

4.2 Розв'язання систем нелінійних рівнянь

Розглянемо задачу розв'язання системи нелінійних рівнянь.

Системи лінійних алгебраїчних рівнянь – це лише окремий випадок систем рівнянь. Більшість математичних моделей різних процесів і явищ записуються в загальному випадку у вигляді систем нелінійних рівнянь, тому дана задача має величезне практичне значення. Розглянемо систему n нелінійних рівнянь з n невідомими x_1, x_2, \dots, x_n :

[illegible]

де $f_i(x_1, x_2, \dots, x_n)$, $i = \overline{1, n}$ – деякі нелінійні функції n змінних.

Як і для систем лінійних рівнянь, розв'язок системи нелінійних рівнянь – це такий вектор x^* , який забезпечує тотожність системи рівнянь. Система рівнянь може не мати розв'язків, мати єдиний розв'язок, скінченну чи нескінченну кількість розв'язків.

На відміну від систем лінійних рівнянь, для систем нелінійних рівнянь невідомі прямі методи розв'язання. Лише в окремих випадках систему можна розв'язати безпосередньо. Наприклад, для системи з двох рівнянь іноді вдається подати одне невідоме через інше й у такий спосіб звести задачу до розв'язання одного нелінійного рівняння щодо одного невідомого. Тому ітераційні методи для нелінійних систем набувають особливої актуальності. Можливість одержання розв'язку, як правило, залежить від того, наскільки *вдало обране початкове наближення* – вектор $\mathbf{x}^0 = (x_1^0, \dots, x_n^0)$. Виділяють три основні ітераційні методи розв'язання систем нелінійних рівнянь, а саме, *метод простої ітерації, метод Зейделя і метод Ньютона*.

4.2.1 Метод простої ітерації розв'язання систем нелінійних рівнянь

Для реалізації цього методу потрібно, шляхом алгебраїчних перетворень, виокремити з кожного рівняння по одній змінній і в такий спосіб привести систему (1) до вигляду

[illegible]

В матричному вигляді

[illegible]

Потім задають вектор початкового наближення $x^{(0)}$ і точність $\varepsilon > 0$.

Підставляючи вектор початкового наближення $x^{(0)}$ у перетворену систему рівнянь (2), з першого рівняння обчислюють нове наближення до першої змінної, з другого – до другої змінної і т.д. Обчислені уточнені значення змінних знову підставляють у рівняння (2). Отже, на $(k+1)$ -му кроці ітераційної процедури формула методу ітерацій для розв’язання системи нелінійних рівнянь має вигляд

$$\left\{ \begin{array}{l} x_1^{(k+1)} = g_1(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) \\ x_2^{(k+1)} = g_2(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) \\ \\ x_n^{(k+1)} = g_n(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) \end{array} \right. . \quad (3)$$

Процес (3) триває доти, доки не виконається умова $\left\|x^{(k)} - G\left(x^{(k)}\right)\right\| \leq \varepsilon$.

Ітераційний процес для системи нелінійних рівнянь (2) збігається до єдиного її розв'язку, якщо норма матриці $G'(x)$ є менша за одиницю, тобто для збіжності розв'язків системи достатньою є умова

$$\|G'(x)\| < 1, \quad (4)$$

де $G'(x)$ – матриця частинних похідних, яку називають *матрицею Якобі* (якобіаном) системи (2):

$$G'(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \frac{\partial g_n}{\partial x_2} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}.$$

Основними недоліками методу простої ітерації є:

- складність переходу від запису системи (1) до виду (2), зважаючи на нелінійність функцій $f_i(x_1, x_2, \dots, x_n)$, $i = \overline{1, n}$;
- збіжність тільки для достатньо близьких до розв'язку початкових наближень $x^{(0)}$;
- невисока швидкість збіжності.

Основною перевагою методу простої ітерації є невисока трудомісткість методу.

4.2.2 Метод Зейделя

Метод Зейделя для систем нелінійних рівнянь, як і для системи лінійних рівнянь, полягає у використуванні уточнених значень змінних вже на поточному ітераційному кроці. Так, для уточнення на (k+1)-му кроці значення першої змінної (x_1) використовуємо усі значення попереднього k-го кроку, для другої змінної – значення x_1 (k+1)-го кроку, а для решти змінних – значення попереднього k-го кроку:

$$\left\{ \begin{array}{l} x_1^{(k+1)} = g_1 \left(x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \dots, x_n^{(k)} \right) \\ x_2^{(k+1)} = g_2 \left(x_1^{(k+1)}, x_2^{(k)}, x_3^{(k)}, \dots, x_n^{(k)} \right) \\ x_3^{(k+1)} = g_2 \left(x_1^{(k+1)}, x_2^{(k+1)}, x_3^{(k)}, \dots, x_n^{(k)} \right) \\ \dots\dots\dots \\ x_n^{(k+1)} = g_n \left(x_1^{(k+1)}, x_2^{(k+1)}, x_3^{(k+1)}, \dots, x_n^{(k)} \right) \end{array} \right. , \quad (5)$$

де k – номер ітерації.

Зауважимо, що умови збіжності розв'язків системи нелінійних рівнянь для методу Зейделя такі самі, як і для методу простих ітерацій ($\|G'(x)\| < 1$ формула (4)).

Для цього методу дуже важко забезпечити збіжність, а інтервал збіжності може бути настільки вузьким, що вибір початкових наближень сильно ускладнюється.

4.2.3 Метод Ньютона

Математичним підґрунтям методу є лінеаризація функцій f_1, f_2, \dots, f_n шляхом розкладання в ряд Тейлора в околі точки початкового наближення до розв'язку системи рівнянь й нехтування всіма членами ряду, окрім лінійних щодо приростів змінних.

Розглянемо систему нелінійних рівнянь (1). Введемо позначення:

$x = (x_1, x_2, \dots, x_n)^T$ – вектор-стовпець розмірності n з елементами x_i ,
 $F(x) = (f_1(x), f_2(x), \dots, f_n(x))^T$ – векторна функція розмірності n , елементами якої є
 функції $f_i(x) = f_i(x_1, x_2, \dots, x_n)$, $i = \overline{1, n}$, тоді систему (1) можна записати у векторному
 вигляді:

$$F(x) = 0. \quad (6)$$

Метод Ньютона будує ітераційну послідовність $\{x^{(k)}\}$, $k = 0, 1, 2, \dots$, наближень розв'язку x^* (початкове наближення $x^{(0)}$ задається) за такою ітераційною формулою:

$$x^{(k+1)} = x^{(k)} - [F'(x^{(k)})]^{-1} \cdot F(x^{(k)}) \quad (7)$$

$$\text{де } F'(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} - \text{матриця Якобі системи (6).}$$

Процес (7) триває доти, доки не виконається умова $\|F(x^{(k)})\| \leq \varepsilon$, де ε – задана точність розв'язку задачі (6).

Ідея методу Ньютона полягає в тому, що на k -ій ітерації ($x^{(k)}$ – поточне наближення розв'язку) наступне наближення розв'язку $x^{(k+1)}$ знаходиться, як розв'язок системи лінійних рівнянь:

$$F_k(x) = 0, \quad (8)$$

де $F_k(x) \equiv F(x^{(k)}) + F'(x^{(k)})(x - x^{(k)})$ (перші два члени розкладання в ряд Тейлора функції $F(x)$ в околі точки $x^{(k)}$), тобто система (8) є лінеаризацією (лінійним наближенням) системи (6). Оскільки система (8) лінійна відносно x , то її розв'язок може бути знайдений аналітично:

$$\begin{aligned} F_k(x) &\equiv F(x^{(k)}) + F'(x^{(k)})(x - x^{(k)}) = 0 \\ F'(x^{(k)})(x - x^{(k)}) &= -F(x^{(k)}). \end{aligned} \quad \text{або}$$

Домножуючи на обернену матрицю Якобі (за умови $\det(F') \neq 0$), отримаємо

$$(x - x^{(k)}) = -[F'(x^{(k)})]^{-1} F(x^{(k)}).$$

звідки і отримуємо розв'язок системи (8)

$$x = x^{(k)} - [F'(x^{(k)})]^{-1} \cdot F(x^{(k)}).$$

Основними недоліками методу Ньютона є:

- *чутливість* до вибору початкового наближення $x^{(0)}$: при невдалому виборі метод може збігатися занадто повільно або навіть почати розбігатися;
- $\det(F') \neq 0$ в околі кореня;
- висока трудомісткість методу, оскільки на кожній ітерації необхідно обчислювати матриці $F'(x^{(k)})$ і $[F'(x^{(k)})]^{-1}$.

Основною перевагою методу Ньютона є висока швидкість збіжності. Швидкість збіжності є *квадратичною* (зменшення похибки на кожному кроці пропорційно $1/n^2$).

Слід зазначити, що в методі Ньютона формулу (7) можна записати у вигляді $x^{(k+1)} = x^{(k)} + h^{(k)}$, де $h^{(k)} = -[F'(x^{(k)})]^{-1} \cdot F(x^{(k)})$. Проте при практичній реалізації методу вектор $h^{(k)}$ ефективніше обчислювати як розв'язок системи лінійних рівнянь виду $F'(x^{(k)}) \cdot h^{(k)} = -F(x^{(k)})$.

У зв'язку з трудомісткістю методу Ньютона була запропонована його модифікація, за якої обернена матриця до $F'(x)$ шукається тільки один раз. За цієї модифікації послідовні наближення шукаються за формулою

$$x^{(k+1)} = x^{(k)} - \left[F'(x^{(0)}) \right]^{-1} \cdot F(x^{(k)}).$$

В цьому методі не треба обчислювати обернену матрицю Якобі на кожному кроці розрахунків, що спрощує алгоритм, але уповільнює збіжність і робить метод більш чутливим до вибору початкового наближення.

4.2.4 Інші методи розв'язання

Багато сучасних методів розв'язання системи (1) будуються на інших ідеях, ніж ті, що розглянуто вище.

А саме, велику популярність у наш час набули методи, що зводять розв'язання системи (1) до задачі відшукування екстремуму функції багатьох змінних.

Наприклад, розглянемо функцію

$$\Phi(x_1, x_2, \dots, x_n) = \sum_{i=1}^n f_i^2(x_1, x_2, \dots, x_n).$$

Ясно, що кожному розв'язку $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ системи (1) відповідає нульовий мінімум функції $\Phi(x)$ та, навпаки, кожна точка нульового мінімуму $\Phi(x)$ дає розв'язок системи (1). Отже, задача відшукування розв'язків системи (1) зводиться до задачі відшукування точок нульового мінімуму допоміжної функції $\Phi(x)$. Для розв'язання цієї задачі **застосовуються градієнтні методи, методи покоординатного спуску** та інші. У цьому курсі не розглянуто ці методи, оскільки вони докладно будуть вивчатись у курсі “Математичні методи дослідження операцій” та супутніх спецкурсах, присвячених вивченню методів оптимізації.