

УДК 519.21:519.61

ПРО ДЕЯКІ ІТЕРАЦІЙНО-РІЗНИЦЕВІ МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ БЕЗУМОВНОЇ МІНІМІЗАЦІЇ

О.Гнатишин, С.Шахно

Львівський національний університет імені Івана Франка
вул. Університетська, 1, м. Львів, 79000, e-mail: ktop@franko.lviv.ua

Досліджено деякі ітераційно-різницеві методи розв'язування задач безумовної мінімізації, які не потребують обчислення матриці других похідних. Доведено теореми, які обґрунтовують збіжність, та визначено швидкість збіжності запропонованих методів. Наведено результати числового експерименту.

Ключові слова: безумовна мінімізація, ітераційно-різницеві методи, поділені різниці, швидкість збіжності.

1. ВСТУП

Математичне моделювання складних фізичних процесів дуже часто потребує розв'язування задач оптимізації. Універсальних методів для успішного розв'язування широкого класу таких задач нема, тому актуальною є проблема побудови нових ефективних методів розв'язування задач на екстремум. Ми пропонуємо ітераційно-різницеві методи, які за швидкістю збіжності близькі до методу Ньютона, однак не потребують обчислення матриці других похідних.

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Розглянемо задачу мінімізації

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (1)$$

де $f(x)$ – двічі неперервно диференційовна функція. Відомим методом розв'язування задачі (1) є метод Ньютона. Однак цей метод на кожній ітерації потребує обчислення матриці других похідних. Тому постає питання про побудову методів мінімізації, які не потребують обчислення матриці других похідних. Такими методами є методи типу Стеффенсена [1,4], у яких матриця других похідних замінена матрицею поділених різниць першого порядку. Недоліком цих методів є необхідність виконання мінімізації за градієнтним методом, що потребує додаткових обчислень.

Ми пропонуємо ітераційно-різницеві методи, які не мають цих недоліків:

$$x_{n+1} = x_n - A_n^{-1} f'(x_n), \quad n = 0, 1, \dots \quad (2)$$

де A_n – поділена різниця першого порядку для градієнта зі спеціально вибраними вузлами або лінійна комбінація таких різниць. Зокрема, досліджено такі випадки:

а) $A_n = f'(x_n, x_{n-1})$, б) $A_n = f'(2x_n - x_{n-1}, x_{n-1})$,

в) $A_n = f'(x_n, x_{n-1}) + f'(x_n, x_{n-2}) - f'(x_{n-1}, x_{n-2})$, де $f'(u, v)$ – поділена різниця першого порядку градієнта $f'(x)$, причому $f'(x, y)(x - y) = f(x) - f(y)$, $f'(x, x) = f''(x)$. Далі будемо також використовувати позначення $f'(x, y, z)$ для

поділених різниць другого порядку. Ці різниці задовольняють умови [3]: $f'(x, y, z)(y - z) = f'(x, y) - f'(x, z)$, $f'(x, y, z)(x - y) = f'(x, z) - f'(y, z)$. Поділені різниці першого порядку можна обчислити, наприклад, за формулою [1]:

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

З єдиної точки зору досліджено збіжність запропонованих модифікацій та визначено порядок збіжності. Оскільки ці методи прості в реалізації і потребують обчислення тільки перших похідних, то вони мають широке практичне застосування.

Відзначимо, що в праці [7] розглянуто методи для розв'язування нелінійної задачі найменших квадратів, які використовують поділені різниці функцій типу а) та б), а в праці [6] досліджено метод для розв'язування систем нелінійних рівнянь, який використовує комбінацію поділених різниць функцій типу в).

3. ОБГРУНТУВАННЯ ЗБІЖНОСТІ

Достатні умови і швидкість збіжності запропонованих модифікацій визначені в таких теоремах.

Теорема 1. Нехай $f(x)$ двічі неперервно-диференційовна функція в області $D \subseteq R^n$. Припустимо, що:

- задача (1) має розв'язок x_* у деякій області $\Omega(x_*, r_*) = \{x : \|x - x_*\| < r_*\} \subset D$, $r_* = 1/(3p)$;
- існує $f''(x_*)^{-1}$;
- $f'(x)$ має поділені різниці першого порядку, які задовольняють умови типу Ліпшица: $\|f''(x_*)^{-1}(f'(x, y) - f'(u, v))\| \leq p(\|x - u\| + \|y - v\|)$, $p > 0$.

Тоді ітераційний процес

$$x_{n+1} = x_n - A_n^{-1} f'(x_n), \quad A_n = f'(x_n, x_{n-1}), \quad (n = 0, 1, \dots, \quad x_{-1}, x_0 \in \Omega) \quad (3)$$

коректно визначений, генерує послідовність $\{x_n\}$, яка належить відкритій області Ω , і збігається до розв'язку x_* зі швидкістю, яку характеризує нерівність

$$\|x_{n+1} - x_*\| \leq \frac{p(\|x_n - x_*\| \|x_{n-1} - x_*\|)}{1 - p(\|x_n - x_*\| + \|x_{n-1} - x_*\|)}. \quad (4)$$

Доведення. Відомо [2], що для довільного оператора $G \in R^{n \times n}$, такого що $\|G\| < 1$, існує $(I - G)^{-1}$ і $\|(I - G)^{-1}\| \leq \frac{1}{1 - \|G\|}$.

Прийmemo $G = I - f''(x_*)^{-1} A_n$. Тоді $\|(I - G)^{-1}\| = \|A_n^{-1} f''(x_*)\|$.

Оцінимо $\|G\|$:

$$\begin{aligned} \|I - f''(x_*)^{-1} A_n\| &= \|f''(x_*)^{-1}(f''(x_*) - A_n)\| = \|f''(x_*)^{-1}(f'(x_*, x_*) - f'(x_n, x_{n-1}))\| \leq \\ &p(\|x_n - x_*\| + \|x_{n-1} - x_*\|) \leq 2pr_* < 1. \end{aligned}$$

$$\text{Тоді } \|A_n^{-1}f''(x_n)\| \leq \frac{1}{1-p(\|x_n - x_*\| + \|x_{n-1} - x_*\|)}. \quad (5)$$

Оскільки

$$(A_n^{-1}f'(x_n), x_{n+1} - x_*) = (A_n^{-1}(f'(x_n) - f'(x_*)), x_{n+1} - x_*) = (A_n^{-1}(f'(x_n, x_*)(x_n - x_*)), x_{n+1} - x_*),$$

то

$$\begin{aligned} \|x_{n+1} - x_*\|^2 &= (x_{n+1} - x_*, x_{n+1} - x_*) = (x_n - A_n^{-1}f'(x_n) - x_*, x_{n+1} - x_*) = \\ &= (x_n - x_* - A_n^{-1}f'(x_n, x_*)(x_n - x_*), x_{n+1} - x_*) = \\ &= ((I - A_n^{-1}f'(x_n, x_*))(x_n - x_*), x_{n+1} - x_*) \leq \|I - A_n^{-1}f'(x_n, x_*)\| \|x_n - x_*\| \|x_{n+1} - x_*\|. \end{aligned}$$

Звідси

$$\begin{aligned} \|x_{n+1} - x_*\| &\leq \|I - A_n^{-1}f'(x_n, x_*)\| \|x_n - x_*\| = \|A_n^{-1}f''(x_*)f''(x_*)^{-1}(A_n - f'(x_n, x_*))\| \|x_n - x_*\| \leq \\ &\leq \|A_n^{-1}f''(x_*)\| \|f''(x_*)^{-1}(f'(x_n, x_*) - A_n)\| \|x_n - x_*\|. \end{aligned}$$

Таким чином, ми отримали оцінку

$$\|x_{n+1} - x_*\| \leq \|A_n^{-1}f''(x_*)\| \|f''(x_*)^{-1}(f'(x_n, x_*) - A_n)\| \|x_n - x_*\| \quad (6)$$

З урахуванням (5) та оцінки

$$\|f''(x_*)^{-1}(f'(x_n, x_*) - A_n)\| = \|f''(x_*)^{-1}(f'(x_n, x_*) - f'(x_n, x_{n-1}))\| \leq p \|x_{n-1} - x_*\| \leq pr_* < 1$$

остаточно отримаємо

$$\|x_{n+1} - x_*\| \leq \frac{p(\|x_n - x_*\| \|x_{n-1} - x_*\|)}{1 - p(\|x_n - x_*\| + \|x_{n-1} - x_*\|)}.$$

Теорему доведено.

Наслідок. Нерівність (4) дає змогу зробити висновок про збіжність ітераційного процесу (2) з вибором $A_n = F(x_n, x_{n-1})$ зі швидкістю $\frac{1+\sqrt{5}}{2} \approx 1.618$.

Теорема 2. Нехай $f(x)$ двічі неперервно-диференційовна функція в області $D \subseteq R^n$. Припустимо, що:

- задача (1) має розв'язок x_* у деякій області $\Omega(x_*, r_*) = \{x : \|x - x_*\| < r_*\} \subset D$,

$$r_* = \frac{2}{5p + \sqrt{25p^2 + 16q}};$$

- існує $f''(x_*)^{-1}$;
- $f'(x)$ має поділені різниці першого та другого порядку, які задовольняють умови типу Ліпшиця:

$$\|f''(x_*)^{-1}(f'(x, y) - f'(u, v))\| \leq p(\|x - u\| + \|y - v\|), \quad p > 0;$$

$$\|f''(x_*)^{-1}(f'(u, x, y) - f'(v, x, y))\| \leq q\|u - v\|, \quad q > 0.$$

Тоді ітераційний процес

$$x_{n+1} = x_n - A_n^{-1}f'(x_n), \quad A_n = f'(2x_n - x_{n-1}, x_{n-1}), \quad (n = 0, 1, \dots, x_{-1}, x_0 \in \Omega) \quad (7)$$

коректно визначений, генерує послідовність $\{x_n\}$, яка належить відкритій області Ω , і збігається до розв'язку x_* зі швидкістю, яку характеризує нерівність

$$\|x_{n+1} - x_*\| \leq \frac{p\|x_n - x_*\| + 4q\left(\|x_{n-1} - x_*\|^2\right)}{1 - 2p\left(\|x_n - x_*\| + \|x_{n-1} - x_*\|\right)}\|x_n - x_*\| \quad (8)$$

Доведення. Схема доведення така ж, як попередньої теореми. Скористаємось отриманою вище нерівністю (6) і оцінимо кожен зі співмножників цієї нерівності:

$$\begin{aligned} \|I - f''(x_*)^{-1} A_n\| &= \|f''(x_*)^{-1}(f''(x_*) - A_n)\| = \|f''(x_*)^{-1}(f''(x_*, x_*) - f'(2x_n - x_{n-1}, x_{n-1}))\| \leq \\ &\leq p(\|x_* - 2x_n + x_{n-1}\| + \|x_{n-1} - x_*\|) \leq p(\|x_* - x_n\| + \|x_n - x_{n-1}\| + \|x_{n-1} - x_*\|) \leq \\ &\leq 2p(\|x_* - x_n\| + \|x_* - x_{n-1}\|) \leq 4pr_* < 1. \end{aligned}$$

Звідси

$$\|A_n^{-1} f''(x_n)\| \leq \frac{1}{1 - 2p(\|x_n - x_*\| + \|x_{n-1} - x_*\|)}.$$

Для другого співмножника справедлива оцінка :

$$\begin{aligned} \|f''(x_*)^{-1}(f'(x_n, x_*) - A_n)\| &= \|f''(x_*)^{-1}(f'(x_n, x_*) - f'(2x_n - x_{n-1}, x_{n-1}))\| = \\ &= \|f''(x_*)^{-1}(f'(x_n, x_*) - f'(x_n, x_n) + f'(x_n, x_n) - f'(2x_n - x_{n-1}, x_{n-1}))\| \leq \\ &\leq p\|x_n - x_*\| + \|f''(x_*)^{-1}(f'(x_n, x_n) - f'(2x_n - x_{n-1}, x_{n-1}))\| = \\ &= p\|x_n - x_*\| + \|f''(x_*)^{-1}(f'(2x_n - x_{n-1}, x_{n-1}) - f'(x_n, x_{n-1}) + f'(x_n, x_{n-1}) - f'(x_n, x_n))\| = \\ &= p\|x_n - x_*\| + \|f''(x_*)^{-1}(f'(2x_n - x_{n-1}, x_{n-1})(x_n - x_{n-1}) - f'(x_n, x_{n-1}, x_n)(x_n - x_{n-1}))\| \leq \\ &\leq p\|x_n - x_*\| + q\|x_n - x_{n-1}\|\|x_n - x_{n-1}\| \leq p\|x_n - x_*\| + q\left(\|x_n - x_*\| + \|x_{n-1} - x_*\|\right) \leq \\ &\leq p\|x_n - x_*\| + 4q\left(\|x_{n-1} - x_*\|^2\right) \end{aligned}$$

Таким чином, ми отримали:

$$\|x_{n+1} - x_*\| \leq \frac{p\|x_n - x_*\| + 4q\left(\|x_{n-1} - x_*\|^2\right)}{1 - 2p\left(\|x_n - x_*\| + \|x_{n-1} - x_*\|\right)}\|x_n - x_*\|,$$

що і треба було показати.

З (8) і визначення r_* випливає, що процес (7) коректно визначений, генерує послідовність $\{x_n\}$, яка належить відкритій області Ω , і $\lim_{n \rightarrow \infty} \|x_n - x_*\| = 0$.

Теорему доведено.

Наслідок. Збіжність ітераційного процесу (7) є квадратичною.

Доведення. З (8) випливає, що швидкість збіжності ітераційного процесу не вища 2, а значить

$$c\|x_n - x_*\| \geq \|x_{n-1} - x_*\|^2, \quad c = \text{const} > 0.$$

Тоді (8) матиме вигляд

$$\|x_{n+1} - x_*\| \leq \frac{(p + 4cq)\|x_n - x_*\|^2}{1 - 2p(\|x_n - x_*\| + \|x_{n-1} - x_*\|)},$$

що і доводить твердження.

Теорема 3. Нехай $f(x)$ двічі неперервно-диференційовна функція в $D \subseteq R^n$. Припустимо, що:

- задача (1) має розв'язок x_* в деякій області $\Omega(x_*, r_*) = \{x : \|x - x_*\| < r_*\} \subset D$,
 $r_* = 2 / \left(3p + \sqrt{9p^2 + 16q} \right)$
- існує $f''(x_*)^{-1}$;
- $f'(x)$ має поділені різниці першого і другого порядку, які задовольняють умовам типу Ліпшиця:

$$\|f''(x_*)^{-1}(f'(x, y) - f'(u, v))\| \leq p(\|x - u\| + \|y - v\|), \quad p > 0;$$

$$\|f''(x_*)^{-1}(f'(u, x, y) - f'(v, x, y))\| \leq q\|u - v\|, \quad q > 0.$$

Тоді ітераційний процес

$$x_{n+1} = x_n - A_n^{-1} f'(x_n), \quad A_n = f'(x_n, x_{n-1}) + f'(x_n, x_{n-2}) - f'(x_{n-1}, x_{n-2}),$$

$$(n = 0, 1, \dots, \quad x_{-2}, x_{-1}, x_0 \in \Omega), \quad (9)$$

коректно визначений, генерує послідовність $\{x_n\}$, яка належить відкритій області Ω , і збігається до розв'язку x_* зі швидкістю, яку характеризує нерівність

$$\|x_{n+1} - x_*\| \leq \frac{p\|x_n - x_*\| + q(\|x_n - x_*\| + \|x_{n-2} - x_*\|)\|x_{n-1} - x_*\|}{1 - 2p\|x_n - x_*\| - q(\|x_n - x_*\| + \|x_{n-2} - x_*\|)\|x_{n-1} - x_*\|} \|x_n - x_*\|. \quad (10)$$

Доведення. По аналогії до теореми 1 оцінимо кожен зі співмножників нерівності (6). Ми отримаємо:

$$\begin{aligned} & \|I - f''(x_*)^{-1} A_n\| = \|f''(x_*)^{-1}(f''(x_*) - A_n)\| = \\ & = \|f''(x_*)^{-1}(f'(x_*, x_*) - f'(x_n, x_{n-1}) - f'(x_{n-2}, x_n) + f'(x_{n-2}, x_{n-1}))\| = \\ & = \|f''(x_*)^{-1}(f'(x_*, x_*) - f'(x_n, x_*) + f'(x_{n-2}, x_*) - f'(x_{n-2}, x_n) + f'(x_n, x_*) - \\ & \quad - f'(x_n, x_{n-1}) - f'(x_{n-2}, x_*) - f'(x_{n-2}, x_{n-1}))\| = \\ & = \|f''(x_*)^{-1} \{ (f'(x_*, x_*) - f'(x_n, x_*)) + (f'(x_{n-2}, x_*) - f'(x_{n-2}, x_n)) + \\ & \quad + (f'(x_n, x_*, x_{n-1}) - f'(x_{n-2}, x_*, x_{n-1}))(x_* - x_{n-1}) \}\| \leq \\ & \leq 2p\|x_* - x_n\| + q\|x_n - x_{n-2}\|\|x_* - x_{n-1}\| \leq \\ & \leq 2p\|x_* - x_n\| + q(\|x_n - x_*\| + \|x_{n-2} - x_*\|)\|x_* - x_{n-1}\| \leq 2pr_* + 2qr_*^2 < 1. \end{aligned}$$

Звідси,

$$\|A_n^{-1} f''(x_n)\| \leq \frac{1}{1 - 2p\|x_* - x_n\| - q(\|x_n - x_*\| + \|x_{n-2} - x_*\|)\|x_* - x_{n-1}\|}.$$

Відповідно оцінка другого співмножника в (6):

$$\begin{aligned} & \|f''(x_*)^{-1}(f'(x_n, x_*) - A_n)\| = \\ & = \|f''(x_*)^{-1}(f'(x_n, x_*) - f'(x_n, x_n) + f'(x_n, x_n) - f'(x_n, x_{n-1}) - f'(x_{n-2}, x_n) + f'(x_{n-2}, x_{n-1}))\| = \\ & = \|f''(x_*)^{-1}((f'(x_n, x_*) - f'(x_n, x_n)) + f'(x_n, x_n, x_{n-1})(x_n - x_{n-1}) - f'(x_{n-2}, x_n, x_{n-1})(x_n - x_{n-1}))\| \leq \\ & \leq p\|x_n - x_*\| + q\|x_n - x_{n-2}\|\|x_n - x_{n-1}\| \end{aligned}$$

Остаточо отримаємо

$$\|x_{n+1} - x_*\| \leq \frac{p\|x_n - x_*\| + q(\|x_n - x_*\| + \|x_{n-2} - x_*\|)\|x_{n-1} - x_*\|}{1 - 2p\|x_n - x_*\| - q(\|x_n - x_*\| + \|x_{n-2} - x_*\|)\|x_{n-1} - x_*\|} \|x_n - x_*\|,$$

що і треба було показати.

З (10) і визначення r_* випливає, що процес (9) коректно визначений, генерує послідовність $\{x_n\}$, яка належить відкритій області Ω , і $\lim_{n \rightarrow \infty} \|x_n - x_*\| = 0$.

Теорему доведено.

Наслідок. На підставі нерівності (10) можна стверджувати, що порядок збіжності ітераційної процедури (8) дорівнює 1.839... . Це значення ми отримуємо як єдиний дійсний розв'язок рівняння $t^3 - t^2 - t - 1 = 0$.

4. РЕЗУЛЬТАТИ ЧИСЛОВОГО ЕКСПЕРИМЕНТУ

Для порівняння швидкості збіжності запропонованих методів та методу Ньютона було проведено чисельний експеримент. Обчислення проводили до виконання умови $\|f'(x_n)\| \leq 10^{-8}$. Додаткові початкові наближення будували за правилом $x_{-1} = x_0 + 0.01 * x_0$, $x_{-2} = x_0 - 0.01 * x_0$. В таблиці наведено кількість ітерацій, затрачених для отримання наближеного розв'язку наведених нижче задач.

Номер прикладу	Метод			
	Ньютона	(3)	(7)	(9)
1	16	24	20	17
2	6	11	9	9
3	15	23	18	16
4	7	11	8	8
5	7	10	9	8

З таблиці видно, що запропоновані нами методи (7) та (9) за кількістю ітерацій мало поступаються методу Ньютона, проте вони не потребують обчислення матриці других похідних. Разом з тим метод (9) зі швидкістю збіжності 1.839... має деякі переваги перед методом (7) зі швидкістю збіжності 2. Це можна пояснити тим, що метод (9) фактично використовує у своїй ітераційній формулі три члени інтерполяційного многочлена Ньютона.

Приклад 1:

$$f(x) = 5x_1 + \frac{50000}{x_1} + 20x_2 + \frac{72000}{x_2} + 10x_3 + \frac{144000}{x_3};$$

$$x_* = (100; 60; 120); \quad f(x_*) = 58000; \quad x_0 = (1; 1; 1).$$

Приклад 2: Функція Брауна [5]:

$$F_i(x) = x_i + \sum_{j=1}^n x_j - n - 1, \quad i = 1, 2, \dots, n - 1,$$

$$F_n(x) = \prod_{j=1}^n x_j - 1, \quad n = 8,$$

$$f(x) = \sum_{i=1}^n F_i^2(x);$$

$$x_* = (1; 1; 1; 1; 1; 1; 1; 1), \quad f(x_*) = 0; \quad x_0 = (0.5; 0.5; 0.5; 0.5; 0.5; 0.5; 0.5; 0.5)$$

Приклад 3:

$$f(x) = 2 - \frac{1}{120} x_1 x_2 x_3 x_4 x_5;$$

$$f(x_*) = 2; \quad x_0 = (2; 2; 2; 2; 2).$$

Приклад 4:

$$f(x) = x_1^2 + 5x_2^2 + 2x_3^3 + \cos(x_1 - x_2 + x_3);$$

$$x_* = (0; 0; 0); \quad f(x_*) = 1; \quad x_0 = (1; 1; 1)$$

Приклад 5:

$$f(x) = x_1^2 + 2x_2^3 - x_1 + 2x_2 + e^{x_1^2 + x_2^2};$$

$$x_* = (0.230898; -0.315910); \quad f(x_*) = 0.292997; \quad x_0 = (1; 1; 1).$$

ЛІТЕРАТУРА

1. Васильев Ф.П. Методы решения экстремальных задач. М.: Наука, 1988. 552с.
2. Деннис Дж., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений. М., Мир, 1988. 440с.
3. Ульм С. Об обобщенных разделенных разностях. I. Изв.АН ЭССР, Физ.-Матем., 1967, т.16, №1, С.13-26.
4. Шахно С.М., Гнатишин О.П. Дослідження методів типу Стеффенсена для безумовної мінімізації функцій, Київ, 1994. Деп. в ДНТБ України 30.08.94 № 1812 – Ук – 94. 9с.
5. More J.J., Garbow B.S., Hillstom K.E. Testing Unconstrained Optimization Software // ACM Transactions on Mathematical Software. 1981. Vol.7. N1. P.17-41.
6. Potra F.A. On an iterative algorithm of order 1.839... for solving nonlinear operator equations. // Numer. Funct. Anal. and Optimiz. 1984-85, N 7(1). P. 75-106.
7. Shakhno S.M., Gnatyshyn O.P. Iterative-Difference Methods for Solving Nonlinear Least-Squares Problem // Progress in Industrial Mathematics at ECMI 98. – Verlag B.G.Teubner GMBH, Stuttgart; 1999. P. 287-294.

**ON SOME ITERATIVE-DIFFERENCE METHODS FOR SOLVING
UNCONSTRAINED MINIMIZATION PROBLEMS****O.Gnatyshyn, S.Shakhno***Ivan Franko National University In Lviv
Universytetska str., 1, Lviv, 79000, e-mail: ktop@franko.lviv.ua*

Some iterative-difference methods for solving unconstrained minimization problems, which do not demand evaluation of second derivatives are presented. We have proved theorems where convergence of the proposed methods is justified and rate of convergence is established. Numerical results are presented.

Key words: unconstrained minimization, iterative–difference method; divided difference; rate of convergence.

*Стаття надійшла до редколегії 01.11.2002
Прийнята до друку 9.04.2003*