

В.М. Задачин

Харківський національний економічний університет ім. С. Кузнеця, Харків

КОМБІНОВАНИЙ МЕТОД ДЛЯ РОЗВ'ЯЗАННЯ ВИРОДЖЕНИХ ЗАДАЧ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ

Представлено метод другого порядку для розв'язання вироджених задач безумовної оптимізації, який є комбінацією двох методів: методу Ньютона та градієнтного методу. На кожній ітерації весь простір представляється як декартовий добуток двох підпросторів: ядра матриці Гессе та ортогонального доповнення до нього. На ядрі матриці Гессе працює градієнтний метод, а на ортогональному доповненні до нього – метод Ньютона. При цьому застосовується параметр регуляризації чисельного методу для розділення простору на два ортогональних підпростори. Розглядається також квазі-ньютонівський варіант представленого методу. Ефективність квазі-ньютонівського варіанту метода підтверджується чисельними експериментами, які були проведені на загальноприйнятих тестових функціях для задач безумовної оптимізації.

Ключові слова: безумовна оптимізація, вироджена точка мінімуму, модифікований метод Ньютона, спектральний розклад матриці.

Вступ

Хоча це вже й дуже стара тематика, але безумовна оптимізація є областю, що завжди актуальна для багатьох вчених. Результати безумовної оптимізації застосовуються в різних галузях науки, а також в цілому на практиці [1–2].

Однак при розв'язанні практичних задач точка екстремуму обраного критерію оптимальності нерідко виявляється виродженою, що значно ускладнює її пошук.

Аналіз останніх досліджень і публікацій. Відомі чисельні методи розв'язання загальної задачі безумовної оптимізації, до другого порядку включно, мають дуже низьку швидкість збіжності в разі розв'язання вироджених задач [3]. Це пояснюється тим, що для істотного підвищення швидкості збіжності в цьому випадку необхідно використання в методі похідних більш високого порядку, ніж другий [3–4]. Однак використання похідних 3-го і, тим більш, 4-го порядку робить чисельний метод дуже трудомістким.

Незважаючи на те, що були розроблені досить ефективні ньютонівські та квазі-ньютонівські методи безумовної оптимізації [5], інтерес до них не згасає й досі [6–15]. При цьому в разі розв'язання вироджених задач безумовної оптимізації застосовується підхід, пов'язаний з регуляризацією чисельних методів [6–7].

Метою даної статті є розробка чисельного методу другого порядку для розв'язання вироджених задач безумовної оптимізації, ідея якого (на відміну від регуляризації) полягає в поділі всього простору на два ортогональних підпростори. Поділ простору засновано на спектральному розкладанні гессіану. На кожному підпросторі є своя поведінка цільової

функції, і тому застосовується відповідний метод мінімізації.

Виклад основного матеріалу

Комбінований метод другого порядку

Розглядається вироджена задача безумовної оптимізації:

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

де $f(x)$ – двічі диференційована функція, є $x^* \in R^n$ – локальна точка мінімуму функції $f(x)$, матриця Гессе $f''(x^*)$ вироджена.

Комбінований метод другого порядку для розв'язання задачі (1) будує ітераційну послідовність наближень точки мінімуму за формулою

$$x^{k+1} = x^k + \alpha_{k1} u_1^k + \alpha_{k2} u_2^k, \quad k = 0, 1, 2, \dots, \quad (2)$$

де x^0 – початкове наближення точки мінімуму, u_1^k, u_2^k – ортогональні вектори, $\alpha_{k1} > 0$, $\alpha_{k2} > 0$ – шагові множники вздовж, відповідно, напрямків u_1^k, u_2^k . Ортогональні вектори u_1^k, u_2^k визначаються наступним чином.

На кожній k -й ітерації обчислюється вектор $g^k = f'(x^k)$ та матриця $H_k = f''(x^k)$. Так як матриця H_k є симетричною, то відповідно до спектрального розкладання представляється у вигляді

$$H_k = Q_k \Lambda_k Q_k^T, \quad (3)$$

де Q_k – ортогональна матриця, $\Lambda_k = \text{diag}(\lambda_i^k)$, а $\lambda_i^k (i = 1, \dots, n)$ – власні значення матриці H_k , що

впорядковані по спадаючій за абсолютною величиною.

Представимо тепер діагональну матрицю Λ_k в блочному вигляді

$$\Lambda_k = \begin{bmatrix} \Lambda_{k1} & 0 \\ 0 & \Lambda_{k2} \end{bmatrix},$$

де $\Lambda_{k1} = \text{diag}(\lambda_i^k)$, $|\lambda_i^k| > \varepsilon$ ($i = 1, \dots, r_k$), $r_k \leq n$, $\varepsilon > 0$ – параметр регуляризації чисельного методу, $\Lambda_{k2} = \text{diag}(\lambda_i^k)$, $|\lambda_i^k| \leq \varepsilon$ ($i = r_k + 1, \dots, n$). Тоді матрицю Q_k також запишемо у блочному вигляді $Q_k = [Q_{k1} \quad Q_{k2}]$, де Q_{k1} – блок розмірності $n \times r_k$, а Q_{k2} – блок розмірності $(n - r_k) \times n$.

Тепер матрицю H_k у зв'язку з (3) можна представити наступним чином:

$$H_k = [Q_{k1} \quad Q_{k2}] \begin{bmatrix} \Lambda_{k1} & 0 \\ 0 & \Lambda_{k2} \end{bmatrix} \begin{bmatrix} Q_{k1}^T \\ Q_{k2}^T \end{bmatrix} =$$

$$= Q_{k1} \Lambda_{k1} Q_{k1}^T + Q_{k2} \Lambda_{k2} Q_{k2}^T = H_{k\varepsilon} + E_{k\varepsilon},$$

де $H_{k\varepsilon} = Q_{k1} \Lambda_{k1} Q_{k1}^T$, $E_{k\varepsilon} = Q_{k2} \Lambda_{k2} Q_{k2}^T = H_k - H_{k\varepsilon}$.

Далі будемо ортогональні проектори: $P_k = I - Q_{k1} Q_{k1}^T$, $P_k^\perp = I - P_k = Q_{k1} Q_{k1}^T$ (I – одинична матриця) на підпростір $\text{Ker} H_{k\varepsilon} = \{x \in R^n \mid H_{k\varepsilon} x = 0\}$ та ортогональне доповнення до нього, відповідно. Відмітимо, що з ортогональності матриці Q_k випливає, що $P_k = Q_{k2} Q_{k2}^T$.

Тепер функція $f(x)$ в околиці точки x^k наближається функцією

$$f_k(x) = f_k(u_1, u_2) = f(x^k) + (P_k^\perp g^k, u_1) + (P_k g^k, u_2) + \frac{1}{2} H_{k\varepsilon} [(u_1)^2] + \frac{1}{2} E_{k\varepsilon} [(u_2)^2],$$

яка виходить з розкладу в ряд Тейлора до другого порядку, з урахуванням того, що: $x - x^k = u_1 + u_2$, $u_1 = P_k^\perp (x - x^k)$, $u_2 = P_k (x - x^k)$, $H_{k\varepsilon} u_2 = 0$, $E_{k\varepsilon} u_1 = 0$.

Тоді вектори u_1^k, u_2^k визначаються як точка мінімуму функції $f(u_1^k, u_2^k)$, тобто з системи рівнянь:

$$\frac{\partial f_k(u_1, u_2)}{\partial u_1} = P_k^\perp g^k + H_{k\varepsilon} u_1 = 0; \quad (4)$$

$$\frac{\partial f_k(u_1, u_2)}{\partial u_2} = P_k g^k + E_{k\varepsilon} u_2 = 0. \quad (5)$$

З рівняння (4) знаходимо: $u_1 = -H_{k\varepsilon}^\perp P_k^\perp g^k$.

Тут $H_{k\varepsilon}^\perp$ – псевдообернена матриця до матриці $H_{k\varepsilon}$ (в загальному випадку, тобто при $r_k \leq n$, вона вироджена), тому з (3) і ортогональності матриці Q_k випливає, що $H_{k\varepsilon}^\perp = (Q_{k1}^T)^\perp \Lambda_{k1}^{-1} (Q_{k1})^\perp = Q_{k1} \Lambda_{k1}^{-1} Q_{k1}^T$. Таким чином, отримуємо, що

$$u_1^k = -H_{k\varepsilon}^\perp P_k^\perp g^k = -Q_{k1} \Lambda_{k1}^{-1} Q_{k1}^T P_k^\perp g^k. \quad (6)$$

Визначимо тепер вектор u_2 з рівняння (5). Так як всі діагональні елементи матриці Λ_{k2} по абсолютній величині менше ніж параметр регуляризації ε , то замінимо їх на ε , тобто замінимо матрицю $E_{k\varepsilon} = Q_{k2} \Lambda_{k2} Q_{k2}^T$ на матрицю $\varepsilon Q_{k2} Q_{k2}^T = \varepsilon P_k$. Тоді з рівняння (5) маємо

$$u_2^k = -\varepsilon^{-1} P_k g^k. \quad (7)$$

Відмітимо, що описаний метод (2; 6–7) розв'язання виродженої задачі (1) є комбінацією двох методів: методу Ньютона та градієнтного методу. На кожній k -й ітерації весь простір R^n представляється як декартовий добуток двох підпросторів: $\text{Ker} H_{k\varepsilon}$ (ядро матриці $H_k = f''(x^k)$) та ортогонального доповнення до нього. На підпросторі $\text{Ker} H_{k\varepsilon}$ працює градієнтний метод, а на ортогональному доповненні до нього – метод Ньютона. При цьому параметр регуляризації чисельного методу ε є деяким критерієм, за яким виконується розподіл простору на два ортогональні підпростори. Також, у зв'язку з цим, виникає проблема визначення двох крокових множників: α_{k1} та α_{k2} з (2), вдовж, відповідно, напрямків u_1^k і u_2^k , тобто в різних ортогональних підпросторах.

Як показали результати чисельних експериментів, найбільш ефективним алгоритмом визначення крокових множників $\alpha_{k1} > 0$, $\alpha_{k2} > 0$ виявився алгоритм, аналогічний методу Гаука-Дживаса.

У даній статті не проводиться теоретичне дослідження швидкості збіжності методу (2–6; 7) в разі розв'язання виродженої задачі (1). Однак, так як цей метод є комбінацією двох методів: методу Ньютона і градієнтного методу, то його швидкість збіжності буде не гірше, ніж швидкість збіжності градієнтного методу, поведінка якого в цьому випадку досліджувалась в роботі [16].

Так як обчислення матриці $H_k = f''(x^k)$ є досить трудомісткою процедурою, то є сенс розглядати і квазі-ньютонівський варіант методу (2–6; 7), що застосовує формулу Broyden–Fletcher–Goldfarb–Shanno (BFGS) [5].

Результати чисельних експериментів

Метод (2–6; 7), а також його квази-ньютонівський варіант, розв'язання виродженої задачі (1) були реалізовані в середовищі математичного пакету R. Для стійкості методу до розходження від'ємні діагональні елементи λ_i^k матриці Λ_k в (3) замінювались на $-\lambda_i^k$. Параметр регуляризації чисельного методу ε брався рівним $10^{-5} \times \max_{i=1, \dots, n} (|\lambda_i^k|)$, тобто вважалось, що матриця H_k близька до виро-

дженої, якщо її число обумовленості $\text{cond}(H_k)$ більше, ніж 10^5 . Похідні $g^k = f'(x^k)$ та $H_k = f''(x^k)$ обчислювались чисельно за симетричними формулами:

$$\frac{\partial f(x)}{\partial x_i} \approx \frac{f(x|x_i+h_i) - f(x|x_i-h_i)}{2h_i},$$

$$\frac{\partial^2 f(x)}{\partial x_i^2} \approx \frac{f(x|x_i+h_i) - 2f(x) + f(x|x_i-h_i)}{h_i^2}, \quad (8)$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \approx \frac{f(x|x_i+h_i, x_j+h_j) - f(x|x_i+h_i, x_j-h_j) - f(x|x_i-h_i, x_j+h_j) + f(x|x_i-h_i, x_j-h_j)}{4h_i h_j}$$

з кроком $h = 10^{-6}$.

Методи були протестовані на наступних тестових функціях для задач безумовної оптимізації [17].

1. Extended Rosenbrock function:

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

початкове наближення $x^0 = (-1.2, 1, \dots, -1.2, 1)^T$, точка мінімуму $x^* = (1, 1, \dots, 1)^T$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $\text{rank}(f''(x^*)) = n$, $\text{cond}(f''(x^*)) = 2508.01$.

2. Extended White & Holst function:

$$f(x) = \sum_{i=1}^{n/2} \left[100(x_{2i} - x_{2i-1}^3)^2 + (1 - x_{2i-1})^2 \right],$$

початкове наближення $x^0 = (-1.2, 1, \dots, -1.2, 1)^T$, точка мінімуму $x^* = (1, 1, \dots, 1)^T$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $\text{rank}(f''(x^*)) = n$, $\text{cond}(f''(x^*)) = 10018.01$.

3. Extended Wood function:

$$f(x) = \sum_{i=1}^{n/4} \left[100(x_{4i-3}^2 - x_{4i-2})^2 + (x_{4i-3} - 1)^2 + 90(x_{4i-1}^2 - x_{4i})^2 + (1 - x_{4i-1})^2 + 10.1((x_{4i-2} - 1)^2 + (x_{4i} - 1)^2) + 19.8(x_{4i-2} - 1)(x_{4i} - 1) \right],$$

початкове наближення $x^0 = (-3, -1, -3, -1, \dots, -3, -1, -3, -1)^T$, точка мінімуму $x^* = (1, 1, \dots, 1)^T$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $\text{rank}(f''(x^*)) = n$, $\text{cond}(f''(x^*)) = 1102.78$.

виродженої, якщо її число обумовленості $\text{cond}(H_k)$ більше, ніж 10^5 . Похідні $g^k = f'(x^k)$ та $H_k = f''(x^k)$ обчислювались чисельно за симетричними формулами:

$$\text{cond}(f''(x^*)) = 1397.957.$$

4. Extended Powell function:

$$f(x) = \sum_{i=1}^{n/4} \left[(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4 \right],$$

початкове наближення $x^0 = (3, -1, 0, 1, \dots, 3, -1, 0, 1)^T$, точка мінімуму $x^* = (0, 0, 0, 0, \dots, 0, 0, 0, 0)^T$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $\text{rank}(f''(x^*)) = n/2$.

5. Extended Freudenstein & Roth function:

$$f(x) = \sum_{i=1}^{n/2} (-13 + x_{2i-1} + ((5 - x_{2i})x_{2i} - 2)x_{2i})^2 + (-29 + x_{2i-1} + ((x_{2i} + 1)x_{2i} - 14)x_{2i})^2,$$

початкове наближення $x^0 = (0.5, -2, \dots, 0.5, -2)^T$, точка мінімуму $x^* = (11.4127790, -0.8968053, \dots, 11.4127790, -0.8968053)$, значення цільової функції в точці мінімуму $f(x^*) = \begin{cases} 97.96851 & \text{if } n = 4; \\ 293.9055 & \text{if } n = 12, \end{cases}$

$\text{rank}(f''(x^*)) = n$, $\text{cond}(f''(x^*)) = 1102.78$.

6. Extended Tridiagonal 1 function:

$$f(x) = \sum_{i=1}^{n/2} (x_{2i-1} + x_{2i} - 3)^2 + (x_{2i-1} - x_{2i} + 1)^4,$$

початкове наближення $x^0 = (2, 2, \dots, 2)^T$, точка мінімуму $x^* = (1, 2, \dots, 1, 2)^T$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $\text{rank}(f''(x^*)) = n/2$.

7. FLETCHCR function (CUTE):

$f(x) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i + 1 - x_i^2)^2$, початкове наближення $x^0 = (0, 0, \dots, 0)^T$, точка мінімуму не строга, значення цільової функції в точці мінімуму $f(x^*) = 0$, $rank(f''(x^*)) = n - 1$.

8. My 1 function:

$f(x) = 1000(x_1 - 1000)^2 + 0.001x_2^4 + \sum_{i=3}^n (x_i - i)^2$, початкове наближення $x^0 = (100, 100, \dots, 100)^T$, точка мінімуму $x^* = (1000, 0, 3, 4, \dots, n)$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $rank(f''(x^*)) = n - 1$.

9. My 2 function: $f(x) = x_1^2 + x_1x_2^2 + x_2^4 + \sum_{i=3}^n x_i^2$,

початкове наближення $x^0 = (10, 14, 10, \dots, 10)$, точка мінімуму $x^* = (0, 0, \dots, 0, 0)$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $rank(f''(x^*)) = n - 1$.

10. Mean-square approximation by polynomials:

$$f(x) = \sum_{j=1}^{101} \left[\sum_{i=1}^n x_i \cdot 0.01(j-1)^{i-1} - \sum_{i=1}^n x_i^* \cdot 0.01(j-1)^{i-1} \right]^2,$$

$n = 5$, початкове наближення $x^0 = (2, 2, 2, 2, 2)$, точка мінімуму $x^* = (1, 1, 1, 1, 1)$, значення цільової функції в точці мінімуму $f(x^*) = 0$, $rank(f''(x^*)) = n - 1$.

Функції 4, 6–10 характерні тим, що матриця Гессе $f''(x^*)$ вироджена в точці мінімуму, у інших функцій матриця Гессе погано обумовлена. Чисельні експерименти для функцій 1–9 проводились для розмірності $n = 4$.

Порівняння проводилось з процедурами **optim** (метод 'L-BFGS-B') і **nlm** математичного пакету R 3.6.0 та процедурою **optim** (метод 'Quasi-Newton with BFGS') математичного пакету Scilab 6.0.2. Обчислення проводились з максимально можливою точністю. Результати чисельних експериментів наведені в табл. 1–2, де:

- Funct – номер тестуємої функції;
- n – розмірність задачі;
- n-r – ранг виродження гессіану в точці мінімуму;

Dx – евклідова норма $\tilde{x} - x^*$, де \tilde{x} – отримане процедурою оптимізації наближення розв'язку задачі;

$$Df = |f(\tilde{x}) - f(x^*)|;$$

Nitr – виконана кількість ітерацій;

Nf – виконана кількість обчислень цільової функції;

Ngr – виконана кількість обчислень градієнта цільової функції;

$$NormGr = \text{евклідова норма } f'(\tilde{x});$$

code, message та exitflag – код закінчення обчислень, що повернула процедура оптимізації;

ACM-Z – описаний вище метод (2–6; 7) розв'язку задачі (1);

ACKNM-Z – квазі-ньютонівський варіант описаного вище методу (2–6; 7) розв'язку задачі (1).

В процедурах **ACM21-Z** і **ACKNM21-Z** параметр Nf враховує кількість обчислень цільової функції, що були витрачені на чисельне обчислення гессіана за формулою (8).

В процедурах **ACM21-Z** і **ACKNM21-Z** код закінчення обчислень code приймає значення:

0 – досягнута задана точність по градієнту (задавалося 10^{-20});

1 – досягнута задана точність по аргументу (задавалося 10^{-10});

2 – досягнута задана точність по функції (задавалося 10^{-40}).

В процедурі **optim** (пакет R) код закінчення обчислень message приймає значення:

1 – "ERROR: ABNORMAL TERMINATION IN LNSRCH";

2 – "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH";

3 – "NEW_X".

В процедурі **nlm** код закінчення обчислень code приймає значення:

1 – "relative gradient is close to zero, current";

2 – "iterate is probably solution";

3 – "last global step failed to locate a point".

В процедурі **optim** (пакет Scilab) код закінчення обчислень err приймає значення:

1 – "Norm of projected gradient lower than...";

5 – "Optim stops: maximum number of iterations is reached";

9 – "End of optimization, successful completion".

В табл. 1 подано результати обчислень на тестових функціях 1–10 для процедур **optim** і **nlm** математичного пакету R, а в табл. 2 – результати обчислень для процедури **optim** пакету Scilab та процедури **ACKNM21-Z**. Не дивлячись на те, що процедура **optim** в цілому дає більш кращий результат,

ніж процедура **nlm** (див. табл. 1), але вони обидві програють як процедурі **ACKNM21-Z**, так і процедурі **optim** математичного пакета Scilab, як по точності розв'язання, так і по кількості обчислень цільової функції (див. табл. 2).

Якщо порівнювати процедуру **ACKNM21-Z** з процедурою **optim** математичного пакета Scilab (див. табл. 2), то можна сказати, що процедура **ACKNM21-Z** дозволяє для більшості функцій знай-

ти розв'язок задачі з значно більшою точністю, як по аргументу, так і по функції. При цьому, в тих випадках, де кількість обчислень цільової функції, що знадобилась процедурі **ACKNM-Z**, більша в порівнянні з процедурою **optim** пакета Scilab, то ці зайві обчислення пішли на уточнення розв'язку задачі, бо саме в околиці виродженої точки мінімуму швидкість збігання методів дуже низька.

Таблиця 1

Результати тестування процедур **optim** і **nlm** пакету R

Funct	R optim (L-BFGS-B)									R nlm						
	N	n-r	Dx	Df	Nitr	Nf	Ngr	NormGr	Code	Dx	Df	Nitr	Nf	Ngr	NormGr	Code
1	4	0	6,3E-04	8,0E-08		162	162	5,6E-04	2	6,3E-06	8,0E-12	32			9,3E-07	2
2	4	0	3,3E-03	1,1E-06		16474	16474	2,5E-03	3	2,7E-03	7,2E-07	31			1,3E-09	2
3	4	0	1,8E-05	1,1E-09		129	129	5,2E-04	2	1,3E-07	1,3E-13	33			2,5E-06	2
4	4	2	5,1E-05	6,7E-18		135	135	1,2E-10	1	8,7E-03	6,6E-09	42			7,2E-04	2
5	4	0	1,8E-05	1,7E-10		48	48	2,6E-04	1	1,9E-07	4,3E-14	17			2,2E-08	2
6	4	2	5,7E-11	2,0E-41		39	39	1,1E-15	2	3,7E-04	3,7E-14	30			4,7E-10	1
7	4	1		2,0E-10		51	51	5,3E-04	1		1,3E-13	12			7,8E-06	2
8	4	1	1,3E-08	3,7E-25		81	81	1,2E-12	2	6,5E-05	2,2E-20	81			1,3E-10	1
9	4	1	2,1E-05	1,5E-19		116	116	1,9E-12	2	5,8E-04	8,3E-13	56			7,7E-10	1
10	5	1	4,3E-13	2,0E-25		89	89	8,6E-12	2	4,3E-03	6,8E-09	19			2,4E-05	2

Таблиця 2

Результати тестування процедури **optim** пакету Scilab і процедури **ACKNM21-Z**

Funct	Scilab optim (Quasi-Newton with BFGS)									ACKNM21-Z						
	N	n-r	Dx	Df	Nitr	Nf	Ngr	NormGr	err	Dx	Df	Nitr	Nf	Ngr	NormGr	Code
1	4	0	2,3E-08	1,1E-16	48	71		7,4E-11	9	9,1E-11	3,8E-21	37	158	38	5,8E-08	1
2	4	0	1,0E-07	1,6E-15	97	1774		1,3E-06	9	4,5E-10	6,7E-20	37	178	38	2,5E-07	1
3	4	0	5,5E-10	1,3E-18	74	122		3,5E-09	9	1,9E-12	2,0E-24	45	190	46	5,4E-08	1
4	4	2	5,1E-09	6,3E-34	87	113		2,5E-23	9	2,8E-09	6,9E-35	35	356	36	2,5E-17	1
5	4	0	5,9E-09	9,9E-14	19	115		1,5E-08	9	2,8E-08	8,5E-14	9	50	10	2,1E-07	2
6	4	2	1,6E-08	1,3E-31	49	179		4,4E-16	9	2,7E-13	1,7E-50	17	168	18	9,1E-22	0
7	4	1		6,2E-20	17	92		4,4E-09	9		2,2E-26	11	62	12	6,7E-08	1
8	4	1	8,1E-08	4,3E-32	91	196		4,4E-16	9	5,3E-12	8,0E-49	21	272	22	4,4E-16	1
9	4	1	1,1E-21	1,3E-84	213	229		3,3E-43	1	5,9E-11	2,2E-41	43	458	44	2,5E-20	1
10	5	1	3,7E-13	8,4E-29	29	86		1,8E-14	9	4,1E-14	5,4E-29	9	118	10	1,2E-14	1

Висновки

Представлено комбінований метод другого порядку для розв'язання вироджених задач безумовної оптимізації, заснований на ортогональному розкладанні гессіану та поділі всього простору на два ортогональних підпростори.

На одному підпросторі (ядрі матриці Гессе) працює градієнтний метод, а на ортогональному доповненні до нього – метод Ньютона. Було розглянуто також квазі-ньютонівський варіант представле-

ного методу. Ефективність обох методів підтверджується чисельними експериментами, які були проведені на загальноприйнятих тестових функціях для задач безумовної оптимізації. Потрібно відзначити, що запропоновані методи дозволяють отримати досить точні розв'язки тестових завдань особливо у разі виродження точки мінімуму. Це дає сподівання на те, що і при розв'язанні практичних задач саме квазі-ньютонівський варіант представленого методу буде дуже ефективним.

Список літератури

1. Ring W. Optimization methods on Riemannian manifolds and their application to shape space / W. Ring, B. Wirth // *SIAM Journal on Optimization*. – 2012. – Vol. 22, № 2. – P. 596-627.
2. Maratos N.G. Some results on the Sign recurrent neural network for unconstrained minimization / N.G. Maratos, M.A. Moraitis // *Neurocomputing*. – 2017. – Vol. 287. – P. 1-21. <https://doi.org/10.1016/j.neucom.2017.09.036>.
3. Белаш К.Н. Методы решения вырожденных задач / К.Н. Белаш, А.А.Третьяков // *Журнал вычислительной математики и математической физики*. – 1988. – Т. 28, № 7. – С. 1097-1102.
4. Szczepanik E. The p-Factor Method for Nonlinear Optimization / E. Szczepanik, A. Prusinska, A. Tret'yakov // *Schedae Informaticae*. – 2012. – № 21. – P. 141-157.
5. Avriel M. *Nonlinear Programming: Analysis and Methods* / M. Avriel. – Dover Publishing, 2003 (ISBN 0-486-43227-0).
6. Regularized newton methods for convex minimization problems with singular solutions / D.H. Li, M. Fukushima, L. Qi, N. Yamashita // *Comput. Optim. Appl.* – 2004. – Vol. 28. – P. 131-147.
7. Shen C. A regularized Newton method for degenerate unconstrained optimization problems / C. Shen, X. Chen, Y. Liang // *Optimization Letters*. – 2012. – № 6. – P. 1913-1933.
8. Li Q. A Modified Fletcher-Reeves-Type Method for Nonsmooth Convex Minimization / Q. Li // *Stat., Optim. Inf. Comput.* – 2014. – Vol. 2. – P. 200-210.
9. Graspa T.N. A modified Newton direction for unconstrained optimization / T.N. Graspa // *Optimization*. – 2014. – Vol. 63(7). – P. 983-1004.
10. Taheri S. Globally convergent algorithms for solving unconstrained optimization problems / S. Taheri, M. Mammadov, S. Seifollahi // *Optimization*. – 2015. – Vol. 64(2). – P. 249-263.
11. Свириденко А.Б. Взаимосвязь и реализация квазиньютоновских и ньютоновских методов безусловной оптимизации / А.Б. Свириденко, Г.А. Зеленков // *Компьютерные исследования и моделирование*. – 2016. – Т. 8, № 1. – С. 55-78. http://crm.ics.org.ru/uploads/crmissues/crm_2016_1/16.0805.pdf.
12. Han X. A new hybrid conjugate gradient algorithm for unconstrained optimization / X. Han, J. Zhang, J. Chen // *Bulletin of the Iranian Mathematical Society*. – 2017. – Vol. 43(6). – P. 2067-2084.
13. Xiangrong Li. A modified nonmonotone BFGS algorithm for unconstrained optimization / Xiangrong Li, Bopeng Wang, Wujie Hu // *Journal of Inequalities and Applications*. – 2017. – № 183. <https://doi.org/10.1186/s13660-017-1453-5>.
14. Newton-SOR Iteration for Solving Large-Scale Unconstrained Optimization Problems with an Arrowhead Hessian Matrices / K. Ghazali, J. Sulaiman, Y. Dasril, D. Gabda // *Journal of Physics: Conference Series*. – 2019. – Vol. 1358, № 1. – P. 1-10.
15. Niri T.D. An efficient improvement of the Newton method for solving nonconvex optimization problems / T.D. Niri, M.M. Hosseini, M. Heydari // *Computational Methods for Differential Equations*. – 2019. – Vol. 7, № 1. – P. 69-85.
16. Задачин В.М. О скорости сходимости градиентного метода при вырожденном минимуме / В.М. Задачин // *Кибернетика*. – 1989. – № 1. – С. 99-101.
17. Andrei N. An Unconstrained Optimization Test Functions Collection / N. Andrei // *Advanced Modeling and Optimization*. – 2008. – Vol. 10, № 1. – P. 147-161.

References

1. Ring, W. and Wirth, B. (2012), Optimization methods on Riemannian manifolds and their application to shape space, *SIAM Journal on Optimization*, Vol. 22, No. 2, pp. 596-627.
2. Maratos, N.G. and Moraitis, M.A. (2017), Some results on the Sign recurrent neural network for unconstrained minimization, *Neurocomputing*, Vol. 287, pp. 1-21. <https://doi.org/10.1016/j.neucom.2017.09.036>.
3. Belash, K.N. and Tret'yakov, A.A. (1988), "Metodu reshenia vyrozhdennykh zadach" [Methods for solving degenerate problems], *Journal of Computational Mathematics and Mathematical Physics*, Vol. 28, No. 7, pp. 1097-1102.
4. Szczepanik, E., Prusinska, A. and Tret'yakov, A. (2012), The p-Factor Method for Nonlinear Optimization, *Schedae Informaticae*, No. 21, pp. 141-157.
5. Avriel, M. (2003), *Nonlinear Programming: Analysis and Methods*, Dover Publishing (ISBN 0-486-43227-0).
6. Li, D.H., Fukushima, M., Qi, L. and Yamashita, N. (2004), Regularized newton methods for convex minimization problems with singular solutions, *Comput. Optim. Appl.*, Vol. 28, pp. 131-147.
7. Shen, C., Chen, X. and Liang, Y. (2012), A regularized Newton method for degenerate unconstrained optimization problems, *Optimization Letters*, No. 6, pp. 1913-1933.
8. Li, Qiong (2014), A Modified Fletcher-Reeves-Type Method for Nonsmooth Convex Minimization, *Stat., Optim. Inf. Comput.*, Vol. 2, pp. 200-210.
9. Graspa, T.N. (2014), A modified Newton direction for unconstrained optimization, *Optimization*, Vol. 63(7), pp. 983-1004.
10. Taheri, S., Mammadov, M. and Seifollahi, S. (2015), Globally convergent algorithms for solving unconstrained optimization problems, *Optimization*, Vol. 64(2), pp. 249-263.
11. Sveredenko, A.B. and Zelenkov, G.A. (2016), "Vzaimosvjazj i realizacija kvazinjutonovskikh i njutonovskikh metodov bezuslovnoj optimizacii" [Interconnection and implementation of quasi-Newton and Newton methods of unconditional optimization], *Computer research and modeling*, Vol. 8, No. 1, pp. 55-78, available at: www.crm.ics.org.ru/uploads/crmissues/crm_2016_1/16.0805.pdf.

12. Han, X., Zhang, J. and Chen, J. (2017), A new hybrid conjugate gradient algorithm for unconstrained optimization, *Bulletin of the Iranian Mathematical Society*, Vol. 43 (6), pp. 2067-2084.
13. Xiangrong, Li, Bopeng, Wang and Wujie, Hu (2017), A modified nonmonotone BFGS algorithm for unconstrained optimization, *Journal of Inequalities and Applications*, No. 183. <https://doi.org/10.1186/s13660-017-1453-5>.
14. Ghazali, K., Sulaiman, J., Dasril, Y. and Gabda, D. (2019), Newton-SOR Iteration for Solving Large-Scale Unconstrained Optimization Problems with an Arrowhead Hessian Matrices, *Journal of Physics: Conference Series*, Vol. 1358, No. 1, pp. 1-10.
15. Niri, Tayebah Dehghan, Hosseini, Mohammad Mehdi and Heydari, Mohammad (2019), An efficient improvement of the Newton method for solving nonconvex optimization problems, *Computational Methods for Differential Equations*, Vol. 7, No. 1, pp. 69-85.
16. Zadachyn, V.M. (1989), "O skorosti skhodimosti gradientnogo metoda pri vyrozhdennom minimume" [On the rate of convergence of the gradient method with a degenerate minimum], *USSR Cybernetics* (ISSN 0023-1274), No. 1, pp. 99-101.
17. Andrei, N. (2008), An Unconstrained Optimization Test Functions Collection, *Advanced Modeling and Optimization*, Vol. 10, No.1, pp. 147-161.

Надійшла до редколегії 07.02.2020

Схвалена до друку 10.03.2020

Відомості про автора:

Задачин Віктор Михайлович

кандидат фізико-математичних наук доцент
доцент кафедри Харківського національного
економічного університету ім. С. Кузнеця,
Харків, Україна
<https://orcid.org/0000-0002-8107-4639>

Information about the author:

Viktor Zadachyn

Candidate of Physics and Mathematics
Associate Professor of Simon Kuznets Kharkiv
National University of Economics,
Kharkiv, Ukraine
<https://orcid.org/0000-0002-8107-4639>

КОМБИНИРОВАННЫЙ МЕТОД ДЛЯ РЕШЕНИЯ ВЫРОЖДЕННЫХ ЗАДАЧ БЕЗУСЛОВНОЙ ОПТИМИЗАЦИИ

В.М. Задачин

Представлен метод второго порядка для решения вырожденных задач безусловной оптимизации, который является комбинацией двух методов: метода Ньютона и градиентного метода. На каждой итерации все пространство представляется как декартово произведение двух подпространств: ядра матрицы Гессе и ортогонального дополнения к нему. На ядре матрицы Гессе работает градиентный метод, а на ортогональном дополнении к нему – метод Ньютона. При этом применяется параметр регуляризации численного метода для разделения пространства на два ортогональных подпространства. Рассматривается также квази-ньютоновский вариант представленного метода. Эффективность квази-ньютоновского варианта метода подтверждается численными экспериментами, которые были проведены на общепринятых тестовых функциях для задач безусловной оптимизации.

Ключевые слова: безусловная оптимизация, вырожденная точка минимума, модифицированный метод Ньютона, спектральное разложение матрицы.

COMBINED METHOD FOR SOLVING DEGENERATE UNCONSTRAINED OPTIMIZATION PROBLEMS

V. Zadachyn

When solving the practical problems, the extremum point of the chosen optimality criterion is often degenerate that significantly complicates the search. Known numerical methods for solving the general problem of unconditional optimization, up to and including the second order, have a very low convergence rate when solving the degenerate problems. This is because, in order to significantly increase the convergence rate in this case, it is necessary to use in the method the derivatives of higher order than the second one. However, the use of 3rd and 4th order derivatives makes the numerical method very time consuming. In most existing methods, the approach associated with the regularization of the numerical method is applied in solving the degenerate problems.

The article presents a second-order method for solving the degenerate unconditional optimization problems, which is a combination of two methods: the Newton's method and the gradient method. At each iteration, all space is represented as a Cartesian product of two subspaces: the kernel of the Hessian matrix and the orthogonal complement to it. The gradient method works on the kernel of the Hessian matrix, and the Newton's method works on the orthogonal complement to it. In this case, the regularization parameter of the numerical method is used to divide a space into two orthogonal subspaces. The quasi-Newton version of the presented method is also considered. The efficiency of the quasi-Newton version of the method is confirmed by numerical experiments which were conducted on conventional test functions for the problems of unconditional optimization. It gives hope that even when solving the practical problems it is the quasi-Newton version of the presented method that will be very effective.

Keywords: unconditional optimization, degenerate minimum point, modified method Newton, spectral decomposition of matrix.