УДК 519.6

М. Я. Бартіш, Н. П. Огородник

ПРО ОДИН ТРИКРОКОВИЙ МЕТОД МІНІМІЗАЦІЇ ФУНКЦІЙ

M. Ya. Bartish, N. Ph. Ogorodnyk. *On some three step method for functions minimization*, Mat. Stud. **34** (2010), 106–112.

The new method for solving unconstrained minimization problems is proposed. The considered three step method is based on well-known methods: method of linear interpolation proposed by Kurchatov and gradient method. The rate of convergence for new method is investigated. Numerical investigation is conducted on the test functions. The result of numerical experiments show that three step method is more effective in sense of amount of calculations. The efficiency of method is growing with increasing of function's dimension.

М. Я. Бартиш, Н. Ф. Огородник. *Об одном трёхшаговом методе минимизации функций* // Мат. Студії. – 2010. – Т.34, №1. – С.106–112.

Исследуется трёхшаговый метод решения задач безусловной минимизации функций многих переменных, который базируется на методе линейной интерполяции, предложенном Курчатовым, и градиентном методе. Иследовано скорость сходимости даного метода. Проведен численный эксперимент. Метод показал свою эффективность по количеству вычислений по сравнению с методом линейной интерполяции. Эффективность метода повышается с увеличением размерности функции.

1. Вступ. При моделюванні фізичних чи економічних процесів виникає потреба розв'язування задач оптимізації, зокрема, задач безумовної мінімізації. В літературі велика увага приділяється ітераційним методам розв'язування таких задач [1], [5]. Різні методи виявляють свою ефективність на різних класах задач, при цьому методи мають свої переваги і недоліки такі як вибір початкового наближення, швидкість збіжності, трудомісткість окремої ітерації, тощо. Зараз не існує універсального алгоритму, тому і надалі актуальною залишається проблема побудови ефективних методів.

В статті розглядаються ітераційно-різницеві методи розв'язування задачі безумовної мінімізації, зокрема трикроковий ітераційний метод, побудований на основі методу лінійної інтерполяції, який був запропонований Курчатовим В.А. [3], [4]. На кожній ітерації трикрокового методу обчислюємо два проміжні наближення – методом лінійної інтерполяції та градієнтним методом. У цьому разі градієнтний метод використовує, вже обчислені в методі лінійної інтерполяції, значення похідних, тому не потребує суттєвих додаткових обчислень. Наступне наближення методу шукаємо як точку мінімуму на прямій, що з'єднує проміжні наближення.

Метод показав свою ефективність у сенсі кількості обчислень як на функціях з "хорошими"властивостями, так і для таких, для яких базові методи працюють "погано".

2000 *Mathematics Subject Classification:* 2000 AMS Mathematics Subject Classification 90C52.

2. Формулювання задачі. Розглянемо задачу

$$f(x) \to \min, \tag{1}$$

де $x \in \mathbb{R}^n$ і $f(x) \in C(\mathbb{R}^n)$.

Для розв'язування задачі (1) можна використовувати метод Ньютона, а також його різницеві аналоги [1] вигляду

$$x_{k+1} = x_k - A_k^{-1} f'(x_k), \qquad k = 0, 1, \ldots, \tag{2}$$

де A_k — поділена різниця першого порядку для градієнта $f'(x)$ зі спеціально вибраними вузлами. Даний клас методів деяким чином наближає матрицю других похідних. Це робить ітерацію методу для певних типів задач менш трудомісткою, в сенсі кількості обчислень, у порівнянні з методами другого порядку.

Виберемо матрицю A_k наступним чином: $A_k = f'(x_{k-1}, 2x_k - x_{k-1})$. Тоді (2) буде методом лінійної інтерполяції [2]–[4]. На базі даного методу побудуємо трикроковий ітераційний метод розв'язування задачі (1)

$$u_k = x_k - (f'(x_{k-1}, 2x_k - x_{k-1}))^{-1} f'(x_k), \tag{3}$$

$$v_k = x_k - \alpha_k f'(x_k), \tag{4}$$

$$x_{k+1} = u_k - \gamma_k (u_k - v_k). \tag{5}$$

Параметр γ_k пропонуємо вибирати із умови

$$\gamma_k = \arg\min_{\gamma} f(u_k - \gamma(u_k - v_k)). \tag{6}$$

Послідовність x_k, отримана за схемою (3)–(6), володіє кращими властивостями, в сенсі швидкості збіжності, ніж послідовності, отримані за схемами методів (3) або (4), якщо їх застосовувати самостійно. Водночас, основні розрахунки відбуваються при визначенні u_k, тому обчислювальні затрати на кожній ітерації суттєво не зростають у порівнянні з методом лінійної інтерполяції (3). Запропонований метод дає можливість краще використовувати позитивні сторони кожного з базових методів.

При дослідженні методу (3)–(6) будемо використовувати наступні означення поділених різниць першого і другого порядку [1]

$$f'(x, y)(x - y) = f'(x) - f'(y), \tag{7}$$

$$f'(x, y, z)(x - y) = f'(x, z) - f'(y, z). \tag{8}$$

Для обчислення поділених різниць першого порядку використовуємо формулу з [1]

$$f'(x, y)_{i,j} = \frac{f'_{x_i}(x_1, ..., x_j, y_{j+1}, ..., y_n) - f'_{x_i}(x_1, ..., x_{j-1}, y_j, ..., y_n)}{x_j - y_j}.$$

3. Обґрунтування збіжності.

Теорема. *Нехай*

1) $f(x) \in C(D)$ — *сильно опукла з константою опуклості* m; $0 < m \le M_1$, $M_1 = const$, $D = \left\{ x \in \mathbb{R}^n, \|x - x_0\| \le 4\sqrt{\frac{2}{m}(f(x_0) - f(x_*))} \right\}$;

2) $f'(x)$ має поділені різниці першого і другого порядків в D, які задовольняють умови

$$\|f'(x,y_1)\| \le M_1,\ \|f'(x,y_1,z_1)\| \le M_2,$$
$$\|f'(x,y_1,z_1) - f'(x,y_2,z_2)\| \le M_3(\|y_1-y_2\| + \|z_1-z_2\|),$$

де $x, y_1, z_1, y_2, z_2 \in D$, $0 < M_2 < \infty$, $0 < M_3 < \infty$;

3) для $x, y \in D$ існує $(f'(x,y))^{-1}$, причому $\|(f'(x,y))^{-1}\| \le B$, де $B = const > 0$;

4) початкові наближення x_0, x_{-1} вибрано так, що виконуються умови

$$\frac{24M_3}{m}(f(x_{-1}) - f(x_*)) \le C\sqrt{f(x_0)-f(x_*)},\ \frac{24M_3}{m}\sqrt{f(x_0)-f(x_*)} \le C\sqrt{\mu},$$
$$\frac{M_1^3B^2}{m^3}\Big(M_2M_1B\sqrt{\frac{2}{m}} + C\Big)^2(f(x_0)-f(x_*)) = \mu < 1,$$

де $C = const > 0$.

Тоді послідовність $\{x_k\}$, визначена методом (3)–(6), збігається до розв'язку x_* задачі (1) і виконується оцінка

$$f(x_k) - f(x_*) \le \mu^{2^k-1}(f(x_0)-f(x_*)).$$

Доведення. Існування і єдиність розв'язку задачі (1) випливають з умови сильної опуклості функції $f(x)$ [5].

На підставі властивості опуклих функцій можемо записати

$$f(x_k) - f(x_*) = (f'(\tilde{x}_k), x_k - x_*) \ge \frac{m}{2}\|x_k - x_*\|^2,$$

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

$$\|x_k - x_*\|^2 \le \frac{2}{m}(f(x_k) - f(x_*)). \tag{9}$$

Припустимо, що наближення x_k, x_{k-1} до розв'язку задачі (1) знайдено. Покажемо, що послідовності $\{x_k\}$, $\{2x_k - x_{k-1}\}$ належить області D. Так як метод будує збіжну до розв'язку послідовність $\{f(x_k)\}$, то враховуючи (9) запишемо

$$\|x_k - x_0\| \le \|x_k - x_*\| + \|x_0 - x_*\| \le \sqrt{\frac{2}{m}(f(x_k)-f(x_*))} + \sqrt{\frac{2}{m}(f(x_0)-f(x_*))} \le$$
$$\le \sqrt{\frac{2}{m}(f(x_0)-f(x_*))} + \sqrt{\frac{2}{m}(f(x_0)-f(x_*))} \le 4\sqrt{\frac{2}{m}(f(x_0)-f(x_*))}.$$
$$\|2x_k - x_{k-1} - x_0\| \le 2\|x_k - x_*\| + \|x_{k-1} - x_*\| + \|x_0 - x_*\| \le$$
$$\le 2\sqrt{\frac{2}{m}(f(x_k)-f(x_*))} + \sqrt{\frac{2}{m}(f(x_{k-1})-f(x_*))} + \sqrt{\frac{2}{m}(f(x_0)-f(x_*))}$$
$$\le 4\sqrt{\frac{2}{m}(f(x_0)-f(x_*))}.$$

Оцінимо $\|f'(u_k)\|$, для цього враховуючи (3), (6) та (7), запишемо

$$f'(u_k) = f'(u_k) - f'(x_k) - f'(x_{k-1}, 2x_k - x_{k-1})(u_k - x_k) =$$
$$= (f'(u_k, x_k) - f'(x_{k-1}, 2x_k - x_{k-1}))(u_k - x_k) =$$
$$= (f'(u_k, x_k) + f'(u_k, x_{k-1}) - f'(u_k, x_{k-1}) - f'(x_{k-1}, 2x_k - x_{k-1}))(u_k - x_k) =$$
$$= (f'(u_k, x_k, x_{k-1})(x_k - x_{k-1}) + f'(u_k, x_{k-1}, 2x_k - x_{k-1})(u_k - 2x_k + x_{k-1}))(u_k - x_k) =$$
$$= f'(u_k, x_{k-1}, 2x_k - x_{k-1})(u_k - x_k)^2 +$$
$$+ (f'(u_k, x_k, x_{k-1}) - f'(u_k, x_{k-1}, 2x_k - x_{k-1}))(x_k - x_{k-1})(u_k - x_k).$$

З (3) і умови 3 отримаємо оцінку $\|u_k - x_k\| \le B\|f'(x_k)\|$. Тоді, враховуючи умову 2

$$\|f'(u_k)\| \le M_2 B^2 \|f'(x_k)\|^2 + M_3 B \|f'(x_k)\| (\|x_k - x_{k-1}\| + \tag{10}$$
$$+ \|x_{k-1} - 2x_k + x_{k-1}\|)\|x_k - x_{k-1}\| = M_2 B^2 \|f'(x_k)\|^2 + 3M_3 B \|f'(x_k)\| \|x_k - x_{k-1}\|^2.$$

З умови 2 маємо $\|f'(x_k)\| \le M_1 \|x_k - x_*\|$. Враховуючи (9), можемо записати

$$\|x_k - x_{k-1}\| \le \|x_k - x_*\| + \|x_{k-1} - x_*\| \le$$
$$\le \sqrt{\frac{2}{m}(f(x_k) - f(x_*))} + \sqrt{\frac{2}{m}(f(x_{k-1}) - f(x_*))} \le 2\sqrt{\frac{2}{m}(f(x_{k-1}) - f(x_*))}.$$

Продовжимо оцінку (10), враховуючи наведені оцінки

$$\|f'(u_k)\| \le M_2 B^2 M_1^2 \|x_k - x_*\|^2 + 3M_3 B M_1 \|x_k - x_*\| \frac{8}{m}(f(x_{k-1} - f(x_*)) \le$$
$$\le \Big(M_2 B M_1 \sqrt{\frac{2}{m}(f(x_k) - f(x_*))} + \frac{24M_3}{m}(f(x_{k-1}) - f(x_*))\Big) M_1 B \sqrt{\frac{2}{m}(f(x_k) - f(x_*))}.$$

Використавши умову 2 і розклад функції $f(x)$ у ряд Тейлора, отримаємо

$$f(u_k) - f(x_*) = (f'(\tilde{x}_k), u_k - x_*) = (f'(\tilde{x}_k, x_*), u_k - x_*) \le \frac{M_1}{2}\|u_k - x_*\|^2, \tag{11}$$

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

На підставі властивості опуклих функцій [5], ми можемо записати $(f'(u_k), u_k - x_*) = (f'(u_k) - f'(x_*), u_k - x_*) \ge m\|u_k - x_*\|^2$. Отже,

$$\|u_k - x_*\| \le \frac{1}{m}\|f'(u_k)\|.$$

Продовжимо (11), враховуючи вище наведені оцінки

$$f(u_k) - f(x_*) \le \frac{M_1}{2m^2}\|f'(u_k)\|^2 \le \frac{M_1}{2m^2}\Big(M_2 B M_1 \sqrt{\frac{2}{m}(f(x_k) - f(x_*))} +$$
$$+ \frac{24M_3}{m}(f(x_{k-1}) - f(x_*))\Big)^2 M_1^2 B^2 \frac{2}{m}(f(x_k) - f(x_*)).$$

З (6) зрозуміло, що $f(x_{k+1}) - f(x_*) \le f(u_k) - f(x_*)$. Покажемо, що враховуючи умову 4, виконується $f(x_1) - f(x_*) \le \mu(f(x_0) - f(x_*))$

$$f(x_1) - f(x_*) \le f(u_0) - f(x_*) \le \frac{M_1}{2m^2}\Big(M_2 B M_1 \sqrt{\frac{2}{m}(f(x_0) - f(x_*))} +$$
$$\frac{24M_3}{m}(f(x_{-1}) - f(x_*))\Big)^2 M_1^2 B^2 \frac{2}{m}(f(x_0) - f(x_*)) \le$$
$$\le \frac{M_1^3 B^2}{m^3}\Big(M_2 M_1 B \sqrt{\frac{2}{m}} + C\Big)^2 (f(x_0) - f(x_*))^2 = \mu(f(x_0) - f(x_*)).$$

Припустимо, що для деякого k виконується $f(x_k) - f(x_*) \leq \mu^{2^k-1}(f(x_0) - f(x_*))$. Покажемо, що оцінка справедлива для $k+1$

$$f(x_{k+1}) - f(x_*) \leq f(u_k) - f(x_*) \leq$$
$$\leq \frac{M_1}{2m^2}(M_2 B M_1 \sqrt{\frac{2}{m}} \mu^{\frac{2^k-1}{2}} \sqrt{f(x_0) - f(x_*)} + \frac{24M_3}{m} \mu^{2^{k-1}-1}(f(x_0) - f(x_*)))^2 \times$$
$$\times M_1^2 B^2 \frac{2}{m} \mu^{2^k-1}(f(x_0) - f(x_*)) \leq \frac{M_1^3 B^2}{m^3}(M_2 M_1 B \sqrt{\frac{2}{m}} + C)^2 \mu^{2^k-1} \mu^{2^k-1}(f(x_0) - f(x_*))^2 =$$
$$= \mu^{2^{k+1}-1}(f(x_0) - f(x_*)).$$

Отже, за методом математичної індукції маємо, що послідовність наближень, побудованих за формулами (3)–(6), збігається до x_*. Теорему доведено. □

Відзначимо, що вибір початкових наближень x_0, x_{-1}, які б задовольняли умову 4, є досить складною проблемою, тому на практиці доцільно використовувати метод вигляду

$$u_k = x_k - \beta_k(f'(x_{k-1}, 2x_k - x_{k-1}))^{-1} f'(x_k), \tag{12}$$

$$v_k = x_k - \alpha_k f'(x_k), \tag{13}$$

$$x_{k+1} = u_k - \gamma_k(u_k - v_k), \tag{14}$$

$$\gamma_k = \arg\min_{\gamma} f(u_k - \gamma(u_k - v_k)), \tag{15}$$

де параметри α_k, β_k повинні забезпечувати монотонне спадання функції.

В методі (12)–(15), при виконанні певних умов, квадратична збіжність буде досягатися локально.

4. Апробація методу. Нами розглянуто ряд прикладів і проведено порівняння методу (12)–(15) з методом лінійної інтерполяції (12).

Обчислення проводилися до виконання умови $\|x_{k+1} - x_k\| \leq \varepsilon$, $\varepsilon = 10^{-8}$. Додаткові наближення будували за правилом $x_{-1} = x_0 + 0.01x_0$. У таблиці 1 наведено кількість ітерацій — N та K — кількість обчислень, еквівалентних кількості обчислень значення функції $f(x)$, які були затрачені для отримання наближеного розв'язку наведених задач. Тестові задачі взяті із статті [6].

1. Штрафна функція 2

$$f(x) = \sum_{i=1}^{n}(x_i - 1)^2 + 10^{-3}\Big(\sum_{i=1}^{n} x_i^2 - 0.25\Big)^2;$$
$$x^0 = (20, 20, \ldots, 20); \qquad x^0 = (-10, -10, \ldots, -10); \qquad n = 2, 3, \ldots.$$

Точка розв'язку залежить від значення n.

2. Розширена функція Розенброка

$$f(x) = \sum_{i=1}^{n/2}[100(x_{2i} - x_{2i-1}^2)^2 + (1 - x_{2i-1})^2];$$
$$x^0 = (-1.2, 1, -1.2, 1 \ldots, -1.2, 1); \qquad x^0 = (-1, 0.8, -1, 0.8 \ldots, -1, 0.8); \qquad n = 2, 4, \ldots;$$
$$x^* = (1, 1, \ldots 1); \qquad f(x^*) = 0.$$

3. Розширена функція Пауела

$$f(x) = \sum_{i=1}^{n/4} [(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4];$$
$$x^0 = (-1, 1, -1, 1 \ldots, -1, 1); \qquad x^0 = (1, 1 \ldots 1); \qquad n = 4, 8, 12, 16, \ldots;$$
$$x^* = (0, 0, \ldots, 0); \qquad f(x^*) = 0.$$

4. Штрафна функція 1

$$f(x) = 10^{-5} \sum_{i=1}^{n} (x_i - 1)^2 + \Big(\sum_{i=1}^{n} x_i^2 - 0.25\Big)^2;$$
$$x^0 = (0.5, 0.5, \ldots, 0.5); \qquad x^0 = (-0.5, -0.5, \ldots, -0,5); \qquad n = 2, 3, \ldots.$$

Розв'язок залежить від значення n.

ф-ція	x_0	$n = 10$				$n = 50$			
		(12)		(12)-(15)		(12)		(12)-(15)	
		N	K	N	K	N	K	N	K
1	1	10	2111	3	830	11	28650	4	10450
1	2	9	1900	3	846	10	26050	3	7850
2	1	21	2562	10	1648	21	54682	11	29114
2	2	20	2485	7	1338	20	52125	8	21424
3	1	52	15654	16	3979	54	151684	17	49236
3	2	50	15051	16	3836	51	143260	16	46109
4	1	9	1901	3	798	11	28650	2	5250
4	2	9	1900	3	820	11	28650	2	5250

5. Висновки. В даній статті запропоновано новий метод розв'язування задач безумовної мінімізації функцій багатьох змінних. Проведено теоретичні та числові дослідження даного методу. В запропонованому трикроковому методі (3)–(6) використовуються переваги базових методів. Доведено квадратичну збіжність методу (3)–(6).

На підставі проведених числових розрахунків та порівнянні отриманих результатів бачимо, що метод (3)–(6) за кількістю обчислень та ітерацій переважає метод лінійної інтерполяції (3). Числові дослідження проведено на різних типах функцій. Відзначимо, що запропонований метод досить ефективно працює у випадку функцій з "поганими"властивостями, такими як функції типу яру (функція Розенброка); та функції з виродженим в точці розв'язку гессіаном (функція Пауелла). При наближенні до точки розв'язку збіжність методу, запропонованого Курчатовим, стає дуже повільною, а запропонований трикроковий метод не втрачає швидкості збіжності. Крім того зі збільшенням розмірності простору — n, ефективність запропонованого методу, в сенсі кількості обчислень, зростає.

ЛІТЕРАТУРА

1. Васильев Ф.П. Численные методы решения экстремальных задач. – М.: Наука, 1988.
2. Гнатишин О. П., Шахно С.М. *Про деякі ітераційно-різницеві методи розв'язування задач безумовної мінімізації*// Вісник Львів. ун-ту. Сер. прикл. матем. та інформ. – 2008. – №.6. – С.28–35.
3. Курчатов В.А. *Об условиях примененния одного метода линейной интерполяции для решения функциональных уравнений*// Известия высших учебных заведений Математика. – 1975. – №8. – С.55–63.
4. Курчатов В.А. *Об одном методе линейной интерполяции решения функциональных уравнений*// Докл. АН СССР. – 1971. – Т.198. – №3. – С.524–526.
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

Надійшло 23.12.2009
Після переробки 03.11.10