



2026, № 1

Журнал обчислювальної та
прикладної математики
Заснований в 1965 році



**Journal of Numerical
& Applied Mathematics**

ISSN (Print): 2706-9680, ISSN (Online): 2706-9699

ЄДРПОУ видавця: 02070944

ROR: 02aaqv166

E-mail:
jnameditorialteam@gmail.com

Головний редактор:
Ляшко Сергій Іванович

Сусол К., Ключин Д.

МОДИФІКАЦІЯ МЕТОДУ ШАМАНСЬКОГО

Адреса редакції:
Кафедра обчислювальної математики,
факультет комп'ютерних наук та кібернетики,
Київський національний університет імені Тараса Шевченка,
Україна, 01601, Київ, вул. Володимирська, 64/13.

©Київський національний університет імені Тараса Шевченка, 2026

Свідоцтво про державну реєстрацію КВ 4246 від 26.05.2000
Включено в Перелік наукових фахових видань у відповідності до Наказу МОН
України № 409 від 17.03.2020

DOI: 10.17721/2706-9699.2026.1.06

УДК 519.6

MSC 90C53

A MODIFICATION OF SHAMANSKII'S METHOD

KAROLINA SUSOL, DMYTRO KLYUSHIN

Faculty of Computer Science and Cybernetics,
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,
E-mail: susolcaroline@gmail.com, ORCID: 0009-0004-2118-2824,
E-mail: dmytroklyushin@knu.ua, ORCID: 0000-0003-4554-1049

МОДИФІКАЦІЯ МЕТОДУ ШАМАНСЬКОГО

КАРОЛІНА СУСОЛ, ДМИТРО КЛЮШИН

Факультет комп'ютерних наук та кібернетики,
Київський національний університет імені Тараса Шевченка, Київ, Україна,
E-mail: susolcaroline@gmail.com, ORCID: 0009-0004-2118-2824,
E-mail: dmytroklyushin@knu.ua, ORCID: 0000-0003-4554-1049

ABSTRACT. In this paper, a new method for solving systems of non-linear equations is introduced, obtained by combining Shamanskii's method with line search. Local applicability and global convergence of the new method are proven. Its efficiency is demonstrated on a ten-dimensional test problem.

KEYWORDS: Shamanskii's method, Newton's method, systems of nonlinear equations, line search, iteration methods.

АНОТАЦІЯ. У статті пропонується новий метод для розв'язання систем нелінійних рівнянь, отриманий шляхом поєднання методу Шаманського та лінійного пошуку. Доведено локальну застосовність та глобальну збіжність нового методу. Продемонстровано його ефективність на десятивимірній тестовій задачі.

КЛЮЧОВІ СЛОВА: метод Шаманського, метод Ньютона, системи нелінійних рівнянь, лінійний пошук, ітераційні методи.

1. ВСТУП

Потреба у розв'язанні систем нелінійних рівнянь виникає усюди. Хоч вигляд систем може бути найрізноманітнішим, більшість можна звести до стандартного вигляду:

$$F(x) = 0,$$

Corresponding author: Кароліна Сусол (susolcaroline@gmail.com).

© Кароліна Сусол, Дмитро Ключин, 2026. This is an open-access article distributed under the terms of **Creative Commons Attribution Licence (CC BY)**.

де $F : \mathbb{R}^N \rightarrow \mathbb{R}^M$. Часто $N = M$ (кількість змінних дорівнює кількості рівнянь), надалі ми будемо працювати саме з цим випадком.

Одним із найвідоміших методів для розв'язання задачі пошуку коренів є метод Ньютона-Рафсона (або ж просто метод Ньютона) [1]. Ключовою його перевагою є квадратична швидкість збіжності. Водночас він вимагає обчислення оберненої матриці Якобі або її LU-факторизації на кожній ітерації. Для великих систем, які виникають, приміром, для нелінійних крайових задач на системах звичайних диференціальних рівнянь, обчислення навіть самої матриці Якобі є занадто громіздкою задачею, щоб виконувати її щоразу. Наприклад, такі системи виникають в задачах про взаємодію рідини та конструкцій та в задачах керування деформацією різницевих сіток [2].

Через цей недолік виникли модифікації методу Ньютона, у яких уникають використання безпосередньо матриці F' , а замінюють її на наближення. Такими модифікаціями є метод Бroyдена [3, 4], а також методи BFGS, DFP, хорд тощо [6, 7].

В методі хорд матриці Якобі використовується її значення на нульовому наближенні, умовно припускаючи, що $F'(x_k) \approx F'(x_0)$ [7]. Таким чином, значно полегшується обчислення, але метод теж має лише q -лінійну збіжність.

Тим часом між хордовим методом та методом Ньютона існує компроміс: чергувати ітерацію методом Ньютона з декількома ітераціями методом хорд. Тобто використовувати матрицю Якобі, отриману на певній ітерації, ще декілька наступних ітерацій. Цей компроміс полегшує обчислювальне ядро і водночас ми не втрачаємо високу швидкість збіжності: алгоритм залишається у класі алгоритмів з q -порядком збіжності ≥ 2 .

Цю ідею запропонував український математик Володимир Євтихійович Шаманський. У своїй статті, опублікованій 1967 року, він довів локальну збіжність та установив її швидкість у термінах r -збіжності [8]. Частота оновлення матриці Якобі може бути нерегулярною, і саме такий загальний випадок розглядався в статті, однак у літературі здебільшого під методом Шаманського розуміють метод з оновленням матриці Якобі кожні m кроків (параметр оновлення позначають здебільшого m чи p). Метод залишається актуальним досі, одна з нещодавніх модифікацій методу Шаманського з'явилася у 2022 році [9]. Всього відомо якнайменше шість модифікацій методу.

Іншим недоліком методу Ньютона є відсутність глобальної збіжності для достатньо широких класів функцій. Очевидно, що локальність є невід'ємною характеристикою методу Ньютона, адже він ґрунтується на лінеаризації, яка є достатньо точною лише у певному околі точки, відносно якої будується лінійне наближення до функції. Звісно, існують теореми про глобальну збіжність для опуклих функцій з деякими додатковими обмеженнями, однак цей клас функцій настільки невеликий, що можна стверджувати, що метод Ньютона не є глобально збіжним. Ця відсутність глобальної збіжності зберігається і для методу Шаманського.

Однак є певні техніки, які дозволяють все ж її досягти. Однією із таких технік є лінійний пошук. Поєднання лінійного пошуку з методом Шаманського уже зустрічається у певних статтях, здебільшого це стосується саме задач оптимізації. У цій роботі увага зосереджена на задачі пошуку коренів.

2. ПОСТАНОВКА ЗАДАЧІ ТА ДОПОМІЖНІ ВІДОМОСТІ

У статті використовуються поняття r - та q -збіжності в розумінні С. Kelley [7].

Нехай є нелінійне рівняння $F(x) = 0$, де $F : X \rightarrow Y$ функція з банахового простору X у банахів простір Y . Тоді метод Ньютона-Рафсона для пошуку коренів можна сформулювати наступним чином:

$$x_{k+1} = x_k - [F'(x_k)]^{-1}F(x_k) \quad (1)$$

$$k = 0, 1, 2, \dots,$$

де F' — похідна Фреше функції F . Надалі, де нічого іншого не вказано, вважатимемо, що X, Y є евклідовими просторами та $X = Y$.

Наведемо теорему про локальну збіжність методу Ньютона [10]:

Теорема 1. Розглянемо відображення $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ і нехай $\varepsilon > 0$ та x^* такі, що F є неперервно диференційовною на кулі $B(x^*, \varepsilon)$, $F(x^*) = 0$ і $F'(x^*)$ є невивродженою. Тоді існує таке $\delta > 0$, що якщо $x \in B(x^*, \delta)$, то послідовність x_k згенерована за формулою

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k)$$

є визначеною, збігається до x^* і задовольняє $x_k \in B(x^*, \delta)$ для усіх k . Якщо додатково $x_k \neq x^*$ для усіх k , то:

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

Іншими словами, x_k збігається до x^* q -надлінійно. Також для довільного $r > 0$ існує таке $\delta_r > 0$, що якщо $x_k \in B(x^*, \delta_r)$, то:

$$\|x_{k+1} - x^*\| \leq r\|x_k - x^*\|,$$

$$\|F(x_{k+1})\| \leq r\|F(x_k)\|.$$

Якщо додатково припустити, що для деяких $M > 0, L > 0$ виконується

$$\|F'(x) - F'(y)\| \leq L\|x - y\| \quad \forall x, y \in B(x^*, \varepsilon),$$

$$\|F'(x)\| \leq M, \quad \forall x \in B(x^*, \varepsilon),$$

то

$$\|x_{k+1} - x^*\| \leq \frac{1}{2}LM\|x_k - x^*\|^2 \quad \forall k = 0, 1, \dots,$$

і x_k збігається до x^* q -надлінійно із порядком мінімум 2.

Нагадаємо, що лінійний пошук полягає у заміні довільного ітераційного методу вигляду:

$$x_{k+1} = x_k + p_k$$

на

$$x_{k+1} = x_k + \alpha_k p_k.$$

Тобто вводиться скалярний параметр α_k , котрий регулює довжину кроку.

Одним із найпопулярніших видів лінійного пошуку є зворотний лінійний пошук:

ALGORITHM 1. Зворотний лінійний пошук

Choose $\bar{\alpha} > 0$, $\rho \in (0, 1)$, $c \in (0, 1)$; Set $\alpha \leftarrow \bar{\alpha}$

repeat

$\alpha \leftarrow \rho\alpha$;

until $f(x_k + \alpha p_k) \leq f(x_k) + c\alpha \nabla f(x_k)^T p_k$

Terminate with $\alpha_k = \alpha$.

У алгоритмі в якості умови зупинки використовується умова:

$$f(x_k + \alpha p_k) \leq f(x_k) + c\alpha \nabla f(x_k)^T p_k. \quad (2)$$

Цю рівність називають умовою чи правилом Арміхо.

Нагадаємо означення диференційовності за Гато [6].

Означення 1. Відображення $F : D \subset \mathbb{R}^N \rightarrow \mathbb{R}^M$ називається диференційовним за Гато у точці $x \in \text{int}(D)$, якщо існує лінійний оператор $A \in L(\mathbb{R}^N, \mathbb{R}^M)$ такий, що $\forall h \in \mathbb{R}^N$ виконується:

$$\lim_{t \rightarrow 0} \frac{\|F(x + th) - F(x) - tA(h)\|}{t} = 0. \quad (3)$$

Відомо, що якщо F є диференційовним за Гато, то оператор A , який задовольняє рівність у визначенні, є єдиним. І такий оператор називають похідною Гато функції F . Якщо представити похідну Гато функції F у вигляді матриці, то це буде матриця Якобі.

Наостанок, наведемо алгоритм Шаманського за [7]:

Нехай x_c поточне наближення. Перехід до наступного наближення x_+ задається формулами нижче (m — фіксоване натуральне число).

$$\begin{aligned} y_1 &= x_c - F'(x_c)^{-1} F(x_c) \\ y_{j+1} &= y_j - F'(x_c)^{-1} F(y_j), \quad j = 1, 2, \dots, m-1 \\ x_+ &= y_m. \end{aligned} \quad (4)$$

У науковій літературі наявні різні способи змінити алгоритм Шаманського задля отримання нового алгоритму, зокрема, поєднання його з такими методами:

1. Лінійний пошук. Такий підхід використовують у статтях [11] та [12], де розглядається задача оптимізації.

2. Метод довірчої області. Цей підхід розглядається у статтях [13] та [14, 15]. У останній використовується також і техніка лінійного пошуку.
3. Кубічна та квадратична регуляризація. Цей підхід описаний в статті [9].
4. Прискорення за рахунок заміни $F(x) = 0$ на $F(x)/F'(x)$. Цей підхід розвинений у статті [16].

Лінійний пошук є цінним за рахунок того, що він часто допомагає забезпечити *монотонність* (спадання $f(x_k)$ у випадку мінімізації та спадання $\|F(x_k)\|$ у випадку пошуку коренів) та *глобальну збіжність* для методів, які були лише локально збіжними.

3. Виведення умови зупинки

Лема 1. *Нехай функція $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ є диференційовною за Гато у точці $x \in \text{int}(D)$ і нехай для деякого $p \in \mathbb{R}^n$ виконується $f'(x)^T p > 0$. Тоді $\forall c_1 \in (0, 1)$ існує таке $\delta > 0$, що $\forall \alpha \in (0, \delta)$*

$$f(x - \alpha p) < f(x) - c_1 \alpha f'(x)^T p.$$

Доведення. З рівності

$$\lim_{\alpha \rightarrow 0} \frac{f(x - \alpha p) - f(x)}{\alpha} + (f'(x), p) = 0$$

впливає, що для довільного фіксованого c_1 на проміжку $(0, 1)$ та для довільного фіксованого $c_2 \in (1, +\infty)$ для достатньо малих α виконується умова:

$$\frac{f(x - \alpha p) - f(x)}{\alpha} \in (-c_2 L, -c_1 L),$$

де $L = f'(x)^T p = (f'(x), p) > 0$. Тому $\exists \delta > 0$ таке, що $\forall \alpha \in (0, \delta)$ має місце ланцюжок:

$$-c_2 \alpha L < f(x - \alpha p) - f(x) < -c_1 \alpha L.$$

Тому

$$f(x - \alpha p) < f(x) - c_1 \alpha f'(x)^T p$$

і

$$f(x) - c_2 \alpha f'(x)^T p < f(x - \alpha p).$$

Лема доведена. $\square$

Наслідком наведеної вище леми є те, що зворотний лінійний пошук з правилом Арміхо в якості умови зупинки буде закінчуватися за скінченну кількість кроків.

Покажемо, що ми можемо застосувати цю теорему до $f(x) = (F(x), F(x))$, та кроку $-p = -F'(x)^{-1} F(x)$ (це крок Ньютона для функції F).

Лема 2. *Нехай функція $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ є диференційовною за Гато у точці x (і $F'(x)$ — похідна F за Гато, або точніше її матричне зображення, яке є матрицею Якобі). Введемо функціонал $f(x) = (F(x), F(x)) : \mathbb{R}^N \rightarrow \mathbb{R}$.*

Тоді f є теж є диференційовною функцією за Гато в точці x і її похідна f'_G знаходиться за формулою:

$$f'_G = 2F(x)^T F'(x).$$

Доведення. Почнемо з того, що знайдемо границю $l = \lim_{t \rightarrow 0} \frac{f(x+th) - f(x)}{t}$,

$$l = \lim_{t \rightarrow 0} \frac{(F(x+th), F(x+th)) - (F(x), F(x))}{t}$$

$$l = l_1 + l_2,$$

де

$$l_1 = \lim_{t \rightarrow 0} \frac{(F(x+th), F(x+th)) - (F(x+th), F(x))}{t}$$

та

$$l_2 = \lim_{t \rightarrow 0} \frac{(F(x+th), F(x)) - (F(x), F(x))}{t}.$$

Ми покажемо, що l_1 та l_2 існують, як границі.

$$l_1 = \lim_{t \rightarrow 0} (F(x+th), \frac{F(x+th) - F(x)}{t})$$

Оскільки скалярний добуток є неперервним функціоналом по обох аргументах, то

$$l_1 = (F(x), F'(x)h).$$

Далі знаходимо

$$l_2 = \lim_{t \rightarrow 0} (\frac{F(x+th) - F(x)}{t}, F(x)),$$

$$l_2 = (F'(x)h, F(x)).$$

Оскільки $\mathbb{R}^N, \mathbb{R}$ є дійсними гільбертовими просторами, то скалярний добуток є симетричним і $l_1 = l_2$. Отже,

$$l = l_1 + l_2 = 2(F(x), F'(x)h) = 2F(x)^T F'(x)h.$$

Таким чином ми знайшли лінійний оператор $A : \mathbb{R}^N \rightarrow \mathbb{R}$ такий, що:

$$\lim_{t \rightarrow 0} \frac{f(x+th) - f(x)}{t} = A(h).$$

Тому f є диференційовною за Гато з похідною $A : \mathbb{R}^N \rightarrow \mathbb{R}$,

$$(A(h) = 2F(x)^T F'(x)h).$$

У матричному зображенні: $f'_G(x) = 2F(x)^T F'(x)$. □

Зауваження 1. f'_G позначає похідну Гато функції, записану у рядок, а f' позначає похідну, записану у стовпець (градієнт).

Перед застосуванням теореми 1, потрібно забезпечити виконання вказаних там умов:

1. f повинна бути визначена у околі x , що еквівалентно визначеності F у околі x .

2. Для диференційовності f буде достатньо диференційовності функції F (а диференційовність F — це одне зі стандартних припущень, адже ми її похідну використовуємо у самій же формулі переходу між ітераціями).
3. Залишається умова $f'(x)^T p > 0$.

$$f'(x)^T p = f'(x)^T F'(x)^{-1} F(x)$$

Врахувавши, що $f'(x) = 2F(x)^T F'(x)$, маємо:

$$f'(x)^T p = 2F(x)^T F'(x) F'(x)^{-1} F(x) = 2F(x)^T F(x) = 2\|F(x)\|^2,$$

$$2\|F(x)\|^2 > 0 \text{ за умови, що } F(x) \neq 0.$$

Таким чином умови леми 1 виконуються, якщо у точці x_c визначена ітерація Ньютона (тобто F є диференційовною, а $F'(x_c)$ є невиродженою), і x_c не є розв'язком рівняння $F(x) = 0$.

Із леми 1 випливає: $\forall c_1 \in (0, 1)$ існує таке $\delta > 0$, що

$$\|F(x - \alpha p)\|^2 = f(x - \alpha p) < f(x) - c_1 \alpha f'(x)^T p = \|F(x)\|^2 - c_1 \alpha (2\|F(x)\|^2),$$

$$\|F(x - \alpha p)\|^2 < (1 - 2c_1 \alpha) \|F(x)\|^2.$$

Теорема 2. Нехай у точці x_k для функції $F : D \rightarrow \mathbb{R}^N$, $\text{int}(x_k) \subseteq D \subseteq \mathbb{R}^N$, визначена ітерація Ньютона і $p_k = -F'(x_k)^{-1} F(x_k)$ — відповідний крок. Також нехай x_k не є коренем рівняння $F(x) = 0$. Тоді для довільного $c \in (0, 2)$ існує таке $0 < \delta < 1/c$, що:

$$\|F(x_{k+1})\| < \sqrt{(1 - c\alpha_k)} \|F(x_k)\|,$$

де $x_{k+1} = x_k + \alpha_k p_k$ та α_k довільне число з проміжку $(0, \delta)$.

Очевидно, що побудована відповідним чином послідовність x_k генеруватиме монотонно спадну послідовність нев'язок $\|F(x_k)\|$, адже:

$$\|F(x_{k+1})\| < \sqrt{(1 - c\alpha_k)} \|F(x_k)\| < \|F(x_k)\|.$$

Підсумовуючи, тепер ми можемо створити новий алгоритм, а саме: обладнати метод Шаманського (в якому кроки визначаються як кроки Ньютона), лінійним пошуком зі зворотним пошуком α аналогічним бектрекінгу наведеному в алгоритмі 1, умовою зупинки якого буде нерівність:

$$\|F(x_{k+1})\| < \sqrt{(1 - c\alpha_k)} \|F(x_k)\|, \quad (5)$$

Випишемо псевдокод, що відповідає нерівності (5):

ALGORITHM 2. Лінійний пошук для методу Ньютона (1)

Parameters: $\rho \in (0, 1)$, $c \in (0, 2)$, $\bar{\alpha} \in (0, 1/c)$, step p_k ; Set $\alpha \leftarrow \bar{\alpha}$

repeat

$\alpha \leftarrow \rho \alpha$;

until $\|F(x_k + \alpha p_k)\| < \sqrt{(1 - c\alpha)} \|F(x_k)\|$

Terminate with $\alpha_k = \alpha$.

4. АЛГОРИТМ

Запропонуємо алгоритм:

$$x_{k+1} = x_k - \alpha_k J_k F(x_k),$$

де $J_k = J_{ip+j} = F'(x_{ip})^{-1}$, $j = 0, 1, \dots, p-1$, $i = 0, 1, \dots$, а α_k обирається за алгоритмом 2, який ми спершу вивели суто для методу Ньютона. Назвемо його SM1, скорочення від Shamanskii Modification 1.

ALGORITHM 3. Модифікація методу Шаманського SM1

Step 0. Choose $x_0 \in \mathbb{R}^N$, $\rho \in (0, 1)$, $c \in (0, 2)$, $\bar{\alpha} \in (0, 1/c)$, a positive integer p .
 Step 1. Set $k := 0$, $i := 0$.
 Step 2.
if $k = ip$ **then**
 Compute $F'(x)^{-1}$ and set $J_k := F'(x)^{-1}$. Set $i := i + 1$.
else
 Set $J_k := J_{k-1}$, $i := i + 1$.
end if
 Step 3. Compute d_k by solving $J_k d_k = -F(x_k)$.
 Step 4. Set $x_{k+1} = x_k + \alpha_k d_k$, where $\alpha_k = \text{Algorithm}[2](\rho, c, \bar{\alpha}, \text{step } d_k)$.
 Step 5. $k := k + 1$. Go to step 2.

Зауваження 2. Те, чи ставити в умові зупинки знак $\leq$ чи $<$, має вирішувати розробник. На теоретичні результати це не впливає.

5. ТЕОРЕМИ ПРО ЗАСТОСОВНІСТЬ ТА ЗБІЖНІСТЬ

Для доведення теоретичних результатів нам знадобиться наступна лема.

Лема 3. Нехай ϵ диференційовна за Гато функція $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$, $F(x_c) \neq 0$. Припустимо, що для матриці $F'(x_c)$ та певної матриці J виконується:

$$\|F'(x_c)J - I\| < 1,$$

де I — одинична матриця. Позначимо $E := F'(x_c)J - I$, $e := \|E\|$. Тоді крок $d = -JF(x_c)$ утворює з антиградієнтом функції $f(x) = (F(x), F(x))$ гострий кут і виконується:

$$(-f'(x_c), d) \geq 2(1 - e)\|F(x_c)\|^2.$$

Зауваження 3. Формулювання теореми можна узагальнити на випадок, коли областю визначення є певна множина D , але тоді слід накласти умову, що розв'язок і кожне наближення, отримане внаслідок ітераційного процесу, лежать у внутрішності D .

Доведення. Скориставшись лемою 2, запишемо, чому дорівнює $(-f'(x_c), d)$:

$$\begin{aligned} (-f'(x_c), d) &= 2F(x_c)^T F'(x_c) J F(x_c) = \\ &= 2F(x_c)^T (I + E) F(x_c) = 2\|F(x_c)\|^2 + 2(F(x_c), EF(x_c)). \end{aligned}$$

За нерівністю Коші $|(a, b)| \leq \|a\| \cdot \|b\|$:

$$(-f'(x_c), d) \geq 2\|F(x_c)\|^2 - 2\|F(x_c)\| \cdot \|EF(x_c)\|.$$

Враховуючи, що $\|EF(x_c)\| \leq \|E\| \cdot \|F(x_c)\|$, а $\|E\| < 1$ за умовою, маємо:

$$(-f'(x_c), d) \geq 2\|F(x_c)\|^2 - 2\|E\| \cdot \|F(x_c)\|^2 > 2\|F(x_c)\|^2 - 2\|F(x_c)\|^2 = 0.$$

Отже, дійсно крок спрямований проти градієнта і виконується:

$$(-f'(x_c), d) \geq 2(1 - e)\|F(x_c)\|^2,$$

що і треба було довести. $\square$

Наступним результатом буде наслідок про те, що нерівність (5) можна застосовувати у лінійному пошуку для методу подібного до методу Ньютона, де замість оберненої матриці Якобі використовуються її наближення.

Лема 4. *Нехай e диференційовна за Гато функція $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$, $F(x_c) \neq 0$. Припустимо, що для матриці $F'(x_c)$ та певної матриці J :*

$$\|F'(x_c)J - I\| < 1, \quad (6)$$

де I — одинична матриця, тобто виконуються умови попередньої леми.) Введемо позначення: $E := F'(x_c)J - I$, $e := \|E\|$, $d := -JF(x_c)$. Тоді $\forall c \in (0, 2)$ існує таке $\delta > 0$, що $\forall \alpha \in (0, \delta)$ виконується:

$$\|F(x_c + \alpha d)\|^2 < (1 - c\alpha(1 - e))\|F(x_c)\|^2.$$

Доведення. Використаємо теорему 1, адже внаслідок леми вище крок d і антиградієнт функції $f(x) = (F(x), F(x))$ у точці x_c утворюють гострий кут.

$\forall c_1 \in (0, 1)$ існує таке $\delta > 0$, що $\forall \alpha \in (0, \delta)$:

$$f(x_c + \alpha d) < f(x_c) + c_1 \alpha f'(x_c)^T d.$$

Враховуючи також, що $(f'(x_c), d) \leq -2(1 - e)\|F(x_c)\|^2$, маємо:

$$\|F(x_c + \alpha d)\|^2 < \|F(x_c)\|^2 + c_1 \alpha f'(x_c)^T d \leq \|F(x_c)\|^2 - 2c_1 \alpha (1 - e) \|F(x_c)\|^2,$$

$$\|F(x_c + \alpha d)\|^2 < (1 - 2c_1 \alpha (1 - e)) \|F(x_c)\|^2.$$

Замінивши довільне $c_1 \in (0, 1)$ на довільне $c \in (0, 2)$, отримаємо шукану нерівність. $\square$

Наслідок 1. *Умову зупинки (5) для методу Шаманського бажано застосовувати лише тоді, коли $J_{ip} = F'(x_{ip})^{-1}$ достатньо добре наближує $F'(x_{ip+1})^{-1}, \dots, F'(x_{ip+(p-1)})^{-1}$, тобто коли виконується нерівність (6). Інакше може трапитися так, що алгоритм лінійного пошуку не знайде потрібного α , навіть якщо працюватиме нескінченно довго.*

Тим не менш, якщо говорити про локальну збіжність, то коли точки x_k належать достатньо малому околу розв'язку x^* умова (6) починає автоматично виконуватися, якщо $F'(x)$ є неперервною функцією (а це є також стандартним припущенням на F).

Лема 5. Нехай $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ неперервно диференційовна функція. І нехай $\bar{x} \in \mathbb{R}^N$ — точка, в певному околі якої матриця Якобі F' є невідродженою. Позначимо цей окіл O . Тоді

1. $\forall q > 0$ існує таке $\delta_1 > 0$, що для довільних x, y з кулі $B(\bar{x}, \delta_1) \subseteq O$ має місце:

$$\|F'(x)F'(y)^{-1} - I\| < q.$$

2. $\forall \varepsilon > 0$ існує таке $\delta_2 > 0$, що для довільних x, y з кулі $B(\bar{x}, \delta_2) \subseteq O$ має місце:

$$\|F'(x)^{-1} - F'(y)^{-1}\| < \varepsilon.$$

Доведення. Всередині околу O візьмімо довільний замкнений обмежений окіл Q точки $\bar{x}$: $\bar{x} \in \text{int}Q \subseteq Q \subseteq O$. Функція $\|F'(x)\|$ досягає на Q свого максимального значення, позначимо його W . Функція $F'(x)^{-1}$ на Q є визначеною і неперервною, оскільки $F'(x)$ є неперервною. Оскільки $\mathbb{R}^{N \times N}$ скінченновимірний простір, то Q є також компактом. Тому $F'(x)^{-1}$ є рівномірно неперервною на Q . Тобто

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \|x - y\| < \delta_\varepsilon \Rightarrow \|F'(x)^{-1} - F'(y)^{-1}\| < \varepsilon.$$

Тоді, взявши $\delta_2 = \delta_\varepsilon/2$ отримаємо твердження пункту 2. Для доведення пункту 1 поставимо $\varepsilon = \frac{q}{W}$. Тоді

$$\exists \delta_W > 0 : \|x - y\| < \delta_W \Rightarrow \|F'(x)^{-1} - F'(y)^{-1}\| < \frac{q}{W}.$$

Шукана куля $B(\bar{x}, \delta_1) = B(\bar{x}, \delta_W/2)$, тому що якщо $x, y \in B(\bar{x}, \delta_W/2)$, то $\|x - y\| \leq \|x - \bar{x}\| + \|y - \bar{x}\| < \delta_W/2 + \delta_W/2 = \delta_W$, а звідси:

$$\|F'(x)F'(y)^{-1} - I\| \leq \|F'(x)\| \cdot \|F'(y)^{-1} - F'(x)^{-1}\| \leq W \cdot \frac{q}{W} = q.$$

Отже, пункт 1 доведено теж. $\square$

Для доведення теореми про локальну застосовність нам знадобиться ще одна лема, яка випливає із леми 4:

Лема 6. Нехай є диференційовна за Гато функція $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$, $F(x_c) \neq 0$. А також задано число $c \in (0, 2)$. Тоді існує таке $\hat{e} > 0$, що якщо для матриці $F'(x_c)$ та певної матриці J виконується:

$$e := \|F'(x_c)J - I\| \leq \hat{e}, \tag{7}$$

де I — одинична матриця, то існує таке $\delta > 0$, що $\forall \alpha \in (0, \delta)$ виконується:

$$\|F(x_c + \alpha d)\|^2 < (1 - c\alpha)\|F(x_c)\|^2.$$

Доведення. Нехай $\hat{c}$ довільне число з проміжку $(c, 2)$ (наприклад, $\frac{c+2}{2}$). Тоді $\frac{c}{\hat{c}} \in (0, 1)$, а тому $1 - \frac{c}{\hat{c}} > 0$. Поставимо $\hat{e} = 1 - \frac{c}{\hat{c}}$. Покажемо, що таке $\hat{e}$ задовольняє лему. За результатом леми 4 маємо, що для $\hat{c}$ існує таке $\delta > 0$, що $\forall \alpha \in (0, \delta)$ виконується:

$$\|F(x_c + \alpha d)\|^2 < (1 - \hat{c}\alpha(1 - e))\|F(x_c)\|^2.$$

Також $e \leq \hat{e}$, тому $(1 - \hat{c}\alpha(1 - e)) \leq (1 - \hat{c}\alpha(1 - \hat{e})) = (1 - c\alpha)$ і тому:

$$\|F(x_c + \alpha d)\|^2 < (1 - \hat{c}\alpha(1 - e))\|F(x_c)\|^2 \leq (1 - c\alpha)\|F(x_c)\|^2.$$

Лему доведено. $\square$

Зауваження 4. Цей результат можна трошки покращити. Помітимо з доведення, що оскільки $\hat{c}$ було довільним числом з проміжку $(c, 2)$, то в якості $\hat{e}$ може бути довільне число на проміжку $(1 - \frac{c}{\hat{c}}, 1 - \frac{c}{2}) = (0, 1 - \frac{c}{2})$. А отже ми можемо замінити нестрогу нерівність $\|F'(x_c)J - I\| \leq \hat{e}$ для $\hat{e} \in (0, 1 - \frac{c}{2})$ на строгу нерівність $\|F'(x_c)J - I\| < 1 - \frac{c}{2}$.

Наведемо ще одну важливу теорему [6].

Теорема 3. Нехай $F : D \subseteq \mathbb{R}^N \rightarrow \mathbb{R}^N$, диференційовна за Гато на певній опуклій множині $D_0 \subseteq D$. Тоді для довільних $x, y \in D_0$,

$$\|F(x) - F(y)\| \leq \sup_{t \in (0,1)} \|F'(x + t(y - x))\| \cdot \|x - y\|.$$

Наслідок 2. Нехай функція $F : D \subseteq \mathbb{R}^N \rightarrow \mathbb{R}^N$ диференційовна за Гато на певній опуклій множині $D_0 \subseteq D$ і її похідна на цій множині обмежена за нормою певною константою M . І нехай x^* розв'язок рівняння $F(x) = 0$, $x^* \in D_0$. Тоді для довільного $x \in D_0$:

$$\|F(x)\| \leq M\|x - x^*\|.$$

Теорема 4 (Про локальну застосовність алгоритму 3). Нехай $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ неперервно диференційовна функція, і для певного x^* $F(x^*) = 0$ та $F'(x^*)$ є невиродженою. Тоді існує таке $\gamma > 0$, що якщо $x_0 \in B(x^*, \gamma)$, то всі наступні ітерації алгоритму 3 будуть визначеними (відповідний лінійний пошук закінчиться за скінченну кількість кроків) (за довільного $c \in (0, 2)$ та інших параметрів).

Доведення. Оскільки доведення є досить довгим, розіб'ємо його на декілька пунктів.

1. Почнемо з того, що скористаємося теоремою про локальну збіжність методу Ньютона (1). Перший набір умов теореми виконується, тож для довільного $r > 0$ існує таке $\delta_r > 0$, що якщо $x_c \in B(x^*, \delta_r)$, то:

$$\|\hat{x}_c - x^*\| \leq r\|x_c - x^*\|,$$

де $\hat{x}_c = x_c - F'(x_c)F(x_c)$. Зафіксуємо певне довільне $r \in (0, 1/2)$, а також відповідне йому $\delta_r > 0$. Якщо J певна квадратна матриця, то

$$\|x_c - JF(x_c) - x^*\| = \|x_c - (F'(x_c)^{-1} + J - F'(x_c)^{-1})F(x_c) - x^*\|,$$

$$\|x_c - JF(x_c) - x^*\| = \|\hat{x}_c - x^* + (J - F'(x_c)^{-1})F(x_c)\|.$$

Тоді застосувавши нерівність трикутника для норм, нерівність $\|\hat{x}_c - x^*\| \leq r\|x_c - x^*\|$, маємо:

$$\|x_c - JF(x_c) - x^*\| \leq r\|x_c - x^*\| + \|J - F'(x_c)^{-1}\| \cdot \|F(x_c)\|. \quad (8)$$

2. Згідно леми 6, для того, щоб користуватися умовою зупинки (5), достатньо, аби виконувалася нерівність: $\|F'(x_c)J - I\| \leq \hat{e}$ для певного $\hat{e} > 0$. Застосуємо лему 5, поклавши $q = \hat{e}$. За цією лемою існуватиме $\delta_1 > 0$, що для довільних x, y з кулі $B(x^*, \delta_1)$ має місце:

$$\|F'(x)F'(y)^{-1} - I\| < q.$$

Сумарно з цих двох лем випливає, що існує таке $\delta_1 > 0$, що для довільних x, y з кулі $B(x^*, \delta_1)$ лінійний пошук 2 для $x + \alpha d$, де $d = -F'(y)^{-1}F(x)$ закінчиться за скінченну кількість кроків. Отже, для доведення теореми, нам потрібно, аби з того, що $x_k \in B(x^*, \delta_1)$ випливало, що і наступне наближення $x_{k+1} \in B(x^*, \delta_1)$. Однак для цього нам доведеться зробити окіл навколо x^* ще меншим.

3. Нехай $\delta_0 > 0$ таке, що у відповідному околі кореня $B(x^*, \delta_0)$ матриця Якобі є невинродженою. Позначимо $M = \sup_{x \in B(x^*, \delta_0)} F'(x)$. Тоді є застосовним є наслідок 2:

$$\|F(x)\| \leq M\|x - x^*\|.$$

Далі, скористаємося лемою 5. За нею для $\forall \varepsilon > 0$ існує $\delta_2 > 0$ таке, що для довільних x, y з кулі $B(x^*, \delta_2) \subseteq O$ має місце:

$$\|F'(x)^{-1} - F'(y)^{-1}\| < \varepsilon.$$

Зафіксуємо $\varepsilon = \frac{r}{M}$ і відповідне δ_2 .

4. Тепер нарешті складемо разом ті факти, які ми установили. Покладемо $\gamma = \min\{\delta_0, \delta_1, \delta_r, \delta_2\}$. І позначимо $B := B(x^*, \gamma)$. Доведемо твердження нашої теореми методом математичної індукції.

Нехай $x_0, \dots, x_k \in B$. Покажемо, що тоді визначеною є наступна ітерація і $x_{k+1} \in B$. x_{k+1} шукається лінійним пошуком у вигляді:

$$x_{k+1} = x_k - \alpha J_k F(x_k).$$

Оскільки $J_k = F'(x_i)^{-1}$, де x_i одна з попередніх ітерацій, то, по-перше:

$$\|F'(x_k)F'(x_i)^{-1} - I\| < \hat{\varepsilon},$$

оскільки $x_k, x_i \in B(x^*, \delta_1)$. По-друге, оскільки виконується попередня нерівність, то для наперед заданого параметру c є застосовним лінійний пошук 2, який використовується нашим алгоритмом 3. Отже, x_{k+1} є визначеним.

По-третє, помітимо, що x_{k+1} є опуклою лінійною комбінацією точок x_k та $\hat{x}_{k+1} = x_k - J_k F(x_k)$. Тому, довівши, що $\hat{x}_{k+1} \in B$ ми покажемо, що x_{k+1} також належить опуклій множині B (адже B куля) як опукла комбінація точок із B . Оскільки $x_k \in B(x^*, \delta_r)$, за нерівністю (8) маємо:

$$\|\hat{x}_{k+1} - x^*\| = \|x_k - J_k F(x_k) - x^*\| \leq r\|x_k - x^*\| + \|J_k - F'(x_k)^{-1}\| \cdot \|F(x_k)\|.$$

Оскільки $x_k \in B(x^*, \delta_0)$, то $\|F(x_k)\| \leq M\|x_k - x^*\|$. Також $x_k, x_i \in B(x^*, \delta_2)$, тому $\|F'(x_k)^{-1} - F'(x_i)^{-1}\| < \varepsilon = r/M$, отже:

$$\|\hat{x}_{k+1} - x^*\| \leq r\|x_k - x^*\| + (r/M) \cdot (M\|x_k - x^*\|) = 2r\|x_k - x^*\|.$$

Враховуючи, що $r \in (0, 1/2)$, маємо $\|\hat{x}_{k+1} - x^*\| < \|x_k - x^*\|$. Отже, $\hat{x}_{k+1} \in B$, бо знаходиться ще ближче до x^* , аніж x_k . Отже, і $x_{k+1} \in B$. Що і треба було довести. $\square$

Теорема 5 (Глобальна застосовність алгоритму 3 при $p = 1$). *Нехай $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ є диференційовною за Гато функцією, а $x_0 \in \mathbb{R}^N$ довільне початкове наближення. Тоді гарантовано виконується одне з тверджень:*

- усі наступні ітерації $x_k, k = 1, 2, \dots$ алгоритму 3 є визначеними (для довільних установлених параметрів цього алгоритму).
- для певного натурального числа m визначеними будуть ітерації $x_1, \dots, x_m$, де $F(x_m) = 0$ або $F'(x_m)$ вироджена (тобто ми дійдемо до кореня за скінченну кількість кроків чи до точки з виродженою матрицею Якобі).

Доведення. Нехай x_k певна досягнута ітерація алгоритмом 3, тобто для усіх попередніх ітерацій лінійний пошук знайшов потрібне α за скінченну кількість кроків. Якщо $F(x_k) = 0$ чи $F'(x_k)$ вироджена, то це відповідає другому пункту цієї теореми. Якщо ж $F(x_k) \neq 0$ і $F'(x_k)$ не вироджена, то ми можемо використати результат леми 4 з $J = F'(x_k)^{-1}$. Тоді $e = 0$ і тому: Тоді $\forall \epsilon \in (0, 2)$ (тобто і для нашого установленого також) існує таке $\delta > 0$, що $\forall \alpha \in (0, \delta)$ виконується:

$$\|F(x_k + \alpha d)\|^2 < (1 - c\alpha(1 - 0))\|F(x_k)\|^2 = (1 - c\alpha)\|F(x_k)\|^2,$$

де $d = -JF(x_k) = -F'(x_k)^{-1}F(x_k)$. Звідси випливає, що лінійний пошук, що відповідає k -тій ітерації, знайде наступне α_k за скінченну кількість кроків. Тому x_{k+1} є визначеним. Таким чином $\forall k$ таких, що визначеним є наближення x_k виконується одне із двох: або $F(x_k) = 0$ чи $F'(x_k)$ вироджена, або наступне наближення x_{k+1} за алгоритмом 3 є також визначеним. $\square$

Зауваження 5. Всі ці теореми можна послабити, якщо замінити визначеність і неперервну диференційовність на всьому просторі на якусь його підмножину.

Власне кажучи, єдина проблема із використанням нового лінійного пошуку для методу Шаманського із $p > 1$, це те, що обернена матриця Якобі на певній ітерації може погано наближувати обернені матриці Якобі наступних ітерацій. Однак метод можна аналізувати за припущенням, що алгоритм таки відпрацював за скінченну кількість кроків (насправді це навіть слабша умова, аніж те, що усі наближення задовольняють нерівність (6), і тим паче, аніж те, що початкове наближення повинно знаходитись у якомусь достатньо малому околі розв'язку).

Перед записом теореми, яка установить глобальну збіжність алгоритму 3 за певних умов, нам знадобиться теорема про середнє значення для функціоналів. Це класичний результат, його можна також знайти у [6].

Теорема 6 (Про середнє значення для функціоналів). *Нехай функція $f : D \subseteq \mathbb{R}^N \rightarrow \mathbb{R}$ є диференційовною за Гато у кожній точці опуклої множини $D_0 \subseteq D$. Тоді для довільних двох точок $x, y \in D_0$ існує таке $t \in (0, 1)$, що*

$$f(y) - f(x) = f'(x + t(y - x))(y - x).$$

Наслідок 3. *Нехай $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ диференційовна за Гато. Тоді $\forall x, y \in \mathbb{R}^N$ існує таке $t \in (0, 1)$, що:*

$$F(y, y) - F(x, x) = 2F(x + t(y - x))^T F'(x + t(y - x))(y - x).$$

Доведення. Доведення випливає з використання вищенаведеної теореми 6 і леми про похідну 2. За лемою про похідну 2 $f(x) = (F(x), F(x))$ буде

також диференційовною за Гато на $\mathbb{R}^N$. Тому ми можемо використати 6 і записати $F(y, y) - F(x, x) = f'(x + t(y - x))(y - x)$. Використовуючи ще раз лему 2 і розписуючи значення $f'(x + t(y - x))$, отримуємо шукану рівність. $\square$

Також запишемо невелику лему, яка нам теж знадобиться:

Лема 7. *Нехай e є послідовність квадратних невідроджених матриць $M_k \in \mathbb{R}^{N \times N}$, що збігається до певної матриці M . Відомо також, що M_k^{-1} є обмеженими у сукупності, тобто $\exists C > 0$, що $\|M_k^{-1}\| \leq C$. Тоді матриця M теж є невідродженою, існує границя у M_k^{-1} і $M^{-1} = \lim_{k \rightarrow \infty} M_k^{-1}$.*

Доведення. Розглянемо послідовність M_k^{-1} . У неї існує хоча б одна часткова границя, оскільки ця послідовність обмежена, а у скінченновимірних просторах, яким є $\mathbb{R}^{N \times N}$, з обмеженості випливає компактність. Позначимо цю часткову границю L :

$$L = \lim_{k \rightarrow \infty, k \in K_L \subseteq \mathbb{N}} M_k^{-1}.$$

Тоді запишемо границю:

$$\lim_{k \rightarrow \infty, k \in K_L \subseteq \mathbb{N}} M_k M_k^{-1} = \lim_{k \rightarrow \infty, k \in K_L \subseteq \mathbb{N}} M_k \cdot \lim_{k \rightarrow \infty, k \in K_L \subseteq \mathbb{N}} M_k^{-1} = ML.$$

З іншого боку це границя послідовності, що складається з одиничних матриць, тож $ML = I$, а отже M є невідродженою, а $L = M^{-1}$. Оскільки L була довільною частковою границею, звідси випливає, що у M_k^{-1} єдина часткова границя. І оскільки M_k^{-1} обмежені, то ця часткова границя є границею: $\lim_{k \rightarrow \infty} M_k^{-1} = L = M^{-1}$. $\square$

Теорема 7 (Про глобальну збіжність алгоритму 3). *Нехай $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ неперервно диференційовна, $\{x_k\}_{k=1}^\infty$ — ітераційний процес, отриманий за алгоритмом 3. Тобто відомо, що $\forall k = 0, 1, \dots$ лінійний пошук закінчився за скінченну кількість кроків, а серед матриць Якобі $F'(x_k)$ не було жодної виродженої. Припустимо також, що множина наближень $X = \{x_k\}_{k=1}^\infty$ є обмеженою (міститься у певній кулі B), а також що матриці J_k є обмеженими в сукупності. Тоді кожна гранична точка $\{x_k\}_{k=1}^\infty$ є розв'язком рівняння $F(x) = 0$.*

Доведення. Для кожного k виконується нерівність:

$$\|F(x_{k+1})\| = \|F(x_k + \alpha_k d_k)\| < \sqrt{(1 - c\alpha_k)} \|F(x_k)\|.$$

Звідси випливає, що послідовність норм $\|F(x_k)\|$ є монотонно спадною. До того ж вона є обмеженою знизу, оскільки норма є функціоналом, що набуває невід'ємних значень. Отже, існує границя:

$$\lim_{k \rightarrow \infty} \|F(x_k)\| = \beta \in \mathbb{R}^+ = [0, +\infty).$$

Розглянемо два випадки. Перший випадок: $\beta = 0$. Нехай $\tilde{x}$ — гранична точка X , тобто для певної нескінченної множини $\tilde{K} \subseteq \mathbb{N}$ виконується $\lim_{k \rightarrow \infty, k \in \tilde{K}} x_k = \tilde{x}$ (а граничні точки у послідовності наближень є, оскільки

кожна обмежена множина у $\mathbb{R}^N$ є компактною). Тоді з неперервності F випливає, що $\lim_{k \rightarrow \infty, k \in \tilde{K}} \|F(x_k)\| = F(\tilde{x})$. З іншого боку

$$\lim_{k \rightarrow \infty, k \in \tilde{K}} \|F(x_k)\| = \beta.$$

Тож $F(\tilde{x}) = \beta = 0$.

Другий випадок: $\beta \neq 0$. Послідовність α_k є обмеженою, бо належить проміжку $(0, \bar{\alpha}]$. Покажемо, що $\alpha_k \rightarrow 0$, довівши, що кожна часткова границя α_k дорівнює нулю.

Нехай a — її довільна часткова границя (по певній множині $K_a \subseteq \mathbb{N}$). Тоді:

$$\lim_{k \rightarrow \infty, k \in K_a} \|F(x_{k+1})\| \leq \lim_{k \rightarrow \infty, k \in K_a} (\sqrt{(1 - c\alpha_k)} \|F(x_k)\|),$$

$$\lim_{k \rightarrow \infty, k \in K_a} \|F(x_{k+1})\| = \beta,$$

$$\lim_{k \rightarrow \infty, k \in K_a} (\sqrt{(1 - c\alpha_k)} \|F(x_k)\|) = \sqrt{(1 - ca)} \cdot \beta,$$

оскільки:

$$\lim_{k \rightarrow \infty, k \in K_a} \sqrt{(1 - c\alpha_k)} = \sqrt{(1 - ca)}, \quad \lim_{k \rightarrow \infty, k \in K_a} \|F(x_k)\| = \beta.$$

Тоді $\beta \leq \sqrt{(1 - ca)} \cdot \beta \implies \beta(1 - \sqrt{(1 - ca)}) \leq 0$. Однак $\beta > 0$, а $1 - \sqrt{(1 - ca)} \geq 0$, оскільки $c, a \geq 0$. Отже, вказана нерівність можлива лише якщо $1 - \sqrt{(1 - ca)} = 0 \iff ca = 0 \iff a = 0$.

Таким чином, кожна часткова границя α_k дорівнює нулю, а отже вся послідовність збігається до нуля.

З цього випливає те, що $\lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0$, адже:

$$\|x_{k+1} - x_k\| = \|- \alpha_k J_k F(x_k)\| \leq |\alpha_k| \cdot \|J_k\| \cdot \|F(x_k)\|,$$

α_k прямує до нуля, $\|J_k\|$ є обмеженими в сукупності за умовою, а неперервна функція $\|F(x)\|$ досягає свого максимального значення на кулі $B \supset \{x_k\}_{k=1}^\infty$.

Наявність прямування $x_{k+1} - x_k$ до нуля (в межах другого випадку) є дуже важливою властивістю, оскільки звідси випливатиме, що множина часткових границь для $x_{ip}, i = 0, 1, \dots$ дорівнює множині часткових границь для всієї послідовності $x_k, k = 0, 1, \dots$. Позначимо ці множини X_p та $\tilde{X}$ відповідно.

Доведемо цей факт. Зрозуміло, що $X_p \subseteq \tilde{X}$. Тож потрібно лише довести, що $\tilde{X} \subseteq X_p$.

Нехай z довільна часткова границя з $\tilde{X}$. Їй відповідає певна підпослідовність $\{x_k, k \in K_z\}$, $K_z \subseteq \mathbb{N}$. Кожне k можна представити у вигляді $k = i \cdot p + j, j \in \overline{0, p-1}, i \in \mathbb{N} \cup \{0\}$. Множину відповідних i позначимо K_i , формально $K_i = \{[k/p], k \in K_z\}$. Розглянемо послідовність $\{x_{ip} | i \in K_i\}$,

$$\|x_k - x_{ip}\| = \left\| \sum_{l=0}^{j-1} (x_{ip+l+1} - x_{ip+l}) \right\|,$$

де $i = [k/p]$, $j = k - ip$,

$$\|x_k - x_{ip}\| \leq \left\| \sum_{l=0}^{p-1} (x_{ip+l+1} - x_{ip+l}) \right\| \leq \sum_{l=0}^{p-1} \|x_{ip+l+1} - x_{ip+l}\|.$$

Кожен із цих p доданків прямує до нуля, бо $\|x_{k+1} - x_k\| \rightarrow 0$. Отже, і $\|x_k - x_{ip}\| \rightarrow 0$, і тоді:

$$\lim_{i \rightarrow \infty, i \in K_i} x_{ip} = \lim_{k \rightarrow \infty, k \in K_z} x_k = z.$$

Враховуючи також, що $\lim_{i \rightarrow \infty, i \in K_i} x_{ip} \in X_p$, маємо $z \in X_p$. Отже, $X_p = \tilde{X}$.

Залишається довести, що всі часткові границі послідовності точок x_{ip} , $i = 0, 1, 2, \dots$ є коренями рівняння $F(x) = 0$.

Оскільки $\alpha_k \rightarrow 0$, то $\exists A \in \mathbb{N}$:

$$\alpha_k < \bar{\alpha}, \forall k \geq A \leftrightarrow \alpha_k \leq \rho \bar{\alpha}, \forall k \geq A.$$

Тобто у лінійному пошуку відбулося присвоєння α значення $\frac{\alpha_k}{\rho}$, однак воно не підійшло, а це означає, що:

$$\|F(x_k + \frac{\alpha_k}{\rho} d_k)\| \geq \sqrt{1 - c \frac{\alpha_k}{\rho}} \|F(x_k)\|.$$

Ця нерівність еквівалентна наступній:

$$\|F(x_k + \frac{\alpha_k}{\rho} d_k)\|^2 - \|F(x_k)\|^2 \geq -c \frac{\alpha_k}{\rho} \|F(x_k)\|^2. \quad (9)$$

Застосуємо наслідок 3 до F та пари точок x_k та $x_k + \frac{\alpha_k}{\rho} d_k$. Тоді існує певне $t = \theta_k$ таке, що:

$$\|F(x_k + \frac{\alpha_k}{\rho} d_k)\|^2 - \|F(x_k)\|^2 = 2F(x_k + \theta_k \frac{\alpha_k}{\rho} d_k)^T F'(x_k + \theta_k \frac{\alpha_k}{\rho} d_k) \frac{\alpha_k}{\rho} d_k. \quad (10)$$

Об'єднавши рівність (10) з останньою нерівністю (9), отримаємо:

$$2F(x_k + \theta_k \frac{\alpha_k}{\rho} d_k)^T F'(x_k + \theta_k \frac{\alpha_k}{\rho} d_k) \frac{\alpha_k}{\rho} d_k \geq -c \frac{\alpha_k}{\rho} \|F(x_k)\|^2.$$

Оскільки $\frac{\alpha_k}{\rho} > 0$, можемо скоротити нерівність на цей дріб:

$$2F(x_k + \theta_k \frac{\alpha_k}{\rho} d_k)^T F'(x_k + \theta_k \frac{\alpha_k}{\rho} d_k) d_k \geq -c \|F(x_k)\|^2. \quad (11)$$

Тепер розглянемо довільну часткову границю $\hat{x} \in X_p$. Нехай їй відповідає підпослідовність $x_k, k \in \hat{K}$, де $\hat{K}$ є нескінченною підмножиною $K_p = \{ip, i = 0, 1, 2, \dots\}$. Оскільки θ_k обмежена, $\frac{1}{\rho}$ константа, а d_k теж обмежена ($\|d_k\| \leq \|J_k\| \cdot \|F(x_k)\| \leq M = \text{const}$), то:

$$\lim_{k \rightarrow \infty, k \in \hat{K}} x_k + \theta_k \frac{\alpha_k}{\rho} d_k = \lim_{k \rightarrow \infty, k \in \hat{K}} x_k = \hat{x}.$$

Тоді з неперервності F та F' :

$$\lim_{k \rightarrow \infty, k \in \hat{K}} 2F(x_k + \theta_k \frac{\alpha_k}{\rho} d_k)^T F'(x_k + \theta_k \frac{\alpha_k}{\rho} d_k) = 2F(\hat{x})^T F'(\hat{x}).$$

Що стосується d_k , то оскільки $k \in \hat{K} \subseteq K_p$,

$$d_k = -J_k F(x_k) = -F'(x_k)^{-1} F(x_k).$$

Оскільки F' неперервна, то існує $\lim_{k \rightarrow \infty, k \in \hat{K}} F'(x_k) = F'(\hat{x})$. Скористаємося лемою 7, поклавши $M_k = F'(x_k)$ (відповідно, $M = F'(\hat{x})$). Всі умови виконуються (по $k \in \hat{K}$), оскільки за умовою $J_k = M_k^{-1}$ є обмеженими. Тому існує

$$\lim_{k \rightarrow \infty, k \in \hat{K}} F'(x_k)^{-1} = \lim_{k \rightarrow \infty, k \in \hat{K}} M_k^{-1} = M^{-1} = F'(\hat{x})^{-1}.$$

Таким чином $d_k \rightarrow -F'(\hat{x})^{-1} F(\hat{x})$, коли $k \rightarrow \infty, k \in \hat{K}$. Тепер спрямуємо у нерівності (11) $k \rightarrow \infty, k \in \hat{K}$.

$$\begin{aligned} \lim_{k \rightarrow \infty, k \in \hat{K}} 2F(x_k + \theta_k \frac{\alpha_k}{\rho} d_k)^T F'(x_k + \theta_k \frac{\alpha_k}{\rho} d_k) d_k &\geq \lim_{k \rightarrow \infty, k \in \hat{K}} -c \|F(x_k)\|^2, \\ -2F(\hat{x})^T F'(\hat{x}) F'(\hat{x})^{-1} F(\hat{x}) &\geq -c \|F(\hat{x})\|^2, \\ -2F(\hat{x})^T I F(\hat{x}) = -2\|F(\hat{x})\|^2 &\geq -c \|F(\hat{x})\|^2, \\ (c - 2)\|F(\hat{x})\|^2 &\geq 0. \end{aligned}$$

Як було показано, параметр $c \in (0, 2)$, тому це можливо, лише якщо $\|F(\hat{x})\| = 0 \leftrightarrow F(\hat{x}) = 0$.

Таким чином, ми установили, що довільна часткова границя послідовності x_{ip} є розв'язком рівняння $F(x) = 0$. Це і слід було довести, адже множини граничних точок у послідовностей $x_k, k = 0, 1, \dots$ та $x_{ip}, i = 0, 1, \dots$ збігаються ($\tilde{X} = X_p$). $\square$

5. ОБЧИСЛЮВАЛЬНИЙ ЕКСПЕРИМЕНТ

На практиці до ситуації, коли лінійний пошук не досягає умови зупинки за визначену максимальну кількість ітерацій (*MaxItLS*), ми підійдемо кількома способами.

1. Зробити маленький крок, тобто αd , де $\alpha = \bar{\alpha} \rho^{MaxItLS-1}$.
2. Зробити повний крок (використовуючи наближення оберненої матриці Якобі).
3. Зробити повний крок (використовуючи обернену матрицю Якобі).

Відповідні реалізації будемо позначати SM1-0, SM1-1, SM1-2. Також слід зазначити, що у тестуванні замість точної матриці Якобі ми використовуємо її різницеве наближення. Це є найбільш універсальним підходом, оскільки часом функції у системах нелінійних рівнянь не є елементарними, або навіть не представляються у аналітичному вигляді. Також це дає змогу краще оцінити час роботи методів.

Зауваження 6. Ітерацією методу Шаманського та всіх його модифікацій називатимемо послідовність з m ітерацій, які використовують одне і те ж наближення оберненої матриці Якобі. Тобто одному обчисленню $F'(x)^{-1}$ відповідатиме одна ітерація і навпаки.

Будемо проводити тестування для $N = 10$ та функції:

$$F(x) = \begin{pmatrix} \sin(10 - (x[0] - x[1])^2) - 0.4121184852417566, \\ x[1]^3 + 3x[2] - 17, \\ x[3] + x[4] - 9, \\ x[9] - \cos(x[1])^3 - 10.072067555747765, \\ \exp(x[9] - x[5]) - 2x[2] - 48.598150033144236, \\ x[4]^3 - x[4] + \cos(x[3]x[7]x[8] - 1000) - 119.583798343844, \\ x[8]x[7] - 58 - x[6]x[1], \\ x[8] - \sin(x[1])^4 - 8.316365693794271, \\ x[6]^2 + x[7] + \exp(10 - x[7]) - 64.38905609893065, \\ x[6] + x[7] + x[0] + x[1] - 18 \end{pmatrix} \quad (12)$$

Параметри методів залишаються такими самими, як у одновимірному тестуванні. Функція підбрана таким чином, щоб вектор

$$(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)^T$$

був її коренем. Однак зауважимо що, як виявилось під час тестування, у функції є поблизу ще два корені:

$$(2.57, 1.57, 4.37, 4.00, 5.00, 6.02, 6.53, 7.33, 9.32, 10.07)^T,$$

$$(3.08, 1.27, 4.99, 4.02, 4.98, 6.02, 6.42, 7.23, 9.15, 10.10)^T$$

(координати наведено з точністю $1e - 3$).

Для того, щоб установити загальну картину успішності методів, проведемо тестування наступним чином:

Накладемо сітку у 10-вимірному просторі біля одного з розв'язків. Нехай це буде: $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)^T$, а сітку утворимо за рахунок усіх можливих точок, у яких перша координата це 1-step або 1+step, друга: 2-step або 2+step, третя: 3-step або 3+step,..., десята: 10+step або 10-step. Оберімо step=1.

Запускатимемо методи Ньютона, Шаманського та SM1-0, SM1-1, SM1-2 з початкових точок на побудованій сітці (всього 1024 початкових точок). $\{1, 2, 5, 10\}$ — множина значень параметра оновлення матриці Якобі для методів Шаманського та реалізацій SM1.

Умовою зупинки алгоритмів є достатньо мале значення нев'язки:

$$\|F(x_k)\| \leq eps = 1e - 10.$$

Максимальна кількість ітерацій у всіх методів: 200.

Максимальна кількість ітерацій всередині лінійного пошуку: 200.

Параметри модифікації методу Шаманського: $\alpha_0 = 1, c = 1/2, \rho = 1/2$.

Результати тестування:

$$m = 1$$

Успішність методів:

SM (NM): 957/1024 (93.46%)

SM1-0: 996/1024 (97.27%)

SM1-1: 1006/1024 (98.24%)

SM1-2: 1006/1024 (98.24%)

$$m = 2$$

Успішність методів:

SM: 616/1024 (60.16%)

SM1-0: 1022/1024 (99.80%)

SM1-1: 1020/1024 (99.61%)

SM1-2: 1020/1024 (99.61%)

$$m = 5$$

Успішність методів:

SM: 549/1024 (53.61%)

SM1-0: 1018/1024 (99.41%)

SM1-1: 1009/1024 (98.54%)

SM1-2: 1018/1024 (99.41%)

$$m = 10$$

Успішність методів:

SM: 378/1024 (36.91%)

SM1-0: 1020/1024 (99.61%)

SM1-1: 956/1024 (93.36%)

SM1-2: 1017/1024 (99.32%)

У записі $N/1024$ N позначає кількість разів, коли за кількість ітерацій, що не перевищує 200, метод отримував вектор, на якому норма значення функції F має достатньо мале значення (не перевищує $1e - 10$).

6. ВИСНОВКИ

В роботі пропонується поєднати метод Шаманського та лінійний пошук.

- Запропоновано новий метод (SM1), що є модифікацією методу Шаманського, з використанням зворотного лінійного пошуку та виведеної умови зупинки.
- Доведено теореми, які установлюють його застосовність (локальна для $p > 1$ та глобальна для $p = 1$).
- Доведена теорема про глобальну збіжність методу. До того ж припущення, накладені у відповідній теоремі, є відносно слабкими.
- Проведено тестування методу Ньютона, методу Шаманського та нової модифікації методу Шаманського (у трьох варіантах реалізації) для 4-х різних значень параметру, що відповідає за частоту оновлення матриці Якобі:

$$p = m_1 = 1, \quad p = m_2 = 2, \quad p = m_3 = 5, \quad p = m_4 = 10.$$

- Показано, що SM1 є конкурентним методом, який в реалізаціях SM1-0 та SM1-2 продемонстрував кращу успішність, ніж методи Ньютона та Шаманського.

Виходячи з вищесказаного можна стверджувати, що SM1 є алгоритмом, вартим використання і подальших досліджень.

Автори заявляють про відсутність конфлікту інтересів щодо публікації цієї статті.

ЛІТЕРАТУРА

1. Dedieu J.P. Newton-Raphson Method. In: Engquist B. (eds) *Encyclopedia of Applied and Computational Mathematics*. Berlin, Heidelberg: Springer, 2015. doi: 10.1007/978-3-540-70529-1_374
2. Shamanskiy A., Simeon B. Mesh moving techniques in fluid-structure interaction: robustness, accumulated distortion and computational efficiency. *Computational Mechanics*. 2021. Vol. 67, P. 1–18. doi: 10.1007/s00466-020-01950-x.
3. Broyden C.G. A Class of Methods for Solving Nonlinear Simultaneous Equations. *Mathematics of Computation*. 1965. Vol.19, no. 92. P. 577–593. doi: 10.1090/S0025-5718-1965-0198670-6
4. Shakhno S., Yarmola H. On the Newton-Broyden method for solving systems of nonlinear equations. *Journal of Applied and Numerical Analysis*. 2023. Vol. 1. P. 80–87.
5. Nocedal J., Wright S.J. *Numerical Optimization*. Springer New York, NY, 2006.
6. Ortega J.M., Rheinboldt W.C. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, 1970.
7. Kelley C.T. *Iterative Methods for Linear and Nonlinear Equations*. Society for Industrial and Applied Mathematics, 1995.
8. Shamanskii V.E. On a Modification of Newton's Method. *Ukrainskyi Matematychnyi Zhurnal*. 1967. Vol. 19. P. 118–122.
9. Doikov N., Chayti E.M., Jaggi M. Second-order optimization with lazy Hessians. *arXiv*, 2023. doi: 10.48550/arXiv.2212.00781
10. Bertsekas D.P. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.
11. Lampariello F., Sciandrone M. Global Convergence Technique for the Newton Method with Periodic Hessian Evaluation. *Journal of Optimization Theory and Applications*. 2001, Vol. 111. P. 341–358. doi:10.1023/A:1011934418390
12. Wang C., Chen Y., Du S. Further insight into the Shamanskii modification of Newton method. *Applied Mathematics and Computation*. 2006. Vol. 180, no. 1. P. 46–52.
13. Fan J. A Shamanskii-like Levenberg-Marquardt method for nonlinear equations. *Computational Optimization and Applications*. 2013. Vol. 56, no. 1. P. 63–80.
14. Huang B., Ma C. A Shamanskii-like self-adaptive Levenberg–Marquardt method for nonlinear equations *Computers & Mathematics with Applications*. 2019. Vol. 77, no. 2. P. 357–373, doi:10.1016/j.camwa.2018.09.039.
15. Chen L., Ma Y. Shamanskii-Like Levenberg-Marquardt Method with a New Line Search for Systems of Nonlinear Equations. *Journal of Systems Science and Complexity*. 2020. Vol. 33, P. 1694–1707.
16. Thander A., Paul S., Maitra P. An improved Shamanskii method for finding zeros of linear and nonlinear equations. *Applied Mathematical Sciences*. 2012. Vol. 6, no. 86. P. 4277–4281.

Надійшла: 20.01.2026 / Переглянута: 23.02.2026 /
Прийнята: 10.04.2026 / Опублікована: 24.04.2026