ЯРОВИЙ Анатолій Трохимович,
канд. фіз.-мат. наук, доцент

СТРАХОВ Євген Михайлович,
канд. фіз.-мат. наук

ОДИН ТРИКРОКОВИЙ МЕТОД МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

Одеський національний університет імені І. І. Мечникова
Кафедра оптимального керування і економічної кібернетики

Одеса – 2015

Постановка задачі

$$\begin{aligned} \varphi(x) &\to \min, \\ x &\in \mathbb{R}^n \end{aligned} \tag{1}$$

$\varphi(x) : \mathbb{R}^n \to \mathbb{R}^1$ – неперервно диференційовна функція.

Позначимо

$$\varphi'(x) = \left(\frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2}, \ldots, \frac{\partial \varphi}{\partial x_n} \right)^T.$$

Чисельні методи розв'язування задач типу (1)

- методи нульового порядку;
- методи першого порядку;
- методи другого порядку.

- однокрокові методи;
- багатокрокові методи.

Класичний метод спряжених градієнтів

$$
\begin{gathered}
x^{k+1} = x^k + \beta_k s^k,\ k = 0, 1, \dots, \\
s^k = \begin{cases} -\varphi'(x^0), & k = 0, \\ -\varphi'(x^k) + \gamma_{k-1} s^{k-1}, & k = 1, 2, \dots \end{cases}
\end{gathered} \tag{2}
$$

x^0, x^1, $\dots, x^k$, $\dots$ – послідовні наближення,
s^0, s^1, $\dots, s^k$, $\dots$ – напрямки спуску,
β_k – величина кроку вздовж напрямку спуску,
γ_{k-1} – числовий параметр.

Способи вибору кроку β_k

1) розв'язування задачі одновимірної мінімізації:

$$\beta_k : \min_{\beta \geqslant 0} \varphi\left(x^k + \beta s^k\right). \tag{3}$$

2) умови Вольфе (Wolfe):

$$\begin{gathered} \varphi\left(x^k + \beta_k s^k\right) - \varphi\left(x^k\right) \leqslant \delta \beta_k \left[\varphi'\left(x_k\right)\right]^T s_k, \\ \left[\varphi'\left(x_k + \beta_k s^k\right)\right]^T s_k \geqslant \sigma \left[\varphi'\left(x_k\right)\right]^T s_k, \end{gathered} \tag{4}$$

де $0 < \delta \leqslant \sigma < 1$ – деякі константи.

Способи вибору параметру γ_{k-1}

1) метод Флетчера – Рівза (Fletcher – Reeves):

$$\gamma_{k-1} = \frac{\left\|\varphi'\left(x^k\right)\right\|^2}{\left\|\varphi'\left(x^{k-1}\right)\right\|^2};$$

2) метод Полака – Ріб'єра – Поляка (Polak – Ribiere – Polyak):

$$\gamma_{k-1} = \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^k\right) - \varphi'\left(x^{k-1}\right)\right)}{\left\|\varphi'\left(x^{k-1}\right)\right\|^2}.$$

Тут і далі $\|\cdot\|$ означає евклідову норму вектора, $(\cdot,\cdot)$ – скалярний добуток векторів.

Трикроковий метод

$$
\begin{aligned}
x^{k+1} &= x^k + \beta_k s^k,\ k = 0, 1, \ldots, \\
s^k &= \begin{cases} -\varphi'\left(x^k\right), & k = 0, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1}, & k = 1, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1} + \gamma_{k-2} s^{k-2}, & k = 2, 3, \ldots \end{cases}
\end{aligned}
\tag{5}
$$

$x^0, \ldots, x^k, \ldots$ – послідовні наближення,
$s^0, \ldots, s^k, \ldots$ – напрямки спуску,
ξ_{k-1}, γ_{k-2} – числові параметри.

Крок β_k будемо визначати з умови (3).

Побудова системи взаємно спряжених напрямків

Означення

Нехай A – симетрична додатно означена матриця.
Вектори s' і s'' називаються спряженими (відносно матриці A), якщо вони відмінні від нуля і $(As', s'') = 0$.
Вектори s^0, s^1, $\ldots, s^k$ називаються взаємно спряженими (відносно матриці A), якщо всі вони відмінні від нуля і $\left(As^i, s^j\right) = 0$, $i \neq j$, $0 \leqslant i, j \leqslant k$.

Побудова системи взаємно спряжених напрямків

Розглянемо квадратичну функцію

$$\varphi(x) = \frac{1}{2}(Ax, x) + (b, x) + c.$$

Побудуємо систему взаємно спряжених напрямків за правилом (5).

$$0 = \left(s^k, As^{k-1}\right) = -\left(\varphi'\left(x^k\right), As^{k-1}\right) +$$
$$+\xi_{k-1}\left(s^{k-1}, As^{k-1}\right) + \gamma_{k-2}\left(s^{k-2}, As^{k-1}\right).$$

Звідси

$$\xi_{k-1} = \frac{\left(\varphi'\left(x^k\right), As^{k-1}\right)}{\left(s^{k-1}, As^{k-1}\right)}. \tag{6}$$

Побудова системи взаємно спряжених напрямків

Далі,

$$0 = \left(s^k, As^{k-2}\right) = -\left(\varphi'\left(x^k\right), As^{k-2}\right) +$$
$$+\xi_{k-1}\left(s^{k-1}, As^{k-2}\right) + \gamma_{k-2}\left(s^{k-2}, As^{k-2}\right).$$

Отримаємо

$$\gamma_{k-2} = \frac{\left(\varphi'\left(x^k\right), As^{k-2}\right)}{\left(s^{k-2}, As^{k-2}\right)}. \tag{7}$$

Властивості алгоритму (5)–(7)

Лема

Для диференційовної функції $\varphi(x)$ *послідовність* $\left\{x^k\right\}$, *що побудована за алгоритмом (5), (6), (7), така, що*

$$\left(\varphi'\left(x^{k+1}\right), s^k\right) = 0, \; k = 0, 1, \ldots$$

Властивості алгоритму (5)–(7)

Теорема

Нехай послідовність $\left\{x^k\right\}$ *побудована за алгоритмом (5), (6), (7). Тоді вектори* $\varphi'\left(x^k\right)$ *і* $\varphi'\left(x^{k+1}\right)$ *ортогональні,* $k = 0, 1, \ldots$

Теорема

Нехай $x^0 \in \mathbb{R}^n$, *точки* $x^1, \ldots, x^{n-1}$ *і вектори* $s^0, s^1, \ldots,\ s^{n-1}$ *отримані за формулами (5), (6), (7) і* $\varphi'\left(x^k\right) \neq 0$ $(i = 0, \ldots, n-1)$. *Тоді вектори* $s^0, s^1, \ldots, s^{n-1}$ *взаємно спряжені, а градієнти* $\varphi'\left(x^0\right)$, $\varphi'\left(x^1\right)$, …, $\varphi'\left(x^{n-1}\right)$ *попарно взаємно ортогональні.*

Трикроковий метод для неквадратичної функції

$$x^{k+1} = x^k + \beta_k s^k,\ k = 0, 1, \ldots,$$

$$s^k = \begin{cases} -\varphi'\left(x^k\right), & k = 0, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1}, & k = 1, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1} + \gamma_{k-2} s^{k-2}, & k = 2, 3, \ldots \end{cases} \tag{5}$$

$$\beta_k : \min_{\beta \geqslant 0} \varphi\left(x^k + \beta s^k\right) \tag{3}$$

$$\xi_{k-1} = \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^k\right) - \varphi'\left(x^{k-1}\right)\right)}{\|\varphi'\left(x^{k-1}\right)\|^2}, \tag{8}$$

$$\gamma_{k-2} = \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^{k-1}\right) - \varphi'\left(x^{k-2}\right)\right)}{\|\varphi'\left(x^{k-2}\right)\|^2}. \tag{9}$$

Обчислювальний експеримент

Задача 1.

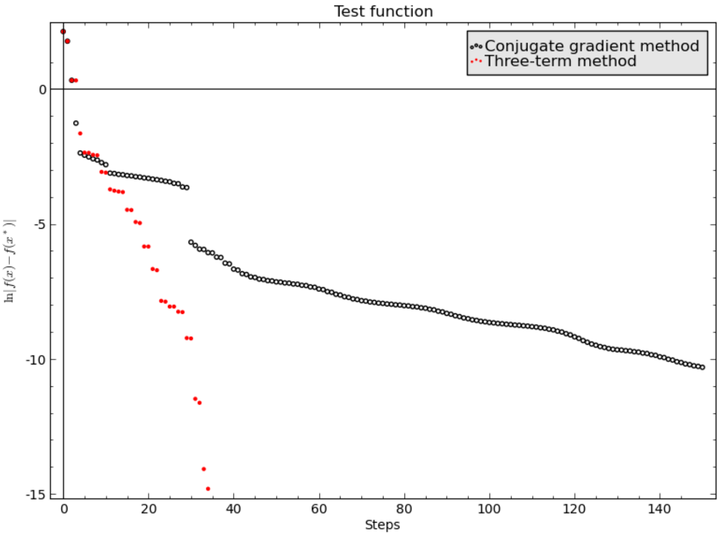
$$\varphi(x) = 100\left[x_3 - \left(\frac{x_1 + x_2}{2}\right)^2\right]^2 + (1 - x_1)^2 + (1 - x_2)^2$$

а) початкова точка $x_0^1 = (-1.2, 2, 0)^T$, $\varphi(x_0^1) = 8.40$;
вибір β з умови мінімуму

МСГ:
('Minimum point', (1.0054, 1.0018, 1.0073))
('Objective function value', 0.0000326625242232925)
('Gradient norm', 0.00860567434228524)
('Steps', 150)

Трикроковий метод:
('Minimum point', (1.0000, 0.9997, 0.9997))
('Objective function value', 9.77552986336505e-8)
('Gradient norm', 0.00422627482790298)
('Steps', 34)

Обчислювальний експеримент



Обчислювальний експеримент

б) початкова точка $x_0^1 = (-1.2, 2, 0)^T$, $\varphi(x_0^1) = 8.40$;
вибір β з умов Вольфе

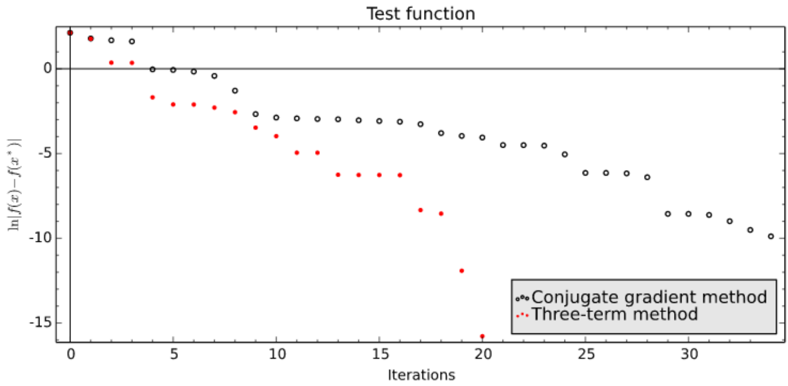
МСГ:

('Minimum point', (1.0038, 1.0060, 1.0099))
('Objective function value', 0.0000509378296007501)
('Gradient norm', 0.00889007602388954)
('Steps', 34)

Трикроковий метод:

('Minimum point', (0.9999, 1.0000, 0.9999))
('Objective function value', 6.96641524466912e-8)
('Gradient norm', 0.00819417509668262)
('Steps', 20)

Обчислювальний експеримент

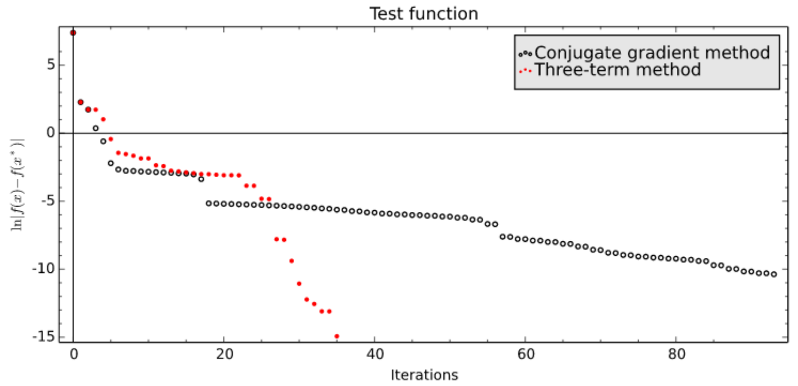


Обчислювальний експеримент

в) початкова точка $x_0^2 = (-2, 2, 4)^T$, $\varphi(x_0^2) = 1610$;
вибір β з умови мінімуму
МСГ:
('Minimum point', (1.0049, 1.0026, 1.0076))
('Objective function value', 0.0000310899302362611)
('Gradient norm', 0.00950630202580792)
('Steps', 93)

Трикроковий метод:
('Minimum point', (1.0001, 0.9995, 0.9996))
('Objective function value', 2.79845362105569e-7)
('Gradient norm', 0.00107100661098176)
('Steps', 35)

Обчислювальний експеримент



Обчислювальний експеримент

г) початкова точка $x_0^2 = (-2, 2, 4)^T$, $\varphi(x_0^2) = 1610$;
вибір β з умов Вольфе

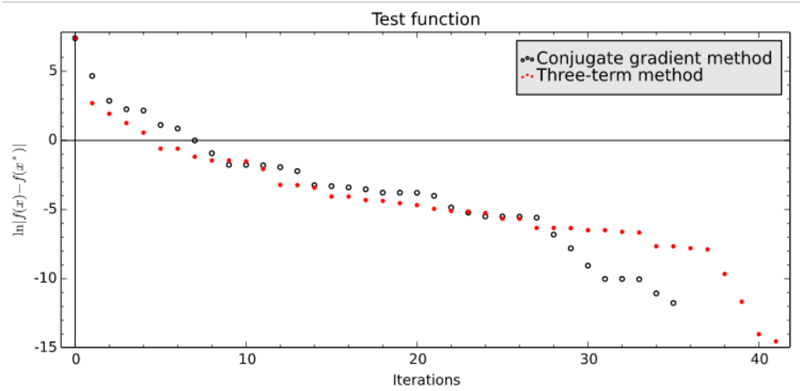
МСГ:

('Minimum point', (0.9989, 1.0026, 1.0015))
('Objective function value', 7.85208283127226e-6)
('Gradient norm', 0.00607748367803774)
('Steps', 35)

Трикроковий метод:

('Minimum point', (1.0001, 1.0001, 1.0002))
('Objective function value', 3.49102795560441e-8)
('Gradient norm', 0.00264475726822080)
('Steps', 41)

Обчислювальний експеримент



Обчислювальний експеримент

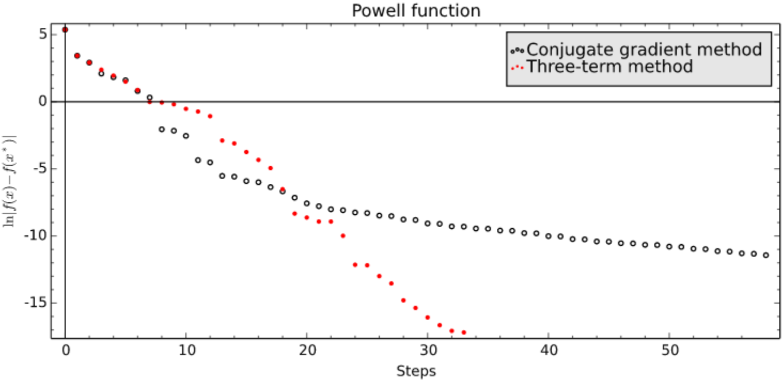
Задача 2. Функція Пауелла, $n = 4$:

$$\varphi(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

а) початкова точка $x_0^1 = (3, -1, 0, 1)^T$, $\varphi(x_0^1) = 215$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 1.0529042590143227e-05)
('Gradient norm', 0.009383531208790138)
('Iterations', 58)

Трикроковий метод:
('Objective function value', 2.3244293224508836e-08)
('Gradient norm', 0.0008646198801780563)
('Steps', 33)

Обчислювальний експеримент

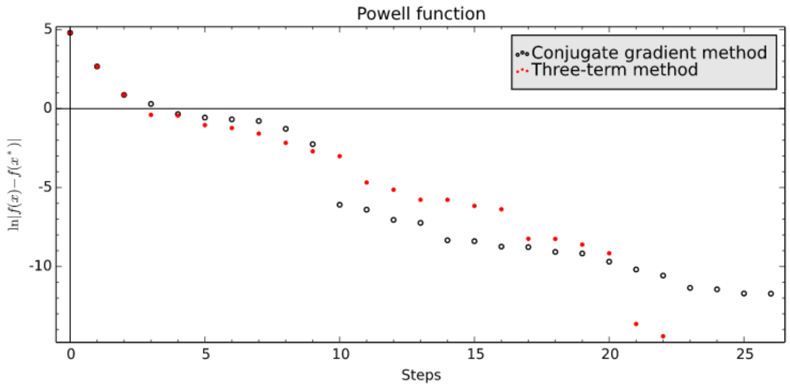


Обчислювальний експеримент

б) початкова точка $x_0^2 = (1, 1, 1, 1)^T$, $\varphi(x_0^2) = 125$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 8.11149176380214e-6)
('Gradient norm', 0.00208296456096100)
('Iterations', 26)

Трикроковий метод:
('Objective function value', 5.46609388197715e-7)
('Gradient norm', 0.000702595949732491)
('Steps', 22)

Обчислювальний експеримент



Обчислювальний експеримент

в) початкова точка $x_0^1 = (3, -1, 0, 1)^T$, $\varphi(x_0^1) = 215$;
вибір β з умов Вольфе

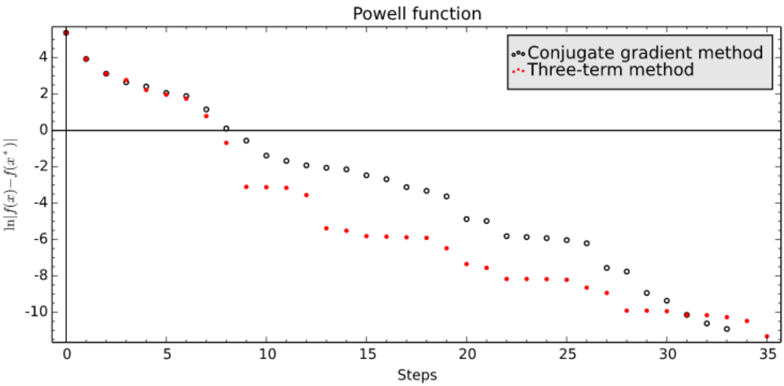
МСГ:

('Objective function value', 0.0000174314915601399)
('Gradient norm', 0.00221580127887057)
('Iterations', 33)

Трикроковий метод:

('Objective function value', 0.0000121049468401826)
('Gradient norm', 0.00261827788302757)
('Steps', 35)

Обчислювальний експеримент



Обчислювальний експеримент

г) початкова точка $x_0^2 = (1, 1, 1, 1)^T$, $\varphi(x_0^2) = 125$;
вибір β з умов Вольфе

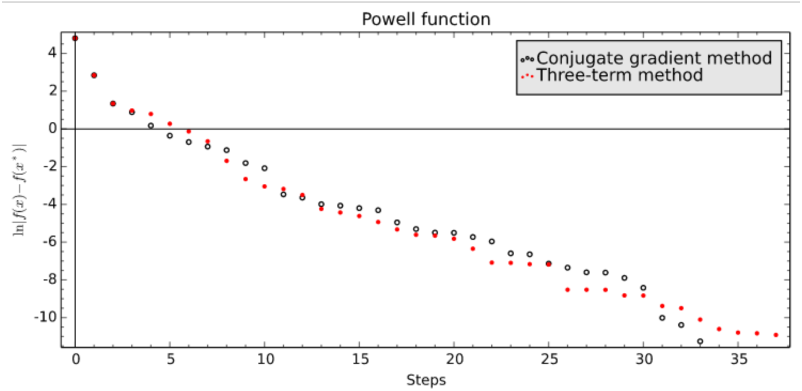
МСГ:

('Objective function value', 0.0000129291047868767)
('Gradient norm', 0.00433387819471842)
('Iterations', 33)

Трикроковий метод:

('Objective function value', 0.0000174601456839417)
('Gradient norm', 0.00629935293289066)
('Steps', 37)

Обчислювальний експеримент



Обчислювальний експеримент

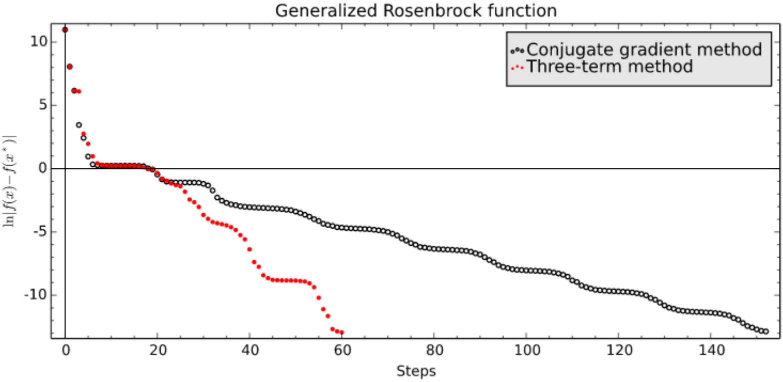
Задача 3. Узагальнена функція Розенброка, $n = 8$:

$$\varphi(x) = \sum_{i=1}^{4} \left[(1 - x_{2i-1})^2 + 100 \left(x_{2i} - x_{2i-1}^2 \right)^2 \right]$$

а) початкова точка $x_0^1 = (2, 4, 2, 4, \ldots)^T$, $\varphi(x_0^1) = 58831$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 2.472494344374354e-06)
('Gradient norm', 0.009753328526332835)
('Steps', 152)

Трикроковий метод:
('Objective function value', 2.3423347528250815e-06)
('Gradient norm', 0.005568626678495983)
('Steps', 60)

Обчислювальний експеримент



Обчислювальний експеримент

б) початкова точка $x_0^1 = (2, 4, 2, 4, \ldots)^T$, $\varphi(x_0^1) = 58831$;
вибір β з умов Вольфе

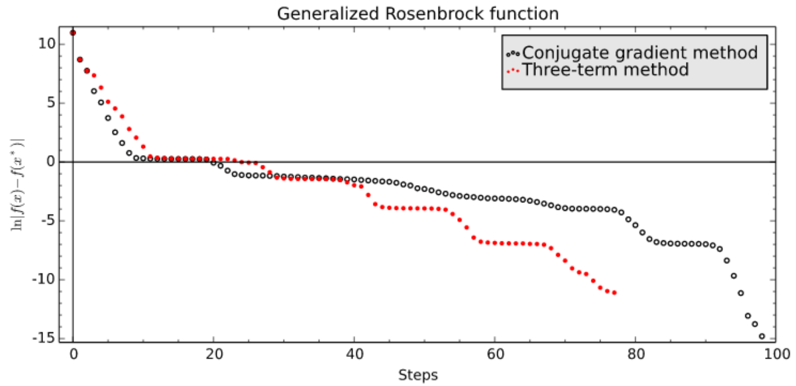
МСГ:

('Objective function value', 2.53290835416919e-7)
('Gradient norm', 0.00722652721724808)
('Steps', 98)

Трикроковий метод:

('Objective function value', 0.0000149272677703038)
('Gradient norm', 0.00517673562084944)
('Steps', 77)

Обчислювальний експеримент



Обчислювальний експеримент

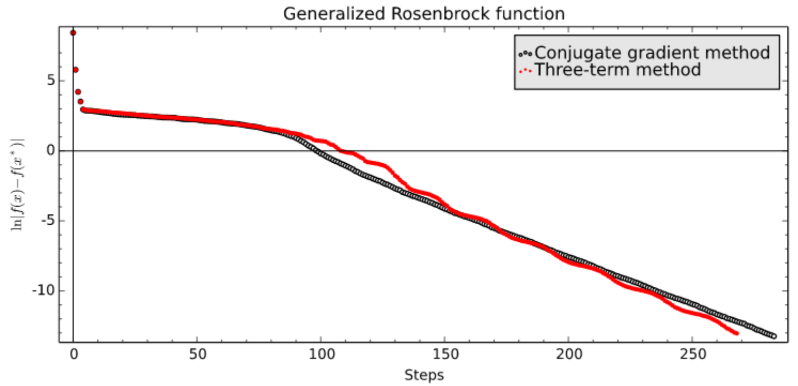
Задача 4. Узагальнена функція Розенброка, $n = 20$:

$$\varphi(x) = \sum_{i=1}^{10} \left[(1 - x_{2i-1})^2 + 100 \left(x_{2i} - x_{2i-1}^2 \right)^2 \right]$$

а) початкова точка $x_0^1 = (-1.2, 1, -1.2, 1, \ldots)^T$, $\varphi(x_0^1) = 4598$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 1.6955823696207517e-06)
('Gradient norm', 0.009839449460052852)
('Steps', 283)

Трикроковий метод:
('Objective function value', 2.0714581410175015e-06)
('Gradient norm', 0.009783084973888882)
('Steps', 268)

Обчислювальний експеримент

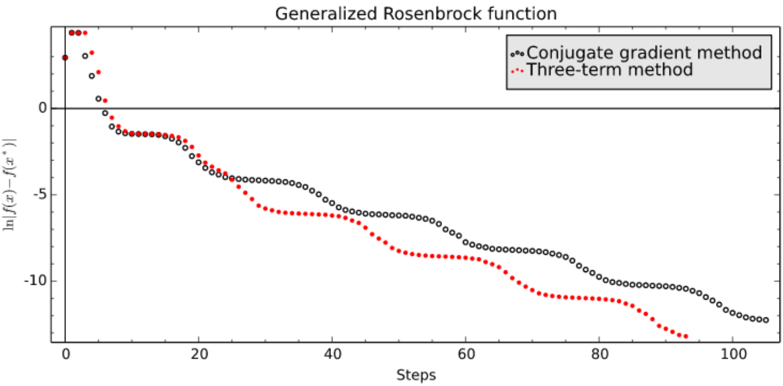


Обчислювальний експеримент

б) початкова точка $x_0^2 = (0, 0, \ldots)^T$, $\varphi(x_0^2) = 19$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 4.668872563229343e-06)
('Gradient norm', 0.009355910979872144)
('Steps', 105)

Трикроковий метод:
('Objective function value', 1.7606163140726494e-06)
('Gradient norm', 0.00854012185872228)
('Steps', 93)

Обчислювальний експеримент



Обчислювальний експеримент

Задача 5. Тридіагональна функція Бройдена, $n = 10$:

$$\varphi(x) = \left(3x_1 - 2x_1^2\right)^2 + \sum_{i=2}^{9} \left[3x_i - 2x_i^2 - x_{i-1} - 2x_{i+1} + 1\right]^2 +$$

$$+ \left(3x_{10} - 2x_{10}^2 - x_9 + 1\right)^2$$

початкова точка $x_0 = (-1, \ldots, -1)^T$, $\varphi(x_0^2) = 52$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 1.95671060687630e-7)
('Gradient norm', 0.00413444281005469)
('Steps', 17)

Трикроковий метод:
('Objective function value', 2.66073750143647e-7)
('Gradient norm', 0.00405672186955735)
('Steps', 15)

Обчислювальний експеримент

Задача 6. Узагальнена функція Біла (Beale), $n = 100$:

$$\varphi(x) = \sum_{i=1}^{n/2} \Big[\left(1.5 - x_{2i-1}\left(1 - x_{2i}\right)\right)^2 + \left(2.25 - x_{2i-1}\left(1 - x_{2i}^2\right)\right)^2 +$$

$$+ \left(2.625 - x_{2i-1}\left(1 - x_{2i}^3\right)\right)^2 \Big]$$

початкова точка $x_0 = (1, 0.8 \ldots, 1, 0.8)^T$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 4.85431854713170e-8) ('Gradient norm', 0.00197700949155216) ('Steps', 11)

Трикроковий метод:
('Objective function value', 4.85431854713170e-8)
('Gradient norm', 0.00197700949155216)
('Steps', 8)

Обчислювальний експеримент

Задача 7. Узагальнена функція Манєвича (Manevich), $n = 200$:

$$\varphi(x) = \sum_{i=1}^{n} \frac{(1 - x_i)^2}{2^i}$$

початкова точка $x_0 = (0, \dots, 0)^T$;
вибір β з умови мінімуму
МСГ:
('Objective function value', 0.000978472890249305)
('Gradient norm', 0.00195528425819557)
('Steps', 9)

Трикроковий метод:
('Objective function value', 0.000978472679849719)
('Gradient norm', 0.00195506945529611)
('Iterations', 9)

Література

- **Карманов В. Г.** Математическое программирование. – М. : ФИЗМАТЛИТ, 2004. – 264 с.
- **Сухарев А. Г., Тимохов А. В., Федоров В. В.** Курс методов оптимизации. – М. : ФИЗМАТЛИТ, 2005. – 368 с.
- **Химмельблау Д.** Прикладное нелинейное программирование. – М. : Мир, 1972. – 536 с.
- **Neculai Andrei.** An Unconstrained Optimization Test Functions Collection // Advanced Modeling and Optimization. – Vol. 10, No. 1. – 2008. – pp. 147–161.
- **Zhongbo Sun, Yantao Tian, and Hongyang Li.** Two Modified Three-Term Type Conjugate Gradient Methods and Their Global Convergence for Unconstrained Optimization // Mathematical Problems in Engineering. – Vol. 2014. – 2014. – http://dx.doi.org/10.1155/2014/394096

Тези

Яровий А. Т., Страхов Є. М. Про один трикроковий метод для задач безумовної оптимізації // Інформатика та системні науки (ІСН-2015): матеріали VI Всеукраїнської науково-практичної конференції за міжнародною участю, (м. Полтава, 19–21 берез. 2015 р.). – Полтава: ПУЕТ, 2015. http://dspace.puet.edu.ua/handle/123456789/2452