УДК 518:517.948

П.С.Сеньо

НОВИЙ ПІДХІД ДО ПОБУДОВИ ІНТЕРВАЛЬНИХ МЕТОДІВ РОЗВ"ЯЗУВАННЯ СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ

Відомо багато методів уточнення коренів систем нелінійних рівнянь. Однак ефективних методів розв"язування таких систем небагато. Майже всі вони, крім методів розв"язування систем спеціальних видів, базуються на ідеях інтервального аналізу.

Основна ідея побудови інтервальних ітераційних методів розв"язування системи нелінійних рівнянь

$$f(x) = 0, \qquad /1/$$

де $f=(f_1,\dots f_n)^T$, $x=(x_1,\dots,x_n)$, проста.

Запишемо $f(x)$ у вигляді

$$f(x)=f(\bar{x})+f'(\xi)(x-\bar{x}). \qquad /2/$$

Тут /2/ при $n>1$ слід розуміти як узагальнення теореми про середнє значення [3] . Нехай $\mathcal{X}_n$ - інтервальний вектор, який містить в собі x і $\bar{x}$. Тоді $\xi \in \mathcal{X}_n$. Замінивши в /2/ $f'(\xi)$ на $f'(\mathcal{X}_n)$, одержимо матричне лінійне рівняння з інтервальними коефіцієнтами

$$f(\bar{x})+f'(\mathcal{X}_n)(z-\bar{x})=0. \qquad /3/$$

Очевидно, розв"язок рівняння /1/ належить множині Z_n , елементи якої $z \in Z_n$ -є розв"язками рівняння /3/. Тому, якщо $x^* \in \mathcal{X}_n$ то $x^* \in Z_n$, де x^* - розв"язок /1/. Далі приймемо $\mathcal{X}_{n+1} = \mathcal{X}_n \cap Z_n$, $(n=0,1,\dots)$.

При $n=1$ описану ідею легко узагальнити для побудови інтервальних ітераційних методів вищих порядків збіжності [1] . Однак при $n>1$ використання аналога теореми Тейлора у термінах узагальненої теореми про середнє значення вимагає розв"язання матричних рівнянь вищих степенів з інтервальними коефіцієнтами, апроксимації їх приводять до $x^* \notin Z_n$. Цим пояснюється використання для розв"язування /1/ лише інтервального метода Ньютона.

Розвиваючи ідеї, започатковані в [4,5], пропонуємо для побудови інтервальних ітераційних методів високих порядків збіжності використовувати поведінку "середніх" точок залишкових членів у "формі Лагранжа" узагальнених рядів Тейлора [3] функції $f(x)$ та співвідношення між цими точками при розкладі в ряди Тейлора функції $f(x)$ і її похідних.

Нехай відображення $f: D \subset R^n \to R^n$ має m - ту похідну Гато в кожній точці випуклої множини $D_0 \subset D$ і $x_1, x_2 \in D_0$. Тоді

$$f(x_2) = f(x_1) + f'(x_1)(x_2 - x_1) + \ldots + \frac{1}{m!} f^{(m)}(x_1 + \theta_m^0 (x_2 - x_1))(x_2 - x_1)^m. \qquad /4/$$

<u>Теорема 1</u>. Якщо існує $(m+1)$, похідна Гато в кожній точці випуклої множини $D_0 \subset D$ і жодна "координата" гіперматриці $f^{(m+1)}(x_2)$ не дорівнює нулю, то

$$\theta_m^0 \xrightarrow[x_1 \to x_2]{} \frac{1}{m+1} \qquad /5/$$

по кожній координаті "вектора" θ_m^0 .

<u>Доведення</u>. В умовах теореми, крім /4/, має місце ще й рівність

$$f(x_2) = f(x_1) + f'(x_1)(x_2 - x_1) + \ldots + \frac{1}{(m+1)!} f^{(m+1)}(x_1 + \theta_{m+1}^0 (x_2 - x_1))(x_2 - x_1)^{m+1}. \qquad /6/$$

Застосуємо в /4/ до функції $f^{(m)}(x)$ узагальнену теорему про середнє значення. Тоді

$$f^{(m)}(x_1 + \theta_m^0 (x_2 - x_1)) = f^{(m)}(x_1) + f^{(m+1)}(x_1 + \theta_1^m \cdot \theta_m^0 (x_2 - x_1)) \cdot \theta_m^0 (x_2 - x_1), \qquad /7/$$

де компоненти гіперматриці $f^{(m+1)}(x_1 + \theta_1^m \cdot \theta_m^0 (x_2 - x_1))$ одержуємо послідовним застосуванням теореми про середнє значення окремо до кожної компоненти гіперматриці $f^{(m)}(x)$. У поняття гіперматриці вкладаємо наступний зміст. При $m=1$

$$f(x_2) = f(x_1) + f'(x_1 + \theta_1^0 (x_2 - x_1))(x_2 - x_1),$$

де

$$f'(x_1 + \theta_1^0 (x_2 - x_1)) = \begin{pmatrix} f_1'(x_1 + \theta_{11}^0 (x_2 - x_1)) \\ \vdots \\ f_n'(x_1 + \theta_{1n}^0 (x_2 - x_1)) \end{pmatrix}. \qquad /8/$$

це матриця Якобі, в якій, взагалі кажучи, всі $\theta^{o}_{1i} \in (0;1)(i=1,\dots,n)$ різні. Тому $f'(x_1+\theta^{o}_1(x_2-x_1))$ не похідна Гато, взята в деякій проміжній точці. Зауважимо, що для всіх елементів кожної окремої стрічки θ^{o}_{1i} однакові. При $m=2$ $f''(x_1+\theta^{o}_2(x_2-x_1))$ – вектор матриць Гессе, де числа $\theta^{o}_{ij} \in (0;1)$ $(i,j=1,\dots,n)$ різні. Отже, $f''(x_1+\theta^{o}_2(x_2-x_1))^{2ij}$ не є другою похідною Гато, взятою в деякій проміжній точці. При $m=3$ $f'''(x_1+\theta^{o}_3(x_2-x_1))$ – це відповідний вектор, "координати" якого є векторами матриць, тобто він "матриця матриць" і т.д.

Підставляючи /7/ в /4/ і прирівнюючи розклад $f(x_2)$ з розкладом /6/, одержуємо

$$\frac{1}{m+1} f^{(m+1)}(x_1+\theta^{o}_{m+1}(x_2-x_1)) = \theta^{o}_m f^{(m+1)}(x_1+\theta^{m}_1\theta^{o}_m(x_2-x_1)). \qquad /9/$$

З існування $(m+1)$-ї похідної Гато у кожній точці випуклої множини $D_0 \subset D$ випливає неперервність всіх функцій, які є елементами гіперматриць $f^{(m+1)}(x_1+\theta^{o}_{m+1}(x_2-x_1)), f^{(m+1)}(x_1+\theta^{m}_1\theta^{o}_m(x_2-x_1))$. Переходячи в /9/ до границі при $x_1 \to x_2$, одержуємо співвідношення /5/. Теорема доведена.

Зауваження 1. Співвідношення /5/ істинне для всіх θ^{o}_m, які відповідають ненульовим елементам гіперматриці $f^{(m+1)}(x_2)$.

Зауваження 2. При виконанні умов теореми 1 послідовне використання аналога теореми про середнє значення дає

$$f(x_2)=f(x_1)+f'(x_1)(x_2-x_1)+\dots+\prod_{i=0}^{m-1}\bar{\theta}_i^{\,m-i} f^{(m+1)}(x_1+\prod_{i=0}^{m}\bar{\theta}_i(x_2-x_1))(x_2-x_1)^{m+1}, \qquad /10/$$

де $\bar{\theta}_0, \bar{\theta}_1, \dots \in (0;1)$. Однак, якщо жодна "координата" всіх гіперматриць $f^{(i)}(x_2)$ $(i=1,2,\dots,m+1)$ не дорівнює нулю, то

$$\bar{\theta}_i \xrightarrow[x_1 \to x_2]{} \frac{1}{2} \quad (i=1,2,\dots,m).$$

Комбіноване застосування розкладів /6/, /10/, теореми 1, зауваження 2 дає змогу оцінювати "зверху" та "знизу" інтервальні розширення апроксимуючих виразів.

У табл. 1-4 наведені значення величин θ^{o}_1 вказаних там функцій при різних значеннях точок x_1, x_2.

В табл. 5 наведені значення θ^{o}_2, θ^{o}_3, θ^{o}_4 функції $y=\ln x$ при різних значеннях точок x_1, x_2. Результати обчислень /табл. 1-5/ підтверджують положення теореми 1, що $\theta^{o}_1 \to \frac{1}{2}$, $\theta^{o}_2 \to \frac{1}{3}$, $\theta^{o}_3 \to \frac{1}{4}$, $\theta^{o}_4 \to \frac{1}{5}$ при $x_1 \to x_2$.

Таблиця 1

$y = \ln x$

x_1	x_2	θ_1°
.5	1	.442695
.75	1	.4760652
.9	1	.4914121

Таблиця 2

$y = \ln x$

x_1	x_2	θ_1°
1	5	.3713353
4.0	5	.4814432
4.5	5	.4912680
4.99	5	.497109

Таблиця 3

$y = x^4$

x_1	x_2	θ_1°
.7	1	.5291096
.8	1	.518442
.9	1	.508766
.993	1	.5004142

Таблиця 4

$y = x^{100}$

x_1	x_2	θ_1°
.1	1	.9506275
.[illegible]	1	.9225106
.9	1	.770097
.999	1	.5099

Таблиця 5

x_1	x_2	θ_2°	θ_3°	θ_4°
5	1	.2073355	.1402292	.1050953
3	1	.2447826	.1719628	.1318949
2	1	.2764997	.1994903	.1556054
1.5	1	.2997938	.2200558	.1735284
1.4	1	.3055087	.2251572	.178283
1.3	1	.3115613	.230663	.1827493
1.2	1	.318351	.2363695	.192254
1	.5	.2764997	.1994903	.1556054
.8	.5	.2945391	.2153472	.1694846
.6	.5	.3183514	.2363697	.1922541
.58	.5	.3210737	.2431902	.1934442

З допомогою теореми 1 можна побудувати для розв"язування систем нелінійних рівнянь інтервальні ітераційні процеси вищих порядків. Перший такий метод /інтервальний метод типу Рунге/ запропонований у [4]. Використовуючи граничні співвідношення між величинами θ^0_{m+1} та θ^1_m у розкладах функцій $f(x)$ та $f'(x)$ в ряди Тейлора, апроксимацію відрізка ряду Тейлора функції $f(x)$ лінійною комбінацією лінійних операторів за методикою з [2], побудовані наступні інтервальні ітераційні процеси.

<u>Алгоритм 1</u>:

$\mathcal{X}_0$ – довільний початковий інтервал,

$$\bar{K}_{n+1} = m(\mathcal{X}_n) - B_n f(m(\mathcal{X}_n)) + [\mathcal{I} - B_n \cdot f'(\mathcal{X}_n)](\mathcal{X}_n - m(\mathcal{X}_n)),$$

$$\mathcal{Y}_n = \bar{K}_{n+1} \cap \mathcal{X}_n,$$

$$K_{n+1} = m(\mathcal{Y}_n) - C_n f(m(\mathcal{Y}_n)) + [\mathcal{I} - C_n \cdot \mathcal{F}'(\mathcal{Y}_n)](\mathcal{Y}_n - m(\mathcal{Y}_n)),$$

$$\mathcal{X}_{n+1} = K_{n+1} \cap \mathcal{X}_n \quad (n=0,1,\dots),$$

де C_n -приблизна інверсія центра матриці $\mathcal{F}'(\mathcal{Y}_n) = \frac{1}{4} f'(m(\mathcal{Y}_n)) + \frac{3}{4} f'(m(\mathcal{Y}_n) + \frac{2}{3}(\mathcal{Y}_n - m(\mathcal{Y}_n)))$; $m(\mathcal{X})$ -середина інтервалу $\mathcal{X}$; B_n - приблизна інверсія центра матриці $f'(\mathcal{X}_n)$; $\mathcal{I}$ - одинична матриця.

<u>Алгоритм 2</u>:

$\mathcal{X}_0$ -довільний початковий інтервал,

$$K_{n+1} = m(\mathcal{X}_n) - C_n f(m(\mathcal{X}_n)) + (\mathcal{I} - C_n \mathcal{F}'(\mathcal{X}_n))(\mathcal{X}_n - m(\mathcal{X}_n)),$$

$$\mathcal{X}_{n+1} = K_{n+1} \cap \mathcal{X}_n \quad (n=0,1,\dots).$$

<u>Алгоритм 3</u>:

$\mathcal{X}_0$ -довільний початковий інтервал,

$$K_{n+1} = m(\mathcal{X}_n) - B_n f(m(\mathcal{X}_n)) - B_n \mathcal{F}'(\mathcal{X}_n)(\mathcal{X}_n - m(\mathcal{X}_n)),$$

$$\mathcal{X}_{n+1} = K_{n+1} \cap \mathcal{X}_n \quad (n=0,1,\dots).$$

При побудові алгоритмів 1–3 використано співвідношення

$$\frac{1}{2} f''(x_1 + \theta^0_2 (x_2 - x_1))(x_2 - x_1)^2 = \int_0^{x_2 - x_1} f''(x_1 + \theta^1_1 t)\, t\, dt \qquad /11/$$

При $x^* \in X_n$ маємо $x^* \in X_{n+1}$, де X_{n+1} визначають за формулами будь-якого з алгоритмів 1 – 3.

Наведені вище алгоритми здійснюють знаходження всіх дійсних коренів системи /1/, які належать початковому інтервалу X_0, в тому числі і кратних, з заданою точністю. Використовуючи алгоритми поділу проміжних інтервалів X_n [8], описані алгоритми спочатку локалізують корені системи /1/ у початковому інтервалі X_0, а далі обчислюють у виділених підінтервалах вже одиничні корені /1/. Якщо у деякому інтервалі коренів /1/ немає, то в усіх наведених алгоритмах за скінченну кількість кроків перетини $K_{n+1} \cap X_n$ дають пусту множину.

Алгоритми 1-3 не еквівалентні. Наприклад, алгоритм 1 має локальний порядок збіжності не нижчий 5, а алгоритми 2,3 – не нижчий 4. Алгоритм 1 більш трудомісткий порівняно з алгоритмами 2,3. Однак він збігається з вказаним порядком з ширшого початкового інтервалу, що часто дає змогу знайти всі корені в інтервалі X_0 за менший час, ніж за допомогою алгоритмів 2,3. З іншого боку, алгоритм 2 шукає корені системи /1/ в X_0 швидше, ніж алгоритм 3. Однак він більш "чутливий" до розміщення коренів системи /1/ відносно інтервалу X_0.

У табл. 6 наведені результати розв"язування за допомогою інтервального метода Ньютона в інтерпретації Кравчика [7] та алгоритмів 1 3 двох систем:
тестової [6]:

$$\begin{cases} 6x^5 - 25.2x^3 + 24x - 6y = 0, \\ 12y - 6x = 0, \end{cases} \qquad /12/$$

та системи

$$\begin{cases} x^4 - 2x^3 u + 2xy^3 - y^4 + x^3 + x^2 y - 5xy^2 + 3y^3 - 87x^2 + 4xy + 83y^2 - 265x - 435y + 350 = 0, \\ x^7 + 3x^6 y - 2x^5 y^2 - 10x^4 y^3 + x^3 y^4 + 11x^2 y^5 - 4y^7 - 4x^6 - 11x^5 y + 10x^4 y^2 + 30x^3 y^3 - 16x^2 y^4 - \\ - 19xy^5 + 10y^6 - 35x^5 - 20x^4 y + 16x^3 y^2 - 58x^2 y^3 - xy^4 + 98y^5 + 140x^4 + 30x^3 y + \\ + 164x^2 y^2 + 102xy^3 - 112y^4 + 259x^3 + 23x^2 y - 94xy^2 - 728y^3 - 1036x^2 - \\ - 403xy + 98y^2 - 225x + 1530y + 900 = 0. \end{cases}$$

/13/

Характерна особливість системи /13/ та, що вона має 17 дійсних коренів у квадраті $\{x,y\} = [-10,10] \times [-10,10]$, причому корінь $(x,y)=(-2,3)$ двократний. Всі корені знаходили з точністю $\varepsilon = 10^{-6}$. Наведені в табл. 6 обчислення виконав П.С.Венгерський.

Таблиця 6

Система	Початковий інтервал	Кількість коренів	Кількість ітерацій за алгоритмом			
			Ньютона	1	2	3
/12/	$\begin{cases}[-2,3]\\ [-2,3]\end{cases}$	5	109	39	51	67
	$\begin{cases}[.6,2.9]\\ [.6,2.9]\end{cases}$	1	42	15	27	-
	$\begin{cases}[3,10]\\ [3,10]\end{cases}$	Немає коренів	16	5	-	-
	$\begin{cases}[.5,5.0]\\ [.5,5.0]\end{cases}$	2	113	15	-	-
/13/	$\begin{cases}[-1.5,4.0]\\ [-3.0,8.0]\end{cases}$	5	2068	937	1043	[illegible]
	$\begin{cases}[-2.0,2.0]\\ [.5,2.5]\end{cases}$	1	318	142	151	-
	$\begin{cases}[-8.0,-1.5]\\ [2.0,2.5]\end{cases}$	1	149	51	79	-

Очевидно, що наведена в табл. 6 кількість не ілюструє ні швидкості збіжності алгоритмів, ні часу, затраченого на знаходження кожного кореня системи /1/ кожним із названих алгоритмів - уточненню кореня завжди передувала локалізація його за неоднакову кількість кроків для різних алгоритмів. Отже, в табл. 6 відображена лише загальна тенденція.

Швидкість збіжності інтервальних ітераційних процесів суттєво залежить від точності знаходження інтервальних оцінок [1] відповідних функцій. Тому детальніше розглянемо можливості підвищення точності цих оцінок. Інший підхід до розв"язання цієї проблеми запропоновано в [9].

Перша можливість випливає з ненаявності в інтервальній арифметиці закону дистрибутивності /має місце лише субдистрибутивність/ та відсутністю в ній єдиного елемента як додавання, так і множення.

Тому перед побудовою інтервального розширення даної функції її потрібно зобразити у вигляді суми найме..ної можливої кількості доданків. Це досягається розкладом виразу на множники, зведенням дробів до спільного знаменника, знищенням подібних членів, використанням "формул згортання" виразів. Наприклад, потрібно оцінити функцію $f(x)=x^3-3x^2+3x-1$ на інтервалі $[-3,2]$. Безпосередньо одержуємо $f([-3,2])=[-64,41]$. Але $f(x)=(x-1)^3$. Однак $[-4,1]^3=[-64,16]\subset[-64,41]$.

Друга можливість підвищення точності інтервального оцінювання функцій полягає в тому, що всі складові в аналітичному заданні даної функції є функціями одного і того ж аргументу. Таким чином, всі "біжучі" точки інтервалів – інтервальних оцінювань складових функцій цілком залежні між собою. Такі інтервали потрібно розглядати в інтервальних операціях як впорядковані множини точок /порядок задає спільний аргумент/. З кожним значенням аргумента пов"язуємо інтервал, яки" містить значення відповідної складової ᵩун..ції при такому значенні аргумента. Наприклад, якщо функція $f(x)$ монотонно зростає на деякому інтервалі $[a,b]$, то за нижню межу її значень на цьому інтервалі можна взяти с"чну, що проходить через точки $(a,f(a)),(b,f(b))$, а за верхню – відрізки дотичн.., проведених у цих же точках до графіка функції $f(x)$.

1. А л е ф е л ь д І., Х е р ц б е р г е р Ю. Введение в ин ервальные вычисления. М., '987. 2. Б а р т і ш М.Я., С е н ь о П.С. Про методи типу Рунге для розв"язування нелінійних операторних рівнянь //Доп. АН УРСР. Сер. А. № 9. 1972. С.771-775. 3. О р т е г а Дж., Р е й н б о л д т В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М., 19'5. 4. С е н ь о П.С. Построение интервального метода типа Рунге //Распараллеливание обработки информации. Тез. докл. Львов, 1985. С.50-51. 5. С е н ь о П.С. Интервальный итерационный процесс решения нелинейных систем уравнений //Применение вычислительной техники и математических методов в научных и экономических исследованиях. Тез. докл. К., 1988. С.64-65. 6. Hansen E., Sengupta S. Boundi.. Solutions of Systems of Equations Using Interval Analysis // BIT. 1981. 21. P.203-211. 7. Krawczyk R. Intervalliterationsverfahren. Ber. Math. // Statist. Sek. Forschungszent. Graz. 1982. Vol. 185-.. 9. P. 1-49. 8. Moore R.E. Interval Metho.s for Nonlinear Systems // Computing. 1980. Vol. 2. P. 113-120. 9. Nickel K. Die Überschätzung des Wertebereiches einer Funktion in der Intervallrechnung mit Anwendungen auf lineare Gleichungssysteme. 1977. Bd. 18. S. 15-36.

Стаття надійшла до редколегії 05.09.88