ISSN 2078–5097. Вісн. Львів. ун-ту. Сер. прикл. матем. та інф. 2013. Вип. 19. С. 3–9
Visnyk of the Lviv University. Series Appl. Math. and Informatics. Issue 19. P. 3–9

ОБЧИСЛЮВАЛЬНА МАТЕМАТИКА

УДК 519.6

РЕКУРСИВНИЙ АНАЛОГ ТРИКРОКОВОГО МЕТОДУ НЬЮТОНА

В. Бартіш

*Львівський національний університет імені Івана Франка,
вул. Університетська, 1, Львів, 79000, e-mail: bartishv@gmail.com*

Запропоновано та досліджено трикроковий рекурсивний метод розв'язування задач мінімізації, побудований на основі трикрокового методу Ньютона. Подано рекомендації стосовно вибору оптимальної, в сенсі кількості обчислень, глибини рекурсії. Наведено результати чисельних експериментів, які підтверджують ефективність запропонованих методів.

Ключові слова: задачі мінімізації, метод Ньютона, градієнтний метод.

1. ВСТУП

Знаходження оптимальних значень певних величин є життєво необхідною частиною функціонування економіки, невід'ємною частиною сучасної механіки та інженерії, відіграє важливу роль у дослідженні будь-яких операцій та опрацюванні статистичних даних. Теорія сучасної оптимізації перегукується із теорією ігор, вивченню економічного положення рівноваги. У мікроекономіці – це проблема максимізації корисності, де споживачі максимізують свою вигоду, а підприємці – свій прибуток. Теорія оптимізації зробила чималий внесок у дослідження та аналіз ризиків.

Для розв'язування задач безумовної мінімізації функції багатьох змінних існує чимало ітераційних методів, проте універсального методу не знайдено. Важливою задачею є побудова таких методів, які точніше та з меншими затратами допомагатимуть знаходити розв'язок. Наша мета – розглянути ряд трикрокових модифікацій, які ґрунтуються на методі Ньютона, або його модифікаціях та градієнтному методі.

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Розглянемо задачу безумовної мінімізації

$$f(x) \to \min, \; x \in R^n, \; f : R^n \to R, \tag{1}$$

де $f(x)$ – двічі диференційовна.

Для розв'язування задачі (1) існує багато методів. Ми пропонуємо новий алгоритм розв'язування задачі (1), побудований на основі трикрокового методу Ньютона [1] з використанням глибини рекурсії

$$\begin{aligned} u_k &= x_k - \alpha_k \left(f''(\tilde{x}_k)\right)^{-1} f'(x_k) \\ v_k &= x_k - \beta_k f'(x_k), \\ x_{k+1} &= \arg\min_t f\left(u_k + t(v_k - u_k)\right). \end{aligned} \tag{2}$$

Тут $\tilde{x}_k = x_{\left[\frac{k}{p}\right]p}$, p – глибина рекурсії, $[a]$ – ціла частина a, $\alpha_k \in (0,1]$, $\beta_k > 0$ такі, що виконуються умови монотонності, їх можна вибрати аналогічно як в [1].

Зауважимо, що алгоритм (2) при глибині рекурсії $p=1$ вироджується в трикроковий метод Ньютона [2].

3. ОБҐРУНТУВАННЯ ЗБІЖНОСТІ

Перепишемо алгоритм (2) у такому вигляді:

$$\begin{aligned} u_k^{(i)} &= x_k^{(i-1)} - \left(f''(x_k)\right)^{-1} f'\left(x_k^{(i-1)}\right) \\ v_k^{(i)} &= x_k^{(i-1)} - \beta_k^{(i)} f'\left(x_k^{(i-1)}\right), \\ x_k^{(i)} &= \arg\min_t f\left(u_k^{(i)} + t\left(v_k^{(i)} - u_k^{(i)}\right)\right), \end{aligned} \tag{3}$$

де $i=\overline{1,p}$, $k=0,1,\dots$, $x_k^{(0)}=x_k$, $x_k^{(p)}=x_{k+1}$.

Теорема. Нехай:

1) $f(x)\in C^2(R^n)$;

2) $\forall x\in D$, $h\in R^n$, $D=\{x\in R^n : f(x)\le f(x_0)\}$

$$m\|h\| \le \left(f''(x)h,h\right) \le M\|h\|,$$

де $0<m\le M<\infty$;

3) $\forall x,y\in D$, $f''(x)$ задовольняє умову Ліпшиця

$$\left\|f''(x)-f''(y)\right\| \le L\|x-y\|,$$

де $0<L<\infty$;

4) початкове наближення x_0 вибирають таким, що виконується умова

$$\mu = \frac{1}{\sqrt[p]{9}} C\left(f(x_0)-f(x_*)\right) < 1 .$$

Тут $C=\dfrac{9ML^2}{2m^4}$.

Тоді послідовність $\{x_k\}$ визначена за (3), збігається до розв'язку x_* задачі (1) і справджується оцінка

$$f(x_k)-f(x_*) \le \mu^{(p+1)^k-1}\left(\prod_{j=0}^{k-1}\left(\prod_{i=1}^{p}\gamma_j^{(i)}\right)^{(p+1)^{k-1-j}}\right)\left(f(x_0)-f(x_*)\right), \tag{4}$$

де $\gamma_j^{(i)}\le 1$; $k=1,2\dots$.

Доведення. Оскільки згідно з умовою 2) функція $f(x)$ сильно опукла, то існує розв'язок x_* і $f'(x_*)=0$, крім того, справджується оцінка [1]

$$\left\|\left(f''(x)\right)^{-1}\right\| \le \frac{1}{m} .$$

Можемо записати

$$f(x)-f(x_*) = \frac{1}{2}\int_0^1 \left(f''\left(x_*+\tau(x-x_*)\right)(x-x_*),(x-x_*)\right)d\tau .$$

Звідси

$$\frac{m}{2}\|x-x_*\|^2 \le f(x)-f(x_*) \le \frac{M}{2}\|x-x_*\|^2$$

і

$$\|x - x_*\|^2 \le \frac{2}{m}\left(f(x) - f(x_*)\right).$$

Тоді

$$u_0^{(1)} - x_* = x_0 - x_* - \left(f''(x_0)\right)^{-1}\left(f'(x_0) - f'(x_*)\right) =$$
$$= \left(f''(x_k)\right)^{-1}\left(\int_0^1 \left(f''(x_0) - f''(x_* + \tau(x_0 - x_*))\right) d\tau\right)(x_0 - x_*)$$

або

$$\left\|u_0^{(1)} - x_*\right\| \le \frac{1}{m} L \int_0^1 (1-\tau) d\tau \|x_0 - x_*\|^2 \le \frac{L}{2m}\|x_0 - x_*\|^2 .$$

З іншого боку,

$$f\left(u_0^{(1)}\right) - f(x_*) \le \frac{M}{2}\left\|u_0^{(1)} - x_*\right\|^2 \le \frac{M}{2}\left(\frac{L}{2m}\|x_0 - x_*\|^2\right)^2 \le$$
$$\le \frac{M}{2}\left(\frac{L}{m^2}\left(f(x_0) - f(x_*)\right)\right)^2 = \frac{M}{2}\left(\frac{L}{m^2}\right)^2 \left(f(x_0) - f(x_*)\right)^2 \le \frac{C}{9}\left(f(x_0) - f(x_*)\right)^2 .$$

Згідно з (3) отримаємо

$$f\left(x_0^{(1)}\right) - f(x_*) \le \frac{C}{9}\gamma_0^{(1)}\left(f(x_0) - f(x_*)\right)^2,$$

де $\gamma_0^{(1)} \in (0,1]$.

Нехай

$$f\left(x_0^{(p-1)}\right) - f(x_*) \le \frac{C^{p-1}}{9}\gamma_0^{(p-1)} \ldots \gamma_0^{(1)}\left(f\left(x_0^{(0)}\right) - f(x_*)\right)^p .$$

Тут $\gamma_0^{(1)} \in (0,1]$, $i = 1,\ldots,p$.

Тоді

$$u_0^{(p)} = x_0^{(p-1)} - \left(f''(x_0)\right)^{-1} f'\left(x_0^{(p-1)}\right)$$

і одержуємо

$$u_0^{(p)} - x_* = x_0^{(p-1)} - x_* - \left(f''(x_0)\right)^{-1}\left(f'\left(x_0^{(p-1)}\right) - f'(x_*)\right) =$$
$$= \left(f''(x_0)\right)^{-1}\left(\int_0^1 \left(f''(x_0) - f''\left(x_* + \tau\left(x_0^{(p-1)} - x_*\right)\right)\right) d\tau\right)\left(x_0^{(p-1)} - x_*\right)$$

або

$$\left\|u_0^{(p)} - x_*\right\| \le \frac{1}{m} L \int_0^1 \left(\|x_0 - x_*\| + \tau\left\|x_0^{(p-1)} - x_*\right\|\right) d\tau \left\|x_0^{(p-1)} - x_*\right\| \le$$
$$\le \frac{3L}{m^2}\sqrt{\left(f(x_0) - f(x_*)\right)\left(f\left(x_0^{(p-1)}\right) - f(x_*)\right)} .$$

Тоді

$$f\left(u_0^{(p)}\right) - f(x_*) \le \frac{M}{2}\left\|u_0^{(p)} - x_*\right\|^2 \le \frac{9L^2 M}{2m^4}\left(f(x_0) - f(x_*)\right)\left(f\left(x_0^{(p-1)}\right) - f(x_*)\right) \le$$
$$\le \frac{9L^2 M}{2m^4}\left(f(x_0) - f(x_*)\right)\frac{1}{9} C^{p-1}\gamma_0^{(p-1)} \ldots \gamma_0^{(1)}\left(f(x_0) - f(x_*)\right)^p \le$$

$$\le \frac{1}{9} C^p \gamma_0^{(p-1)} \ldots \gamma_0^{(1)} \left(f\left(x_0^{(0)}\right) - f(x_*) \right)^{p+1} = \mu^p \gamma_0^{(p-1)} \ldots \gamma_0^{(1)} \left(f(x_0) - f(x_*) \right).$$

Отже, можна записати так:

$$f(x_1) - f(x_*) \le \mu^{(p+1)^1 - 1} \left(\prod_{i=1}^{p} \gamma_0^{(i)} \right) \left(f(x_0) - f(x_*) \right).$$

Нехай

$$f(x_k) - f(x_*) \le \mu^{(p+1)^k - 1} \left(\prod_{j=0}^{k-1} \left(\prod_{i=1}^{p} \gamma_j^{(i)} \right)^{(p+1)^{k-1-j}} \right) \left(f(x_0) - f(x_*) \right).$$

Аналогічно до попереднього викладення отримаємо

$$f(x_{k+1}) - f(x_*) \le \frac{1}{9} C^p \gamma_k^{(p)} \ldots \gamma_k^{(1)} \left(f(x_k) - f(x_*) \right)^{p+1} \le$$

$$\le \frac{1}{9} C^p \gamma_k^{(p)} \ldots \gamma_k^{(1)} (\mu^{(p+1)^k - 1} \prod_{j=0}^{k-1} \left(\prod_{i=1}^{p} \gamma_j^{(i)} \right)^{(p+1)^{k-1-j}} \left(f(x_0) - f(x_*) \right))^{p+1} \le$$

$$\le \mu^{(p+1)^{k+1} - 1} \left(\prod_{j=0}^{k} \left(\prod_{i=1}^{p} \gamma_j^{(i)} \right)^{(p+1)^{k-j}} \right) \left(f(x_0) - f(x_*) \right).$$

Теорема доведена.

Легко бачити, що частковим випадком методу (2) є трикроковий метод Ньютона ($p=1$), розглянутий в [2], і рекурсивний метод Ньютона з глибиною рекурсії p, де $u_k^{(j)} = x_k^{(j)}$ [3], а також метод Ньютона при $p=1$, $u_k^1 = x_{k+1}$ [4].

Використавши оцінку (4), можна розглядати питання про оптимальну глибину рекурсії в сенсі кількості обчислень. У нашому випадку для обчислення $x_k^{(1)}$ витрачаємо Q_1 обчислень, а для $x_k^{(j)}$ $(j > 1)$ – $Q_2 < Q_1$. Тоді для отримання результату з точністю ε треба вибирати p як розв'язок задачі

$$k\left(Q_1 + (p-1)Q_2\right) \to \min_p$$

$$\mu^{(p+1)^1 - 1} \prod_{j=0}^{k-1} \left(\prod_{i=1}^{p} \gamma_j^{(i)} \right)^{(p+1)^{k-1-j}} \left(f(x_0) - f(x_*) \right) \le \varepsilon . \tag{5}$$

На жаль, у цьому вигляді обмеження досить складні і розв'язати задачу (5) практично не можливо. Тому зробимо так: виберемо за одиницю обчислень обчислення функції $f(x)$, приймемо $\gamma_j^{(i)} = 1$. У такому випадку можна вважати $Q_1 = \frac{n(n+1)}{2} + n + \delta$; $Q_2 = n + \delta$ обчислень. У цьому випадку δ – кількість обчислень $f(x)$ при одновимірній мінімізації, тобто досить невелика. Приймаємо $10 \le \delta \le 20$. Оскільки решта обчислень не впливає на результат, то ними можна знехтувати. Тоді задачу (5) перепишемо у вигляді

$$\left(\left(\frac{n(n+1)}{2} + n + \delta \right) + (p-1)(n+\delta) \right) \to \min_p \tag{6}$$

$$\mu^{(p+1)^{k+1} - 1} \left(f(x_0) - f(x_*) \right) \le \varepsilon .$$

Розв'язок задачі (6) приводить до наближеного розв'язку (5).

4. АПРОБАЦІЯ МЕТОДІВ

Для дослідження властивостей згаданої модифікації (2) виконано числові експерименти з критерієм зупинки $\left\| f'\left(x_k^{(i)}\right) \right\| \le \varepsilon$. Для визначення γ_k на третьому кроці кожної ітерації, у всіх визначених модифікаціях використовували аналог методу ДСК одновимірної мінімізації.

Для проведення числових експериментів обрали такі приклади.

Приклад 1. Розширена функція Бейля

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

$x_0' = (3.5,-0.5,\ldots,3.5,-0.5)$; $x_0'' = (6,-0.5,\ldots,6,-0.5)$; $x_* = (2.125,0,\ldots,2.125,0)$; $f(x_*) = 0.6562\dfrac{n}{2}$.

Приклад 2. Розширена функція Розенброка

$$f(x) = \sum_{i=1}^{n/2} \left[100\left(x_{2i} - x_{2i-1}^2\right)^2 + \left(1 - x_{2i-1}\right)^2 \right];$$

$x_0' = (2,2,\ldots,2)$; $x_0'' = (3,-0.5,\ldots,3,-0.5)$; $x_* = (1,1,\ldots,1)$; $f(x_*) = 0$.

Далі у таблиці подаємо кількість обчислень функції необхідних для досягнення результату згідно з запропонованим алгоритмом порівняно з трикроковим аналогом методу Ньютона

$$\begin{aligned} u_k &= x_k - \alpha_k \left(f''(x_k)\right)^{-1} f'(x_k), \\ v_k &= x_k - \beta_k f'(x_k), \\ x_{k+1} &= \arg\min_{\gamma} f\left(u_k + \gamma(v_k - u_k)\right),\ k = 0,1,\ldots \end{aligned} \tag{7}$$

та класичним методом Ньютона

$$x_{k+1} = x_k - \alpha_k \left(f''(x_k)\right)^{-1} f'(x_k) \tag{8}$$

(у таблиці кількість ітерацій позначено як *i*, а глибину рекурсії – *p*).

Як свідчать обчислення, використання рекурсивного аналогу трикрокового методу дає змогу отримати результат розв'язку задачі (1) за меншої кількості обчислень. Ефективність рекурсивного методу особливо простежується при великих n.

5. ВИСНОВКИ

Ми розглянули рекурсивний аналог трикрокового методу Ньютона, який ґрунтується на класичному методі Ньютона та градієнтному методі. Сформулювали та довели теорему про збіжність. Проведені теоретичні дослідження та практична реалізація показали ефективність методу в сенсі кількості обчислень порівняно з класичним методом Ньютона, особливо при великих n.

Метод	(8)	(7)	(2)	(2)	(2)	(2)	(2)
Пр.1 x_0' $n=2$	64 [i = 9]	187 [i = 7]	187 [i = 7; p = 2]	215 [i = 8; p = 3]	264 [i = 9; p = 5]	310 [i = 11; p = 6]	325 [i = 11; p = 7]

Пр.1 x_0' $n=10$	604 [i = 9]	607 [i = 7]	468 [i = 8; P = 2]	435 [i = 8; p = 3]	440 [i = 9; p = 5]	502 [i = 11; p = 6]	517 [i = 11; p = 7]
Пр.1 x_0' $n=50$	11944 [i = 9]	9430 [i = 7]	5671 [i = 8; p = 2]	4421 [i = 8; p = 3]	3235 [i = 9; p = 5]	3385 [i = 11; p = 6]	3394 [i = 11; p = 7]
Пр.1 x_0' $n=100$	46368 [i = 9]	36205 [i = 7]	21171 [i = 8; p = 2]	16146 [i = 8; p = 3]	11235 [i = 9; p = 5]	11485 [i = 11; p = 6]	11497 [i = 11; p = 7]
Пр.1 x_0'' $n=10$	604 [i = 9]	527 [i = 6]	252 [i = 4; p = 2]	277 [i = 5; p = 3]	238 [i = 5; p = 5]	238 [i = 5; p = 7]	238 [i = 5; p = 8]
Пр.1 x_0'' $n=50$	11944 [i = 9]	8090 [i = 6]	2855 [i = 4; p = 2]	2920 [i = 5; p = 3]	1659 [i = 5; p = 5]	1659 [i = 5; p = 7]	1659 [i = 5; p = 8]
Пр.1 x_0'' $n=100$	46369 [i = 9]	31040 [i = 6]	10605 [i = 4; p = 2]	10720 [i = 5; p = 3]	5684 [i = 5; p = 5]	5684 [i = 5; p = 6]	5684 [i = 5; p = 8]
Пр.2 x_0' $n=10$	1073 [i = 14]	1675 [i = 15]	2692 [i = 24; p = 2]	2685 [i = 29; p = 3]	3375 [i = 37; p = 5]	2864 [i = 37; p = 7]	4469 [i = 51; p = 8]
Пр.2 x_0' $n=50$	20581 [i = 14]	21233 [i = 15]	18326 [i = 24; p = 2]	16060 [i = 29; p = 3]	14629 [i = 37; p = 5]	11788 [i = 37; p = 7]	15156 [i = 52; p = 8]
Пр.2 x_0' $n=100$	82433 [i = 14]	77962 [i = 15]	64803 [i = 24; p = 2]	55260 [i = 29; p = 3]	46675 [i = 37; p = 5]	36291 [i = 37; p = 7]	44202 [i = 52; p = 8]
Пр.2 x_0'' $n=10$	1475 [i = 19]	2427 [i = 21]	4024 [i = 34; p = 2]	4194 [i = 41; p = 3]	4406 [i = 47; p = 5]	4870 [i = 55; p = 7]	5231 [i = 62; p = 8]
Пр.2 x_0'' $n=50$	28887 [i = 19]	29195 [i = 21]	26125 [i = 34; p = 2]	22918 [i = 41; p = 3]	18485 [i = 47; p = 5]	17210 [i = 55; p = 7]	18209 [i = 64; p = 8]
Пр.2 x_0'' $n=100$	113345 [i = 19]	109213 [i = 21]	92016 [i = 34; p = 2]	77847 [i = 41; p = 3]	58592 [i = 47; p = 5]	50165 [i = 55; p = 7]	51674 [i = 64; p = 8]

Список використаної літератури

1. В*асильев Ф.П.* Методы оптимизации / Ф.П. Васильев. – М.: Факториал Пресс, 2002. – С. 824.
2. *Бартіш М.Я.* Трикроковий алгоритм мінімізації функції / М. Бартіш, Н. Огородник // Математичні студії. – 2008. – Т. 29. – С. 108–112.
3. *Бартіш М.Я.* Модифікації рекурсивного методу Ньютона / М. Бартіш, Н. Огородник // Вісн. ЛНУ. Сер. Прикл. матем. та інформ. – 2010. Вип. 16. – С. 3 -9.

4. *Nocedal J.* Numerical optimization / J. Nocedal, S.J. Wright. – New York: Springer, 2006. – P. 665.

Стаття: надійшла до редколегії 17.10.2012
доопрацьована 28.11.2012
прийнята до друку 05.12.2012

РЕКУРСИВНЫЙ АНАЛОГ ТРИШАГОВОГО МЕТОДА НЬЮТОНА

В. Бартиш

Львовский национальный университет имени Ивана Франко,
ул. Университетская, 1, Львов, 79000, e-mail: bartishv@gmail.com

Предложено и исследовано тришаговый рекурсивный метод решения задач минимизации, базирующийся на тришаговом методе Ньютона. Подано рекомендации относительно выбора оптимальной, по количеству исчислений, глубины рекурсии. Наведено результаты численных экспериментов, которые подтверждают эффективность предложенных методов.

Ключевые слова: задача минимизации, метод Ньютона, градиентный метод.

RECURSIVE ANALOGUE OF THREE STEPPED NEWTON'S METHOD

V. Bartish

Ivan Franko National University of Lviv,
Universytetska str., 1, Lviv, 79000, e-mail: bartishv@gmail.com

A three stepped recursive methods of solving minimization problems based on three stepped Newton`s method, is presented and investigated in this paper. Recommendations on selection of optimal, by amount of calculations, recursion depth are provided. Also results of numerical experiments which prove effectiveness of proposed methods are presented.

Key words: minimization problem, Newton`s method, gradient method.