 2, \dots, 2), \qquad x^* = (1, 1, \dots, 1), \qquad f(x^*) = 0.$$

6. $f(x) = \sum_{i=1}^{n}(1 - x_i)^{2i}e^{(1-x_i)^{2i}}$ ([6])

$$x^0 = (0.5, 0.5, \dots, 0.5), \qquad x^0 = (2, 2, \dots, 2), \qquad x^* = (1, 1, \dots, 1), \qquad f(x^*) = 0.$$

7. $f(x) = \sum_{i=1}^{n}(\operatorname{ch} x_i - 1)^2 + \sum_{i=1}^{n} x_i^4$ ([6])

$$x^0 = (0.5, 0.5, \dots, 0.5), \qquad x^* = (0, 0, \dots, 0), \qquad f(x^*) = 0.$$

8. Розширена функція Міела і Кантрела ([5])

$$f(x) = \sum_{i=1}^{n/4}[(e^{x_{4i-3}} - x_{4i-2})^2 + 100(x_{4i-2} - x_{4i-1})^6 + tg^4(x_{4i-1} - x_{4i}) + x_{4i-3}^8]$$

$$x^0 = (1, 2, 1, 2 \dots, 1, 2, 1, 2), \qquad x^0 = (1, 0, 1, 0 \dots, 1, 0, 1, 0), \qquad n \in \{4, 8, 12, 16, \dots\}$$

$$x^* = (0, 1, 1, 1 \dots, 0, 1, 1, 1) \qquad f(x^*) = 0.$$

9. Штрафна функція 1 ([5])

$$f(x) = 10^{-5}\sum_{i=1}^{n}(x_i - 1)^2 + (\sum_{i=1}^{n} x_i^2 - 0.25)^2, \qquad x^0 = (1, 1, \dots), \qquad x^0 = (0.5, 0.5, \dots)$$

Точка розв'язку залежить від значення n.

В таблицях наведено кількість ітерацій затрачених для отримання наближеного розв'язку наведених задач. Знак '–' означає, що метод для даного випадку розбігається, або не збігався за 1000 ітерацій.

5. Висновки. Отже, на підставі проведених розрахунків та порівнянні отриманих результатів, бачимо, що трикроковий метод (12) за кількістю ітерацій переважає узагальнений метод Ньютона (13) та градієнтний метод (2), на базі яких він побудований. В таблицях 1, 2 представлені результати роботи методів для функцій з різними властивостями. З таблиці 2 видно, що метод (12) для функцій з виродженим гессіаном працює значно краще від методу Ньютона. У випадку, коли точність збільшується, при розв'язуванні прикладів 5–9, ефективність запропонованого алгоритму зростає.

Табл. 1: Результати обчислень для функцій 1–4 при $\varepsilon = 10^{-8}$

Функція	x_0	$n = 4$			$n = 50$		
		(2)	(13)	(12)	(2)	(13)	(12)
1	1	76	9	5	80	9	5
1	2	932	10	5	-	11	5
2	1	6	5	3	11	11	4
2	2	4	5	2	4	8	3
3	1	354	18	5	490	18	5
3	2	315	20	9	359	20	9
4	1	193	16	6			
4	2	95	10	4			

Табл. 2: Результати обчислень для функцій 5–9

Ф.	x_0	$n = 4$, $\varepsilon = 10^{-3}$		$n = 4$, $\varepsilon = 10^{-8}$		$n = 50$, $\varepsilon = 10^{-3}$		$n = 50$, $\varepsilon = 10^{-8}$	
		(13)	(12)	(13)	(12)	(13)	(12)	(13)	(12)
5	1	30	5	25	6	272	18	435	22
5	2	25	7	90	9	302	28	449	26
6	1	9	4	25	6	414	24	509	45
6	2	12	5	30	7	411	31	491	47
7	1	19	5	47	20	22	6	51	21
8	1	16	7	126	14	81	15	421	30
8	2	17	8	123	25	87	16	447	32
9	1	7	2	9	2	9	2	12	2
9	2	32	20	35	21	36	20	38	21

ЛІТЕРАТУРА

1. Базара М., Шетти К. Нелинейное программирование. Теория и алгоритмы. – М.: Мир, 1982.
2. Бартіш М.Я. Методи оптимізації. Теорія і алгоритми.– Л.: Видавничий центр ЛНУ ім. І. Франка, 2006.
3. Васильев Ф.П. Численные методы решения экстремальных задач. – М.: Наука, 1988.
4. Дэннис Дж., мл., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений. – М.: Мир, 1988.
5. Пшеничный Б.Н., Данилин Ю.М. Численные методы в экстремальных задачах.– М.: Наука, 1975.
6. Koko J. *A conjugate gradient method with quasi-Newton approximation* // Aplicationes mathematicae – 2000.– №27.– P. 153–165.

Львівський національний університет імені Івана Франка
факультет прикладної математики та інформатики
gut.natalochka@gmail.com

Надійшло 3.04.2007