

Чисельні методи. Застосування . Основні принципи побудови.

Досвід розв'язування науково-дослідних і прикладних задач показує, що незалежно від їхньої складності кінцевої мети можна досягти або постановкою експерименту, або методом математичного моделювання. Кожен з цих методів має свої переваги і недоліки.

За допомогою експерименту можна розв'язувати навіть дуже складні задачі, при цьому достовірність результатів тим вища, чим ретельніше відпрацьована методика експерименту. Водночас здобуті результати будуть стосуватися тільки тих умов, за яких проводився експеримент, внаслідок чого узагальнення результатів на інші умови не коректне. Крім того, треба враховувати економічний бік постановки складного експерименту. Щодо цього, то більші можливості має метод математичного моделювання за допомогою ЕОМ, коли аналізують не реальну задачу, а її модельне зображення.

Процес математичного моделювання зображують у такій послідовності: фізична постановка задачі; математична постановка задачі; математичне дослідження задачі; аналіз і осмислення математичного розв'язку та порівняння його з експериментом.

Розглянемо докладніше математичну постановку і математичне дослідження задачі.

Математична постановка полягає у формуванні математичної моделі досліджуваної задачі, яка звичайно є системою рівнянь математичної фізики (диференціальних, Інтегральних, інтегрально-диференціальних).

Математичне дослідження задачі власне зводиться до розв'язування системи рівнянь і аналізу здобутих результатів. Для порівняно простих задач вдається розв'язати вихід :у систему рівнянь і розв'язок подати у вигляді залежностей, виражених через елементарні та інші відомі функції. Якщо це можливо, то говорять, що знайдено аналітичний (точний) розв'язок задачі. Однак переважна більшість практично важливих задач аналітичних розв'язків не має. До таких належать, наприклад, задачі будівництва: визначення напружено-деформованого стану пластин, плит, фундаментів; задачі стійкості, теплопровідності для твердих тіл; напрямленої дифузії тощо. У цих випадках використовують чисельні методи, які, оперуючи системою алгебраїчних рівнянь (аналогів рівнянь математичної фізики), дають можливість побудувати деяку послідовність арифметичних операцій, збільшення кількості яких до нескінченності дає точний розв'язок. Оскільки на практиці здійснюють скінченне число кроків (операцій), то знайдений розв'язок є наближеним. А через те що обчислювальні операції виконують над числами, то відповідні методи дістали назву *чисельних*. Найбільшого розвитку чисельні методи набули останнім часом завдяки застосуванню ЕОМ, що мають високу швидкість обчислень і велику ємність оперативної пам'яті. Проте основна роль при цьому відводиться, звичайно, людині, яка повинна вміти сформулювати і поставити задачу, описати її математичними залежностями (створити математичну модель об'єкта), скласти алгоритм розв'язання задачі на ЕОМ, написати програму на алгоритмічній мові, зрозумілій машині, розв'язати задачу й оцінити результати.

Щодо оцінювання результатів розрахунку, то слід зазначити, що поєднання чисельних методів і ЕОМ дає можливість зробити це ефективно й оперативно, варіюючи найсуттєвіші параметри розрахункової схеми задачі з наступним чисельним аналізом впливу їх на кінцевий результат. Фактично йдеться про чисельний експеримент, оскільки умови задачі можна змінювати багато разів.

Незважаючи на відмінності в методології, до чисельного експерименту щільно примикають фізичний експеримент і фізичне дослідження, особливо у тій частині, де потрібна оцінка достовірності здобутих результатів.

Математична модель об'єкта — це та сукупність рівнянь, за допомогою якої досліджують реальні фізичні об'єкти (процеси, явища). Математична модель не тотожна досліджуваному об'єкту, а є лише його наближеним описом, оскільки її будують з деякими спрощеннями та ідеалізацією. У моделі враховують найважливіші моменти і взаємозв'язки, найхарактерніші для досліджуваного реального об'єкта. Разом з тим внаслідок заміни реального об'єкта відповідною йому математичною моделлю стало можливим сформулювати задачу як математичну і скористатися для її розв'язання тим чи іншим математичним апаратом.

Алгоритм — це зрозумілий і точний припис (вказівка) виконавцеві здійснювати послідовність дій, спрямованих на досягнення зазначеної мети або розв'язання поставленої задачі.

Точність розв'язку — це міра близькості чисельного розв'язку до аналітичного.

Збіжність розв'язку — це поступове наближення його до точного.

Після вибору математичної моделі об'єкта і опису її на алгоритмічній машинній мові здійснюють чисельну реалізацію задачі на ЕОМ. Останнім часом при реалізації практичних задач здебільшого застосовують ЕОМ, що можуть виконувати від кількох сотень до мільйонів операцій за секунду. Найбільшого застосування в інженерних розрахунках набули ЕОМ, які мають не тільки високу швидкість обчислень, сучасне програмне забезпечення, а й розвинуту сервісну частину, яка дає можливість оперативно діагностувати похибки, графічно відображати результати обчислень, здійснювати розрахунки в режимі діалогу. Великої популярності у користувачів набули також міні- та мікро-ЕОМ, персональні комп'ютери.

Розв'язування багатьох інженерних задач зводиться до обчислення коренів одного нелінійного рівняння або до розв'язання систем нелінійних рівнянь. В обох випадках нелінійні рівняння, що утворюються, можна поділити на два типи — алгебраїчні та трансцендентні.

Алгебраїчними називають рівняння, що містять лише алгебраїчні функції (цілі, раціональні, ірраціональні).

Нелінійні рівняння, що містять тригонометричні, логарифмічні, показникові, степеневі чи інші спеціальні функції, називають трансцендентними.

Нелінійне рівняння можна подати у таких формах запису;

або

$$f(x) = 0 \quad (2.1)$$

$$\varphi(x) = g(x), \quad (2.2)$$

де функції, що входять до (2.1), (2.2), визначені і неперервні на множині x , що називається областю визначення рівняння. Сукупність значень змінної x , при яких

методи розв'язування цих рівнянь, які дають змогу за допомогою скінченного набору арифметичних операцій обчислити корені будь-якого нелінійного рівняння чи системи нелінійних рівнянь з достатньою точністю. Особливо ефективні наближені методи при реалізації на ЕОМ, оскільки використовувані для цього алгоритми є простими, зручними, легко програмуються.

Наближеним значенням кореня x нелінійного рівняння з точністю до ϵ вважають будь-яке число між a і b , при якому виконується умова $b - a \leq \epsilon$. Числа a і b — це наближені значення кореня x відповідно з недостачею і з надлишком з точністю до ϵ . Наприклад, якщо корінь лежить між числами 1,133 і 1,134, то за наближене значення кореня з точністю до 0,001 можна взяти будь-яке число у межах між цими числами, наприклад число 1,1335.

Універсальні алгоритми обчислення коренів нелінійних рівнянь ґрунтуються на тому, що виходять з будь-якого вже відомого наближеного значення $x^{(0)}$ одного з коренів, якщо розв'язується одне рівняння, або значень $x_1^{(0)}, x_2^{(0)} \dots x_n^{(0)}$ якщо розв'язується система рівнянь (n — порядок системи). Ці значення далі уточнюються до заданого ступеня точності. При цьому несуттєво, як знайдено початкове значення обчислюваного кореня (коренів). Часто початкове наближення дістають за допомогою грубого попереднього підрахунку, за допомогою графіка або з аналізу фізичної суті задачі.

Алгоритм наближеного обчислення коренів нелінійного рівняння складається з двох етапів.

1. Відшукування достатньо малих відрізків (інтервалів), у кожному з яких міститься один і тільки один корінь. Цей етап називають відокремленням коренів (або визначенням відрізків ізоляції кореня).

2. Обчислення кореня з наперед заданою точністю ϵ , якщо відоме його деяке початкове наближення в інтервалі, що не містить інших коренів. Цей етап називають уточненням наближених значень коренів.

Відокремити корені можна кількома способами — графічним, аналітичним або методом послідовного перебирання. Останній метод зручний при використанні ЕОМ.

**Чисельне розв'язання трансцендентних рівнянь.
Опис методів дихотомії (половинного ділення), хорд, дотичних,
комбінованого методу хорд та дотичних.**

Трансцендентні рівняння- такі, що містять тригонометричні, показникові, степеневі чи інші спеціальні функції.

Для їх розв'язання необхідно наступні дії

Постановка задачі.

Задано функцію

$$f(x) = 0 \quad (1)$$

де $f(x)$ деяка функція аргументу x , що визначена на інтервалі $[a;b]$.

Коренем рівняння (1) називається всяке число $\xi \in [a;b]$, що перетворює функцію $f(x)$ в нуль, тобто $f(\xi) = 0$. Задача пошуку кореня рівняння поділяється на два етапи:

1. Віднімання кореня, т.б. виділення відрізка, на якому розміщено тільки один корінь. При цьому один з кінців відрізка або його середину вибирають за початкове наближення.

В багатьох випадках відділення кореня можна провести графічно. Приймаючи до уваги, що дійсні корені рівняння (1) - це точки перетину графіка функції $f(x)$ з віссю абсцис, достатньо побудувати графік $f(x)$ і відмітити на осі Ox відрізки, які містять один корінь. Побудову графіків часто вдається сильно спростити, замінивши рівняння (1) рівносильним йому рівнянням $\varphi(x) = \psi(x)$ (2).

В цьому випадку будуються графіки функцій $\varphi(x)$ і $\psi(x)$, а потім на осі Ox відмічають відрізки, які локалізують абсциси точок перетину цих графіків.

Умова існування кореня

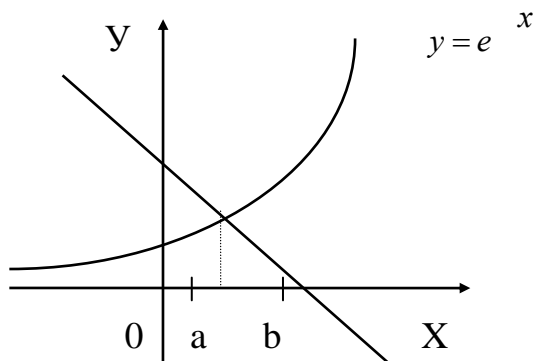
- 1) Якщо неперервна на відріжку $[a;b]$ функція $f(x)$ приймає на його кінцях значення різних знаків, тобто $f(a) \cdot f(b) < 0$, то рівняння (1) має на цьому відріжку по меншій мірі один корінь.

- 2) Якщо функція $f(x)$ строго монотонна, то на $[a;b]$ корінь єдиний.

Приклад 1.

$$e^x + x - 2 = 0 \quad \text{- виділити корені}$$

$$\varphi(x) = e^x, \quad \psi(x) = x - 2$$



Визначаємо:

$$a = 0.3 \quad f(a) = -0,35014$$

$$b = 0.8 \quad f(b) = 1.0255$$

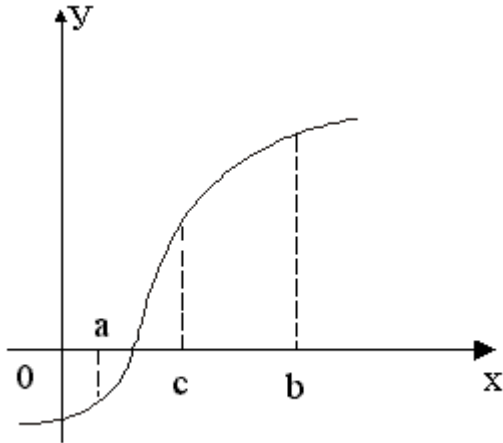
$y = x - 2$ Отже, корінь існує на проміжку $[a;b]$.

Достатня умова - постійність знаку похідної на $[a;b]$.

2. Уточнення кореня - це звуження границь виділеного відрізка ізоляції кореня за допомогою одного з методів доти, поки довжина відрізка не стане меншою, ніж насамперед задана точність ϵ .

Метод половинного ділення.
(дихотомії, Мюллера)

Нехай рівняння (1) на відрізку $[a;b]$ має єдиний корінь і функція на ньому неперервна. Поділимо відрізок $[a;b]$ пополам точкою $c=(a+b)/2$.



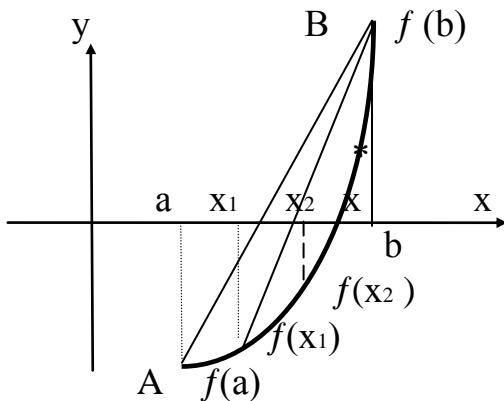
Якщо $f(c) \neq 0$, то можливі два випадки:
- або $f(x)$ змінює знак на $[a;c]$
- або $f(x)$ змінює знак на $[c;b]$.

Вибираючи в кожному випадку той з відрізків, на якому функція змінює знак, і продовжуючи процес поділу далі, можна дійти до скільки завгодно малого проміжку, що містить корінь.

Метод хорд.

Метод хорд - один з поширених ітераційних методів. Його ще називають методом лінійного інтерполювання, методом пропорційних частин, або методом хибного положення.

При виконанні попереднього методу більш раціональним було б ділити $[a;b]$ не навпіл, а пропорційно значенню функції в точках a та b . В цьому випадку точка поділу відрізка буде знаходитися на перетині хорди AB з віссю Ox .



Ідея методу хорд в тому, що на досить малому відрізку дуга кривої $y=f(x)$ замінюється хордою і абсциса точки перетину хорди з віссю Ox є наближеним значенням кореня. Нехай для визначеності $f'(x)>0$, $f''(x)>0$, $f(a)<0$, $f(b)>0$

Візьмемо за початкове наближення шуканого кореня x^* значення $x_0=a$. Через точки A і B

проведемо хорду і за перше наближення кореня x^* візьмемо абсцису x_1 точки перетину з віссю Ox . Тепер наближене значення x_1 кореня можна уточнити якщо застосувати метод хорд до відрізка $[x_1;b]$. Абсциса x_2 точки перетину хорди AB буде другим наближенням кореня. Продовжуючи цей процес необмежено, дістанемо послідовність $x_1, x_2, \dots, x_k, \dots$

наближених значень кореня x^* даного рівняння. Для виведення формули методу хорд запишемо рівняння прямої, що проходить через точки $A_k(x_k, f(x_k))$ і $B(b, f(b))$:

$$\frac{(y - f(x_k))}{(f(b) - f(x_k))} = \frac{(x - x_k)}{(b - x_k)}$$

поклавши $y=0$ знайдемо абсцису точки перетину хорди A_kB з вісю Ox :

$$x = x_k - \frac{f(x_k)}{f(b) - f(x_k)} (b - x_k), \quad k=0,1,2,3,\dots$$

Значення можна взяти за наступне наближення, тобто:

$$x_{k+1} = x_k - \frac{f(x_k)}{f(b) - f(x_k)} (b - x_k), \quad k=0,1,2,3,\dots$$

У цьому разі і тоді, коли $f(a)>0, f(b)<0, f'(x)<0, f''(x)<0$ кінець в $[a;b]$ є нерухомим. У загальному випадку нерухомим буде той кінець відрізка ізоляції кореня в якому знак функції $f(x)$ збігається із знаком другої похідної, а за початкове наближення x_0 можна взяти точку відрізка $[a;b]$, в якій $f(x) f''(x)<0$. Отже, метод хорд можна записати:

$$x_{k+1} = x_k - \frac{f(x_k)}{f(x_k) - f(c)} * (x_k - c), \quad k=0,1,2,3,\dots \quad (3)$$

$c = b$, якщо $f(b), f''(b)>0$

$c = a$, якщо $f(a), f''(a)>0$

З формули (3) видно, що метод хорд є метод ітерацій $x_{k+1} = \varphi(x_k)$, в якому

$$\varphi(x) = x - \frac{f(x)}{f(x) - f(c)} * (x - c) \quad (4)$$

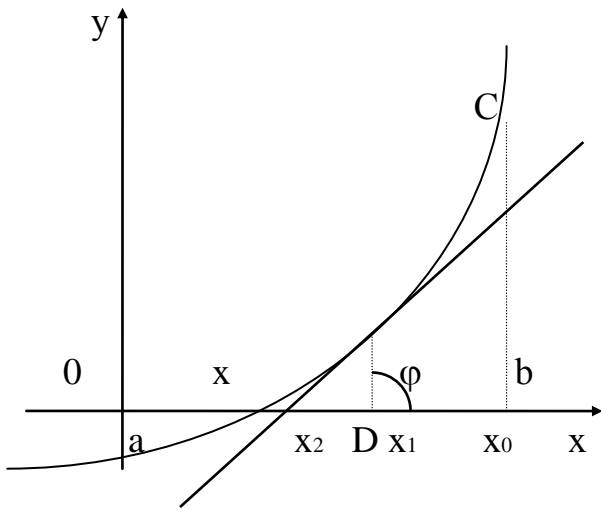
Метод дотичних (Ньютона).

Ідея методу полягає в послідовній заміні ділянки кривої $f(x)$ дотичною в точці c , що належить відріжку $[a;b]$ і перетинає вісь Ox в точці x_k . Точка c вибирається з умови: $f(c) * f''(c)>0$, яка гарантує збіжність процесу. При цьому необхідно, щоб

1. $f(x)=0$ малоєдиний корінь на $[a;b]$;

2. $f(x)$ була неперервна на $[a;b]$.

$f'(x)$ і $f''(x)$ не змінювали на ньому знак.



Отримаємо розрахункову формулу методу.

Розглянемо трикутник DCB. З малюнку видно, що наступне наближення x_1 отримуємо як:

$$x_1 = x_0 - DB$$

$$\text{З трикутника DCB: } DB = \frac{BC}{\operatorname{tg} \varphi}$$

$$\text{але } BC \rightarrow f(x_0); \operatorname{tg} \varphi = f'(x_0);$$

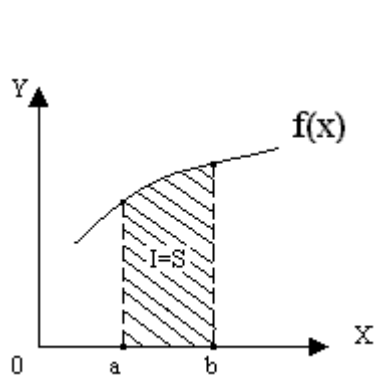
$$\text{таким чином } DB = \frac{f(x_0)}{f'(x_0)}$$

$$\text{а загальна формула } x_{k+1} = x_k - \frac{f(x_k)}{f'(x_{k+1})}$$

Чисельне інтегрування. Методи лівих, правих та середніх прямокутників.

Методи трапецій та парабол.

Чисельне інтегрування функцій.

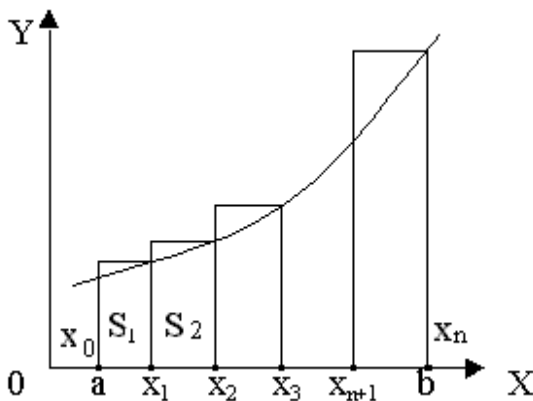


$$I = \int_a^b f(x) dx = F(b) - F(a). \text{ Якщо } f(x) \text{ має первісну, то}$$

значення інтегралу знаходиться за формулою Ньютона – Лейбніца. Але визначення первісної в більшості випадках інженерних задач є неможливою. Тому для визначення інтегралу використовують чисельні методи. Всі ці методи базуються на геометричній інтерпретації визначеного інтегралу, значення якого чисельно дорівнюють площі фігури, що обмежена

зверху – графіком функції $f(x)$, знизу віссю Ox , зліва та права межами інтегрування a та b . Для знаходження площі відрізок AB розбивають на рівні частини, довжиною h , де $h = (b-a)/N$ – крок інтегральної функції.

Метод прямокутників.



а)

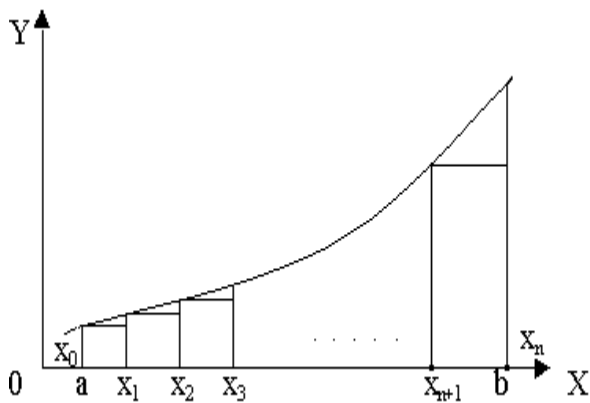
Замінімо елементарні (криволінійні) трапеції в діапазоні $[a;b]$ прямокутниками, і обчислимо загальну площу фігури, як суму площ окремих прямокутників. Для випадку а) знайдену площу назвемо площею лівих, а для в) – правих прямокутників.

Для а) маємо:

$$S^{\wedge} = f(x_1)(x_1 - x_0) + f(x_2)(x_2 - x_1) + \dots +$$

$$+ f(x_n)(x_n - x_{n-1}) = \frac{b-a}{N} \sum_{k=1}^N f(x_k)$$

Для б) маємо :



$$S^n = f(x_0)(x_1 - x_0) + f(x_1)(x_2 - x_0) + \dots + f(x_{n-1})(x_n - x_{n-1}) = \frac{b-a}{N} \sum_{k=0}^{N-1} f(x_k)$$

Істинне значення інтегралу обчислимо, як середнє арифметичне значення площ лівих і правих прямокутників: $I = (S^{\wedge} + S^n) / 2$

$$I = \frac{b-a}{2N} \left[\sum_{k=0}^N f(x_k) + \sum_{k=1}^N f(x_k) \right] = \frac{b-a}{2N} \left[f(a) + f(b) + 2 \sum_{k=1}^{N-1} f(x_k) \right]$$

б) Отримана формула наз. формулою лівих і правих прямокутників, або формулою трапеції.

Метод парабол.

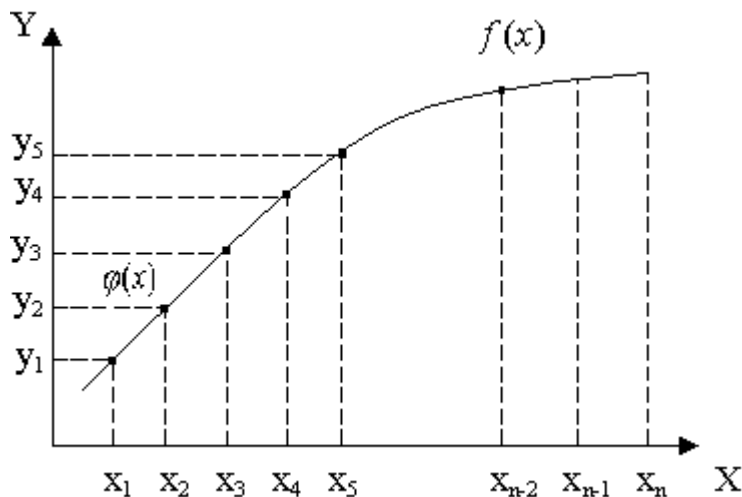
Більш точним методом визначення інтегралу є метод парабол (Сімпсона).

По цьому методу відрізок АВ ділять на $2n$ рівних частин, тобто кожен з проміжків ділять пополам. Розглянемо n -послідовність точок проміжку АВ,

x_1, x_2, \dots, x_n , де $x_1 < x_2 < x_n$, $x_2 = x_1 + h$, $x_3 = x_1 + 2h$, $x_n = x_1 + (n-1)h$. Вважаємо, що в цих точках задані значення деякої функції $f(x) = y$ і необхідно знайти інтеграл

$$\int_{x_1}^{x_n} f(x) dx.$$

Формула Сімпсона для наближеного інтегрування.



Через кожні послідовні три точки проводимо параболу $y = a + bx + cx^2$ і обчислюється інтеграл від функції, вираженої у вигляді цієї параболі. Цей інтеграл і вважається наближеним значенням шуканого інтегралу. Розглянемо три перші точки $(x_1; y_1)$, $(x_2; y_2)$, $(x_3; y_3)$.

Проведемо через них параболу: $y = a + bx + cx^2$ і обчислимо інтеграл:

$$\begin{aligned} S_1 &= \int_{x_1}^{x_3} (a_1 + b_1 x + c_1 x^2) dx = a_1(x_3 - x_1) + b_1 \left(\frac{x_3^2 - x_1^2}{2} \right) + c_1 \left(\frac{x_3^3 - x_1^3}{3} \right) = \\ &= \frac{x_3 - x_1}{6} (6a_1 + 3b_1(x_3 + x_1) + 2c_1(x_3^2 + x_3 x_1 + x_1^2)) = \frac{x_3 - x_1}{6} (c_1 x_1^2 + b_1 x_1 + a_1 + \\ &+ 4(c_1 \left(\frac{x_3 + x_1}{2} \right)^2 + b_1 \frac{x_3 + x_1}{2} + a_1) + c_1 x_3^2 + b_1 x_3 + a_1) = \frac{x_3 - x_1}{6} (y_1 + 4y_2 + y_3) = \\ &= \frac{h}{3} (y_1 + 4y_2 + y_3) \end{aligned}$$

Проводячи параболу через наступні три точки $(x_3; y_3)$, $(x_4; y_4)$, $(x_5; y_5)$ і обчислюючи інтеграл, отримуємо: $S_2 = \frac{h}{3} (y_3 + 4y_4 + y_5)$

Для отримання наближеного значення інтегралу по всій області від x_1 до x_n необхідно знайти суму отриманих значень.

$$S = S_1 + S_2 + S_{n-1} = \frac{h}{3} (y_1 + 4y_2 + y_3 + 4y_4 + y_5 + \dots + y_{n-2} + 4y_{n-1} + y_n) =$$

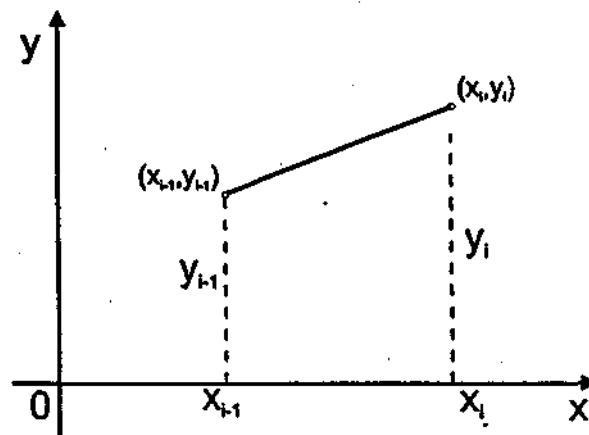
$$= \frac{h}{3} [y_1 + y_n + 4(y_2 + y_4 + \dots + y_{n-1}) + 2(y_3 + y_5 + \dots + y_n)] = \frac{h}{3} [f(x_1) + f(x_n) +$$

$$+ 4 \sum_{i=2}^n f(x_i) + 2 \sum_{i=3}^n f(x_i)]$$

Якщо інтеграл необхідно обчислити із заданою точністю, то необхідно знайти n (число проміжків) для її забезпечення.

Формула трапецій.

Розглянемо інший спосіб побудови квадратурних формул, що пов'язаний з аппроксимацією підінтегральної функції інтерполяційним многочленом. Розглянемо найпростіший випадок. Метод трапецій використовує лінійну інтерполяцію, тобто графік функції $y=f(x)$ подається у вигляді ламаної, що з'єднує точки (x_i, y_i) . В цьому випадку площа всієї фігури (криволінійної трапеції) складається з площ елементарних прямокутних трапецій (мал.3).



мал. 3

Площа кожної такої трапеції дорівнює добутку напівсуми основи на висоту:

$$S_i = \frac{y_{i-1} + y_i}{2} h_i, \quad h=1,2,3,\dots,n \quad (8)$$

Просумувавши ці рівності, отримуємо формулу трапецій для чисельного інтегрування:

$$I \equiv \int_a^b f(x) d(x) \approx \frac{1}{2} \sum_{i=1}^n h_i (y_{i-1} + y_i) \quad (9)$$

Для рівномірної сітки ($h_i=h$) ця формула має такий вигляд:

$$I \equiv \int_a^b f(x) dx \approx h \left(\frac{y_0 + y_n}{2} + \sum_{i=1}^n y_i \right) \equiv I_2 \quad (10)$$

Залишковий член має вигляд:

$$R_2[f] = I - I_2 = -\frac{(b-a)^3}{12} f''(\eta) \quad ; \quad \eta \in (a,b) \quad (11)$$

Використовуючи вираз (11) для залишкового члена, оцінку похибки квадратурної формули (10) можна надати у вигляді:

$$\Delta_1 = \left| \int_a^b f(x) dx - \frac{(b-a)}{2} (f(a) + f(b)) \right| \leq \frac{(b-a)^3}{12} M_2, \quad (12)$$

де $M_2 = \max_{[a,b]} |f''(x)|$.

Оцінка обчислювальної похибки при розрахунках за формулою (10) для випадку, коли значення функції обчислені з однаковою точністю ε , має вигляд:

$$\Delta_2 \leq \frac{b-a}{2} (\varepsilon + \varepsilon) = (b-a)\varepsilon \quad (13)$$

Така система називається зведеною до нормального вигляду.

3) Представимо систему (3.4) в матричному вигляді:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{n2} \\ \alpha_{311} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n11} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix}, \quad (3.5)$$

або векторному

$$\bar{x} = \bar{\beta} + \bar{\alpha} * \bar{x}. \quad (3.6)$$

Якщо деяким чином визначити, так званий, вектор початкових значень $\bar{x}^{(0)}$, який знаходиться в правій частині (3.6), то можна отримати певні значення вектора \bar{x} .

В якості вектора початкових наближень $\bar{x}^{(0)}$ вибирають:

- вектор, в якого всі координати x_i дорівнюють 0;
- вектор, в якого всі координати x_i дорівнюють 1;
- вектор, координати x_i якого дорівнюють координатам вектора вільних членів β_i ;
- координати вектору \bar{x} вибирають в результаті аналізу особливостей об'єкту дослідження та задачі, яка розв'язується.

4) Якщо вектор початкових наближень $\bar{x}^{(0)}$ підставити в праву частину системи (3.5) або (3.6), то вона прийме вигляд:

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \\ \dots \\ x_n^{(1)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{n2} \\ \alpha_{311} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n11} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{bmatrix} * \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ \dots \\ x_n^{(0)} \end{bmatrix}$$

$$\text{або } \bar{x}^{(1)} = \bar{\beta} + \bar{\alpha} * \bar{x}^{(0)},$$

легко розв'язується, тому що в правій частині містить всі визначені елементи, і дозволяє отримати розв'язок системи, який називається *вектором першого наближення* $\bar{x}^{(1)}$.

5) Перевіряється виконання умови закінчення ітераційного процесу пошуку розв'язку системи (3.2) виду:

$$|\bar{x}^{(1)} - \bar{x}^{(0)}| \leq \varepsilon \quad (3.7),$$

де ε - задана похибка результатів розв'язання задачі.

Якщо умова (3.7) не виконується, то $\bar{x}^{(1)}$ підставляється в праву частину (3.5) або (3.6) і знаходиться $\bar{x}^{(2)}$ з системи виду:

$$\begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \\ \dots \\ x_n^{(2)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{bmatrix} * \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \\ \dots \\ x_n^{(1)} \end{bmatrix}$$

або $\bar{x}^{(2)} = \bar{\beta} + \bar{\alpha} * \bar{x}^{(1)}$.

б) Знову перевіряється виконання умови закінчення ітераційного процесу пошуку розв'язку системи (3.2)

$$|\bar{x}^{(2)} - \bar{x}^{(1)}| \leq \varepsilon.$$

Якщо умова не виконується, то $x^{(2)}$ підставляється в праву частину (3.5) і знаходиться $x^{(3)}$ з системи виду:

$$\begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \\ \dots \\ x_n^{(3)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{bmatrix} * \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \\ \dots \\ x_n^{(2)} \end{bmatrix}$$

і т.д.

7) Етапи 4 та 5 повторюються до тих пір поки на якому-небудь k -ому кроці не виконується умова

$$|\bar{x}^{(k)} - \bar{x}^{(k-1)}| \leq \varepsilon. \quad (3.8)$$

Таким чином, процес пошуку розв'язку системи (3.2) наближеними методами з заданою похибкою ε є **ітераційним**, а умовою виходу з цього процесу є умова (3.8).

3.3 Умови збіжності ітераційного процесу

Описаний вище алгоритм дозволяє отримати розв'язок системи (3.2) близький до точного (з заданою похибкою ε) тільки в тому випадку, коли ітераційний процес пошуку розв'язку СЛАР збігається.

Теорема про збіжність. Ітераційний процес пошуку розв'язку системи лінійних алгебраїчних рівнянь виду (3.5) наближеними методами збігається, якщо будь-яка канонічна норма матриці $\|\alpha\| < 1$.

Канонічною нормою матриці називається будь-яке дійсне додатне число, яке визначається за такими умовами:

перша канонічна норма – це максимальна з сум модулів елементів матриці коефіцієнтів α по стрічкам:

$$\|\alpha\|_1 = \max_j \sum_{i=1}^n |\alpha_{ij}|, \quad (3.9)$$

друга канонічна норма – це максимальна з сум модулів елементів матриці коефіцієнтів α по стовбцям:

$$\|\alpha\|_2 = \max_j \sum_{i=1}^n |\alpha_{ij}|, \quad (3.10)$$

третья канонична норма – це корінь квадратний з сум квадратів модулів всіх елементів матриці коефіцієнтів α :

$$\|\alpha\|_3 = \sqrt{\sum_i \sum_j |\alpha_{ij}|^2}$$

Наслідок 1: Ітераційний процес розв'язання системи (3.4) збігається, якщо сума модулів елементів стрічок матриці коефіцієнтів α або сума модулів елементів її стовбців менш одиниці, тобто виконується умова

$$\max_i \sum_{j=1}^n |\alpha_{ij}| < 1 \quad \text{або умова} \quad \max_j \sum_{i=1}^n |\alpha_{ij}| < 1. \quad (3.11)$$

Наслідок 2: Ітераційний процес розв'язання системи (3.4) збігається, якщо елементи головної діагоналі більше суми модулів елементів відповідної стрічки крім діагонального елемента цієї стрічки, тобто виконується умова:

$$|\alpha_{ii}| > \max_j \sum_{j=1}^n |\alpha_{ij}| \quad \text{або умова} \quad |\alpha_{jj}| > \max_i \sum_{i=1}^n |\alpha_{ij}|. \quad (3.12)$$

Приклад.

Визначити, чи збігається ітераційний процес для системи рівнянь

$$\begin{cases} 8x_1 + x_2 + x_3 = 20 \\ x_1 + 5x_2 - x_3 = 7 \\ x_1 - x_2 + 5x_3 = 7 \end{cases} \quad \text{або} \quad \begin{cases} x_1 = 3,25 - 0,125x_2 - 0,125x_3 \\ x_2 = 1,4 - 0,2x_1 + 0,2x_3 \\ x_3 = 1,4 - 0,2x_1 + 0,2x_2 \end{cases}$$

Матриця $\bar{\alpha}$ даної системи має вигляд:

$$\alpha = \begin{bmatrix} 0 & -0,125 & -0,125 \\ -0,2 & 0 & 0,2 \\ -0,2 & 0,2 & 0 \end{bmatrix}$$

Визначаємо норми:

$$\|\alpha\|_1 = \max \begin{cases} |\alpha_{11}| + |\alpha_{12}| + |\alpha_{13}| \\ |\alpha_{21}| + |\alpha_{22}| + |\alpha_{23}| \\ |\alpha_{31}| + |\alpha_{32}| + |\alpha_{33}| \end{cases} = \max \begin{cases} 0 + 0,125 + 0,125 \\ 0,2 + 0 + 0,2 \\ 0,2 + 0,2 + 0 \end{cases} = \max \begin{cases} 0,25 \\ 0,4 \\ 0,4 \end{cases} = 0,4 < 1$$

$$\|\alpha\|_2 = \max \begin{cases} |\alpha_{11}| + |\alpha_{21}| + |\alpha_{31}| \\ |\alpha_{12}| + |\alpha_{22}| + |\alpha_{32}| \\ |\alpha_{13}| + |\alpha_{23}| + |\alpha_{33}| \end{cases} = \max \begin{cases} 0 + 0,2 + 0,2 \\ 0,125 + 0 + 0,2 \\ 0,125 + 0,2 + 0 \end{cases} = \max \begin{cases} 0,4 \\ 0,325 \\ 0,325 \end{cases} = 0,4 < 1$$

Таким чином, перша та друга каноничні норми менш одиниці, тобто ітераційний процес для даної системи збігається.

Розглянемо особливості алгоритмів наближених методів.

3.4 Метод послідовних наближень (метод Якобі)

Нехай задана система лінійних алгебраїчних рівнянь виду (3.2). Метод послідовних наближень (метод Якобі) відноситься до ітераційних методів, тому потребує перетворити дану систему до нормального вигляду (3.5) та знайти каноничні норми матриці $\bar{\alpha}$, для того щоб визначити умови збіжності ітераційного процесу (3.9) - (3.11) пошуку розв'язку системи с заданою похибкою ϵ відповідно теоремі про збіжність. Якщо жодна з умов (3.3) – (3.11) не виконується, то дану систему необхідно перетворити по певним правилам, та знову перевірити умови збіжності ітераційного процесу (3.9) – (3.11). Якщо жодна з умов знову не виконується, то метод послідовних наближень не має сенсу використовувати. Якщо хоча б одна з умов (3.9) –

(3.11) виконується, то ітераційний процес пошуку розв'язку системи с заданою похибкою ε збігається і метод послідовних наближень можна використовувати.

По-перше, вибирається певне значення вектору початкових наближень $\bar{x}^{(0)}$, яке підставляється в праву частину системи рівнянь виду:

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \\ \dots \\ x_n^{(1)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{bmatrix} * \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ \dots \\ x_n^{(0)} \end{bmatrix}, \quad (3.13)$$

що легко розв'язується для знаходження вектора розв'язку першого наближення $\bar{x}^{(1)}$, тому що в правій частині містить всі визначені елементи.

По-друге, перевіряється виконання умови закінчення ітераційного процесу виду:

$$|\bar{x}^{(1)} - \bar{x}^{(0)}| \leq \varepsilon$$

де ε - задана похибка результатів розв'язання задачі. Якщо умова не виконується, то $x^{(1)}$ підставляється в праву частину системи (3.5) і знаходиться $x^{(2)}$:

$$\begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \\ \dots \\ x_n^{(2)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{bmatrix} * \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \\ \dots \\ x_n^{(1)} \end{bmatrix}$$

та знову перевіряється виконання умови закінчення ітераційного процесу виду:

$$|\bar{x}^{(2)} - \bar{x}^{(1)}| \leq \varepsilon.$$

За аналогією будь-яке $(K+1)$ -е наближення можна обчислити за формулою:

$$x^{(\bar{k}+1)} = \bar{\beta} + \bar{\alpha} * x^{(\bar{k})}, \quad \text{де } \bar{k}=0,1,2,\dots \quad (3.14)$$

Якщо послідовність $x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(k)}$, що отримана в результаті ітераційного процесу,

має границю $\lim_{k \rightarrow \infty} x^{(k)}$, то ця границя є розв'язком системи. Умова закінчення ітераційного процесу має вигляд:

$$|x^{(\bar{k}+1)} - x^{(k)}| < \varepsilon, \quad (3.15)$$

де ε - задана похибка результатів розв'язання системи.

Алгоритмічно перевірка умови (3.15) представляє собою алгоритм пошуку максимального відхилення між координатами вектора $\bar{x}^{(k)}$ і $\bar{x}^{(k-1)}$ і порівняння його з заданою похибкою ε .

Алгоритм методу послідовних наближень зображено на рисунку 3.1

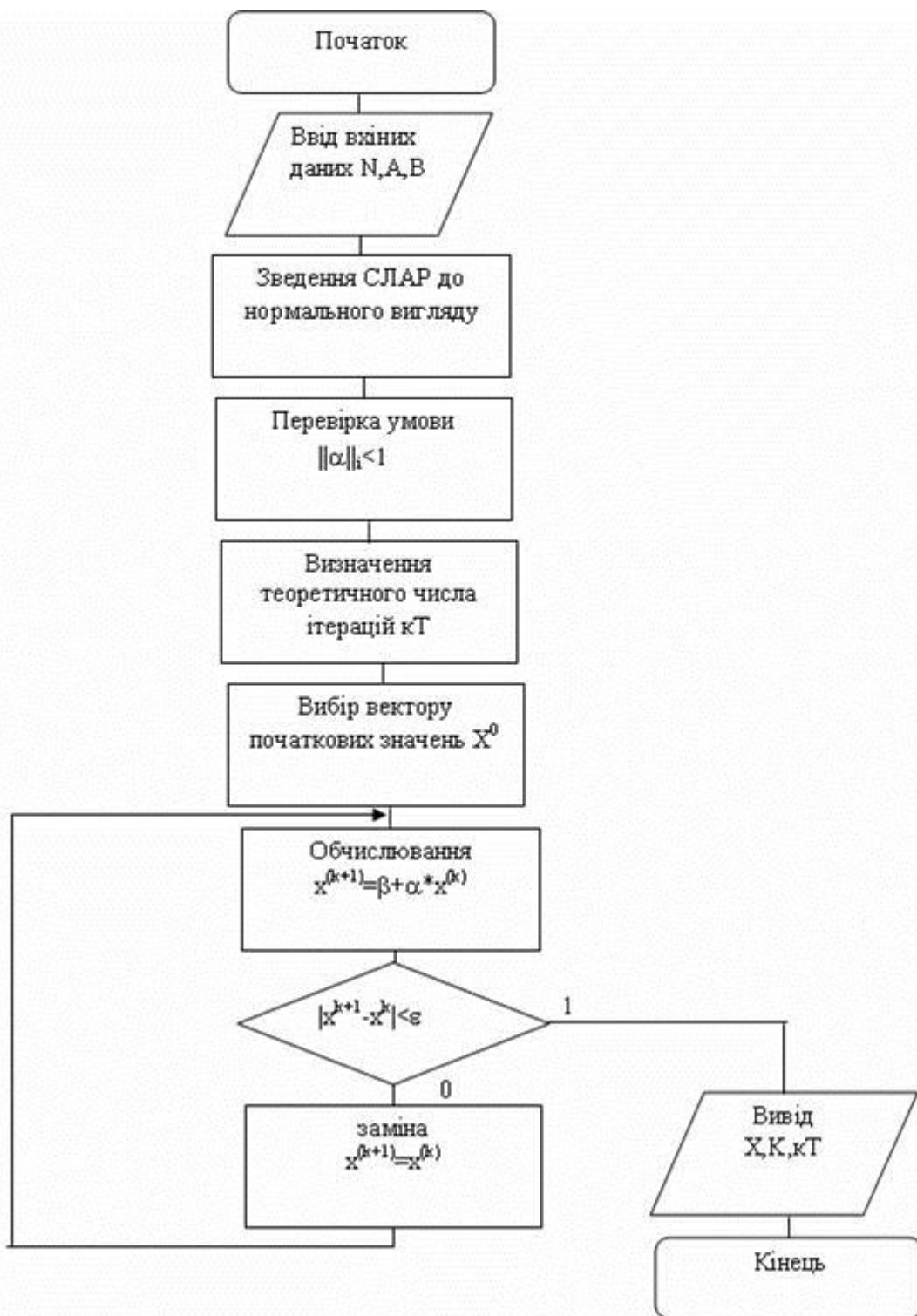


Рисунок 3.1. – Схема алгоритму методу послідовних наближень

Оцінка похибки метода Якобі

Якщо задана допустима похибка обчислень ε і x - вектор точного розв'язку системи лінійних рівнянь, а $x_j^{(k)}$ - k -те наближення до вектору точного розв'язку, то для оцінки похибки методу послідовних наближень використовується формула:

$$\|x_j - x_j^{(k)}\| \leq \frac{\|\alpha\|^{(k+1)}}{1 - \|\alpha\|} \|\beta\|, \quad (3.16)$$

де $\|\alpha\|$ - одна з 3 норм матриці α ; $\|\beta\|$ - аналогічна норма вектора β ; k - кількість ітерацій, необхідна для досягнення потрібної точності ε .

3.5 Метод Гауса-Зейделя

Метод Зейделя являє собою модифікацію метода послідовних наближень, при чому у методі Зейделя при обчисленні i -ої координати вектора розв'язку $(k+1)$ -го наближення використовуються значення всіх $(i-1)$ координат вектора $(k+1)$ -го наближення обчислені раніше. Розглянемо метод більш детально.

Нехай початкова система лінійних алгебраїчних рівнянь приведена до нормального вигляду:

$$\begin{cases} x_1 = \beta_1 + \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n; \\ x_2 = \beta_2 + \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n; \\ \dots \\ x_n = \beta_n + \alpha_{n1}x_1 + \alpha_{n2}x_2 + \dots + \alpha_{nn}x_n. \end{cases} \quad (3.17)$$

Алгоритм Гауса-Зейделя

Вибрати значення координат вектора початкових наближень $\bar{x} = \{x_1^{(0)}, \dots, x_n^{(0)}\}$

Визначити значення першої координати $x_1^{(1)}$ вектора першого наближення з першого рівняння системи:

$$x_1^{(1)} = \beta_1 + \alpha_{11}x_1^{(0)} + \alpha_{12}x_2^{(0)} + \dots + \alpha_{1n}x_n^{(0)}.$$

Підставити в друге рівняння системи значення першої координати $x_1^{(1)}$, яке обчислено на попередньому кроці

$$x_2^{(1)} = \beta_2 + \alpha_{21}x_1^{(1)} + \alpha_{22}x_2^{(0)} + \dots + \alpha_{2n}x_n^{(0)}.$$

Отримані значення координат першого наближення $x_1^{(1)}, x_2^{(1)}$ підставляємо у третє рівняння системи (3.14)

$$x_3^{(1)} = \beta_3 + \alpha_{31}x_1^{(1)} + \alpha_{32}x_2^{(1)} + \alpha_{33}x_3^{(0)} + \dots + \alpha_{3n}x_n^{(0)}$$

для знаходження третій координати і т.д.

1. Для знаходження останньої координати вектора першого наближення $x_n^{(1)}$ в останнє рівняння системи треба підставити значення всіх $(n-1)$ координат $(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \dots, x_{n-1}^{(1)})$, які отримані на попередніх кроках та значення координати $x_n^{(0)}$

$$x_n^{(1)} = \beta_n + \alpha_{n1}x_1^{(1)} + \alpha_{n2}x_2^{(1)} + \dots + \alpha_{n,n-1}x_{n-1}^{(1)} + \alpha_{nn}x_n^{(0)}.$$

6. Аналогічно будують друге, третє та інші наближення. Так для вектора $(k+1)$ -го наближення за методом Зейделя використовують наступні формули:

$$\begin{cases} x_1^{(k+1)} = \beta_1 + \sum_{j=1}^n \alpha_{1j}x_j^{(k)}; \\ x_2^{(k+1)} = \beta_2 + \sum_{j=2}^n \alpha_{2j}x_j^{(k)} + \alpha_{21}x_1^{(k+1)}; \\ x_3^{(k+1)} = \beta_3 + \alpha_{31}x_1^{(k+1)} + \alpha_{32}x_2^{(k+1)} + \sum_{j=3}^n \alpha_{3j}x_j^{(k)}; \\ \dots \\ x_n^{(k+1)} = \beta_n + \sum_{j=1}^{n-1} \alpha_{nj}x_j^{(k+1)} + \alpha_{nn}x_n^{(k)}. \end{cases} \quad (3.18)$$

Умови збіжності ітераційного процесу Зейделя

Даний процес розв'язання СЛАР - ітераційний, тому важливим є аналіз умов збіжності ітераційного процесу. Процес Зейделя для системи лінійних рівнянь $\bar{x} = \bar{b} + \bar{a}x$ збігається до точного розв'язку с заданою похибкою ε при будь-якому виборі вектора початкових наближень, якщо будь яка норма матриці \bar{a} менша 1, тобто якщо:

$$\|\alpha\|_1 = \max_j \sum_{i=1}^n |\alpha_{ij}| < 1 \quad (3.19)$$

$$\|\alpha\|_2 = \max_i \sum_{j=1}^n |\alpha_{ij}| < 1 \quad (3.20)$$

$$\|\alpha\|_3 = \sqrt{\sum_i \sum_j |\alpha_{ij}|^2} < 1 \quad (3.21)$$

Відомо, що процес Зейделя сходиться до точного розв'язку СЛАР швидше ніж метод послідовних наближень.

Приклад. Знайти першу та другу норми та проаналізувати умови збіжності ітераційного процесу для матриці α , яка має вигляд :

$$\bar{\alpha} = \begin{bmatrix} 0,24 & -0,05 & -0,24 \\ -0,22 & 0,09 & -0,44 \\ 0,13 & -0,02 & 0,42 \end{bmatrix}$$

$$\|\alpha\|_{\infty} = \max_j \sum_{i=1}^n |\alpha_{ij}| = \max\{0,53; 0,75; 0,57\} = 0,75 < 1.$$

Очевидно, що процес ітерації для даної системи сходиться до точного розв'язку, не дивлячись на те, що

$$\|\alpha\|_1 = \max_i \sum_{j=1}^n |\alpha_{ij}| = \max\{0,59; 0,16; 1,1\} = 1,1 > 1.$$

Оцінка похибки методу Гауса-Зейделя

Якщо \bar{x} - точне значення вектора розв'язку системи лінійних рівнянь; а $\bar{x}^{(k)}$ - к-е наближення, обчислене за методом Гауса-Зейделя, то для оцінки похибки цього метода використовується формула:

$$\|\bar{x} - \bar{x}^{(k)}\|_1 \leq \frac{\|\alpha\|_1^{(k)}}{1 - \|\alpha\|_1} \|\bar{x}^{(1)} - \bar{x}^{(0)}\|_1 \quad (3.22)$$

3.6 Метод верхньої релаксації

В основі метода верхньої релаксації використовується алгоритм та обчислювальна схема метода Гауса-Зейделя, але на відміну від нього нові значення координат вектора к-го наближення визначаються за формулами:

$$x_i^{(k+1)} = x_i^{(k)} + \omega(\bar{x}_i^{(k+1)} - x_i^{(k)}), \quad (3.23)$$

де $\bar{x}_i^{(k+1)}$ - уточнене значення змінної по методу Гауса-Зейделя, ω - параметр релаксації, значення якого визначається з інтервалу $1 \leq \omega \leq 2$. При $\omega=1$ метод тотожний методу Гауса-Зейделя. Швидкість збіжності ітераційного процесу залежить від значення ω .

ЧИСЕЛЬНІ МЕТОДИ РОЗ'В'ЯЗАННЯ СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ

Достатньо велика кількість інженерних задач на проміжному етапі вирішення зводиться до вирішення системи нелінійних рівнянь. Це одна з найважчих задач з точки зору реалізації її на ЕОМ. Одним із найбільш простих алгоритмів її рішення є метод Ньютона. Це найбільш розповсюджений метод розв'язання систем нелінійних рівнянь. Його популярність обумовлена тим, що в порівнянні з методом простої ітерації він забезпечує найбільш швидку збіжність. В основі методу Ньютона лежить представлення всіх n рівнянь у вигляді рядів Тейлора.

Розглянемо алгоритм методу Ньютона.

Нехай дана система нелінійних рівнянь виду

$$\begin{cases} f_1(x) = 0, \\ f_2(x) = 0, \\ \dots\dots\dots \\ f_m(x) = 0, \end{cases} \quad (4.20)$$

де f_1, f_2, \dots, f_m – неперервно-диференційні функції.

2. Алгоритм методу базується на розкладі кожної функції системи в околі точки з координатами $\bar{X} = \{x_1, x_2, \dots, x_n\}$ в ряд Тейлора.

$$f_i(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) = f_i(x_1, x_2, \dots, x_n) + \Delta x_1 \frac{\partial f_i}{\partial x_1} + \Delta x_2 \frac{\partial f_i}{\partial x_2} + \dots + \Delta x_n \frac{\partial f_i}{\partial x_n} +$$

порядків (f^1, f^2 тощо).

1. Початкова система буде мати вигляд:

$$\begin{cases} f_1(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) = f_1(x_1, x_2, \dots, x_n) + \Delta x_1 \frac{\partial f_1}{\partial x_1} + \Delta x_2 \frac{\partial f_1}{\partial x_2} + \dots + \Delta x_n \frac{\partial f_1}{\partial x_n}, \\ f_2(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) = f_2(x_1, x_2, \dots, x_n) + \Delta x_1 \frac{\partial f_2}{\partial x_1} + \Delta x_2 \frac{\partial f_2}{\partial x_2} + \dots + \Delta x_n \frac{\partial f_2}{\partial x_n}, \\ \dots\dots\dots \\ f_m(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) = f_m(x_1, x_2, \dots, x_n) + \Delta x_1 \frac{\partial f_m}{\partial x_1} + \Delta x_2 \frac{\partial f_m}{\partial x_2} + \dots + \Delta x_n \frac{\partial f_m}{\partial x_n}. \end{cases} \quad (4.21)$$

4.21)

2. Припустимо, що прирости Δx_i вибрані таким чином, що точки з координатами $x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n$ є коренями даної системи рівнянь з заданим ступенем наближення ϵ . Тоді ліву частину рівнянь системи (4.21) можна прирівняти до нуля, тобто система рівнянь (4.21) буде мати вигляд:

$$\begin{cases} \frac{\partial f_1}{\partial x_1} \Delta x_1 + \frac{\partial f_1}{\partial x_2} \Delta x_2 + \dots + \frac{\partial f_1}{\partial x_n} \Delta x_n = -f_1(x_1, x_2, \dots, x_n), \\ \frac{\partial f_2}{\partial x_1} \Delta x_1 + \frac{\partial f_2}{\partial x_2} \Delta x_2 + \dots + \frac{\partial f_2}{\partial x_n} \Delta x_n = -f_2(x_1, x_2, \dots, x_n), \\ \dots\dots\dots \\ \frac{\partial f_m}{\partial x_1} \Delta x_1 + \frac{\partial f_m}{\partial x_2} \Delta x_2 + \dots + \frac{\partial f_m}{\partial x_n} \Delta x_n = -f_m(x_1, x_2, \dots, x_n). \end{cases} \quad (4.22)$$

Або в матричній формі система (4.22) буде мати вигляд:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \times \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_n \end{bmatrix} = \begin{bmatrix} -f_1(x_1, x_2, \dots, x_n) \\ -f_2(x_1, x_2, \dots, x_n) \\ \dots \\ -f_m(x_1, x_2, \dots, x_n) \end{bmatrix} \quad (4.23)$$

де $\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$ – матриця Якобі.

3. В результаті таких перетворень система рівнянь може розглядатися як система лінійних алгебраїчних рівнянь відносно Δx_i . В такому випадку, якщо врахувати, що заданий вектор x початкових наближень виду:

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \dots \\ x_n^{(0)} \end{bmatrix},$$

можливо розв'язувати систему відносно вектора приросту Δx_i , та знайти розв'язок системи, як сума попереднього значення та вектора Δx_i :

$$\begin{cases} x_1^{(1)} = x_1^{(0)} + \Delta x_1, \\ x_2^{(1)} = x_2^{(0)} + \Delta x_2, \\ \dots, \\ x_n^{(1)} = x_n^{(0)} + \Delta x_n. \end{cases} \quad (4.24)$$

Дану задачу можна розв'язати з будь-якої точки, вибравши вектор початкових наближень.

4. Процес розв'язання системи нелінійних рівнянь (4.20) з використанням системи лінійних алгебраїчних рівнянь (4.23) відносно Δx_i - ітераційний, та буде продовжуватись до тих пір, поки всі координати вектору приростів не стануть менше за абсолютною величиною заданої похибки ε , тобто $|\Delta x| \leq \varepsilon$.

Схема алгоритму метода Ньютона представлена на рисунку 4.20.

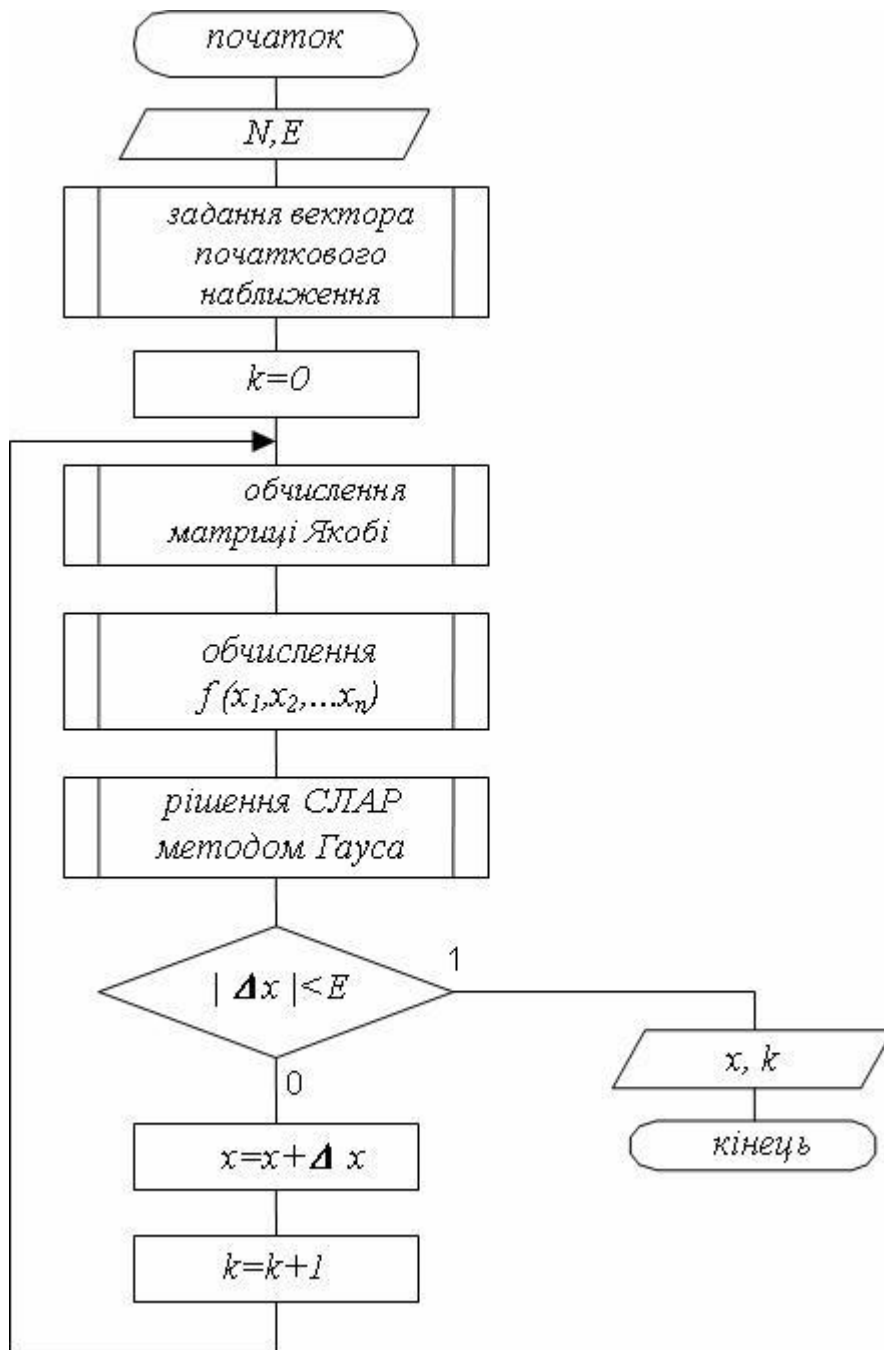


Рисунок 4.20 – Схема алгоритму розв'язання систем нелінійних рівнянь методом Ньютона

Приклад. Стійкість складної конструкції жорстких балок описується системою нелінійних рівнянь виду:

$$\begin{cases} x_1 + 2x_2 + x_3 + 4x_4 = 20,700 \\ x_1^2 + 2x_1x_2 + x_4^3 = 15,880 \\ x_1^3 + x_3^2 + x_4 = 21,218 \\ 3x_2 + x_3x_4 = 7,900 \end{cases}$$

Розв'язати дану систему нелінійних рівнянь методом Ньютона.

Розв'язок. Для того, щоб розв'язати дану систему нелінійних рівнянь методом Ньютона необхідно перетворити систему до вигляду (4.23). Для цього знайдемо спочатку частині похідні всіх рівнянь системи в початковій точці для формування матриці Якобі, яка буде мати вигляд:

$$J = \begin{bmatrix} 1 & 2 & 1 & 4 \\ 2x_1 + 2x_2 & 2x_1 & 0 & 3x_4^2 \\ 3x_1^2 & 0 & 2x_3 & 1 \\ 0 & 3 & x_4 & x_3 \end{bmatrix}$$

За допомогою програми, що складена відповідно алгоритму, який представлений на рисунку 4.20, для знаходження вектора розв'язку X багаторазово методом Гауса розв'яжемо на ЕОМ систему виду:

$$J = \begin{bmatrix} 1 & 2 & 1 & 4 \\ 2x_1 + 2x_2 & 2x_1 & 0 & 3x_4^2 \\ 3x_1^2 & 0 & 2x_3 & 1 \\ 0 & 3 & x_4 & x_3 \end{bmatrix} \times \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_n \end{bmatrix} = \begin{bmatrix} -1x_1 - 2x_2 - x_3 - 4x_4 + 20,700 \\ -1x_1^2 - 2x_1x_2 - x_4^3 + 15,880 \\ -x_1^3 - x_3^2 - x_4 + 21,218 \\ -3x_2 - x_3x_4 + 7,900 \end{bmatrix}$$

відносно вектору поправки $\Delta \bar{x}$ з початковими значеннями змінних $x_1 = x_2 = x_3 = x_4 = 1,0$ та з заданою похибкою $\varepsilon = 0,0001$, з яких починається ітераційний процес.

В результаті можна отримати наступні результати:

Номер ітерації	x_1	x_2	x_3	x_4
0	1,0000	1,0000	1,0000	1,0000
1	2,75037	4,6763	7,8958	0,1753
2	1,3448	5,2971	5,9494	0,7029
3	1,4775	3,8437	4,3419	1,7983
4	1,5427	6,2434	4,1204	0,6376
5	1,2364	5,7274	4,3436	0,9163
6	1,2024	5,5986	4,2995	1,0002
7	1,2000	5,6000	4,3000	1,0000
8	1,2000	5,6000	4,3000	1,0000

ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ НА ЕОМ

Інженеру часто доводиться зіштовхуватись з диференціальними рівняннями і системами диференціальних рівнянь при розробці нових виробів чи технологічних процесів, так як більша частина законів фізики формалізується саме у вигляді диференціальних рівнянь. Будь-яка задача проектування, яка зв'язана з розрахунком потоків енергії чи руху тіл, в кінцевому рахунку зводиться до розв'язку диференціальних рівнянь. Нажаль, лише дуже малу частину з них можливо вирішити без допомоги обчислювальних машин. Тому чисельні методи розв'язку диференціальних рівнянь відіграють важливу роль у практиці інженерних розрахунків.

9.1. Основні визначення та поняття

Рівняння, у якому невідома функція входить під знаком похідної чи диференціала, називається *диференціальним рівнянням*. Наприклад,

$$\frac{dy}{dx} = 2(y-3); \quad \frac{d^2y}{dt^2} = t+1; \quad \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$
$$y' = x^2; \quad xdy = y^3 dx$$

Якщо невідома функція, що входить у диференціальне рівняння, залежить тільки від однієї незалежної змінної, то диференціальне рівняння називається *звичайним*. Наприклад, диференціальні рівняння

$$x^2 \cdot \frac{d^2y}{dx^2} = 2; \quad 2s dt = t ds$$

відносяться до звичайних.

Якщо ж невідома функція, що входить у диференціальне рівняння, є функцією двох чи більшого числа незалежних змінних, то таке рівняння називається *диференціальним рівнянням у частинних похідних*. Наприклад, диференціальне рівняння

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$

відноситься до рівняння в частинних похідних.

Порядком диференціального рівняння називається найвищий порядок похідної (чи диференціала), що входить у рівняння.

Розглянемо звичайні диференціальні рівняння.

Звичайне диференціальне рівняння n -го порядку в самому загальному випадку містить незалежну змінну, невідому функцію і її похідні чи диференціали до n -го порядку включно і має вид

$$F(x, y, y', y'', \dots, y^{(n)}) = 0 \quad (9.1)$$

У цьому рівнянні x - незалежна змінна, y - невідома функція, $y', y'', \dots, y^{(n)}$ - похідні цієї функції.

Розв'язком (чи *інтегралом*) рівняння (9.1) називається будь-яка диференціюєма функція $y = \varphi(x)$, що задовольняє цьому рівнянню, тобто така, після підстановки, якої у рівняння (9.1) воно перетворюється в тотожність.

Графік розв'язку звичайного диференціального рівняння називається *інтегральною кривою* цього рівняння.

Розв'язок диференціального рівняння, що містить стільки незалежних довільних (постійних) параметрів, який його порядок, називається *загальним розв'язком* (чи *загальним інтегралом*) цього рівняння.

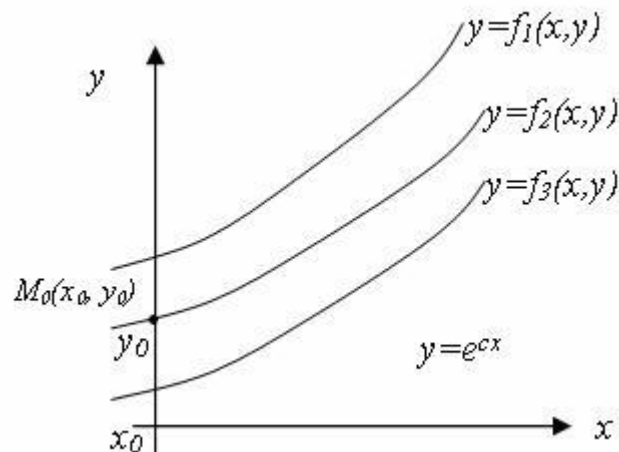


Рисунок 9.1. – Сімейство інтегральних кривих диференціального рівняння (9.1)

Геометрично загальний розв'язок диференціального рівняння являє собою сімейство інтегральних кривих цього рівняння (рис. 9.1).

Частинним розв'язком диференціального рівняння називається будь-який розв'язок, що може бути отриманий з загального при визначених числових значеннях довільних постійних (рис. 9.1). Довільні постійні, що входять в загальний розв'язок, визначаються з *початкових* або *крайових умов*.

Задача з *початковими умовами* ставиться так: знайти розв'язок $y = \varphi(x)$ рівняння $y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)})$, що задовольняє додатковим умовам, які складаються в тім, що розв'язок $y = \varphi(x)$, повинний приймати разом зі своїми похідними до $(n-1)$ -го порядку задані числові значення $y_0, y'_0, y''_0, \dots, y_0^{(n-1)}$ при заданому числовому значенні $x = x_0$ незалежної змінної x .

Такі умови називаються *початковими умовами*, а задача відшукування розв'язку $y = \varphi(x)$ диференціального рівняння (9.1), що задовольняє заданим початковим умовам - *задачею з початковими умовами*, або *задачею Коші*.

Задача з *крайовими умовами* ставиться так: знайти розв'язок $y = \varphi(x)$ рівняння $y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)})$, що задовольняє додатковим умовам, які складаються в тім, що розв'язок $y = \varphi(x)$, повинний приймати разом зі своїми похідними до $(n-1)$ -го порядку задані числові значення $y_0, y'_0, y''_0, \dots, y_0^{(n-1)}$ при заданому числовому значенні $x = x_0$ та $y_n, y'_n, y''_n, \dots, y_n^{(n-1)}$ при заданому числовому значенні $x = x_n$ незалежної змінної x .

Такі умови називаються *крайовими умовами*, а задача відшукування розв'язку $y = \varphi(x)$ диференціального рівняння (9.1), що задовольняє заданим крайовим умовам – *крайовою задачею*.

У випадку рівняння першого порядку, тобто при $n=1$, одержуємо задачу Коші для рівняння $y' = f(x, y)$ з початковою умовою $x = x_0, y = y_0$.

Геометрично задача Коші для рівняння першого порядку полягає в тому, що з усіх інтегральних кривих, що представляють собою загальний розв'язок, потрібно знайти ту інтегральну криву, що проходить через точку M_0 з координатами $x = x_0, y = y_0$. (рис.9.1).

Часто в задачі Коші у ролі незалежної змінної виступає час t . Прикладом може бути задача про вільні коливання тіла, яке підвішене на пружині. Рухи такого тіла описуються диференціальним рівнянням, в якому незалежною змінною є час t . Якщо додаткові умови задані у вигляді значень переміщень чи швидкості при $t=0$, то це також задача Коші.

Задача Коші має єдиний розв'язок, що задовольняє умові $v(x_0) = y_0$, якщо функція $f(x, y)$ неперервна в деякій області $R_{[a,b]} = \{|x - x_0| < a, |y - y_0| < b\}$ і задовольняє в цій області умові Ліпшица,

$$|f(x, \bar{y}) - f(x, y)| \leq N|\bar{y} - y|,$$

де N - постійна Ліпшица, що залежить від a і b (a і b - границі області).

Методи точного інтегрування диференціальних рівнянь придатні лише для порівняно невеликої частини рівнянь, що зустрічаються на практиці.

Тому в задачах моделювання та дослідження складних технічних систем, наприклад, систем автоматичного управління, великого значення набувають методи наближеного розв'язання диференціальних рівнянь, що у залежності від форми представлення розв'язку можна розділити на дві групи:

1) **аналітичні методи**, що дають наближений розв'язок диференціального рівняння у виді аналітичного виразу;

2) **чисельні методи**, що дають наближений розв'язок у вигляді таблиці.

Похибки. Перед тим, як перейти до розглядання методів чисельного розв'язання диференціальних рівнянь, зупинимось на джерелах похибок, пов'язаних з чисельною апроксимацією. Таких джерел три:

1. **Похибка округлення** зумовлена обмеженнями на представлення чисел в ЕОМ, тому що число значущих цифр, що запам'ятовується і використовується в обчисленнях, обмежене.

2. **Похибка відсічення** зв'язана з тим, що для апроксимації функції замість

$$y = y_0, \quad y' = y_0', \quad y'' = y_0'', \dots, \quad y^{(n-1)} = y_0^{(n-1)} \quad \text{при } x = x_0 \quad (9.2)$$

нескінчених рядів часто використовується лише декілька перших їх членів.

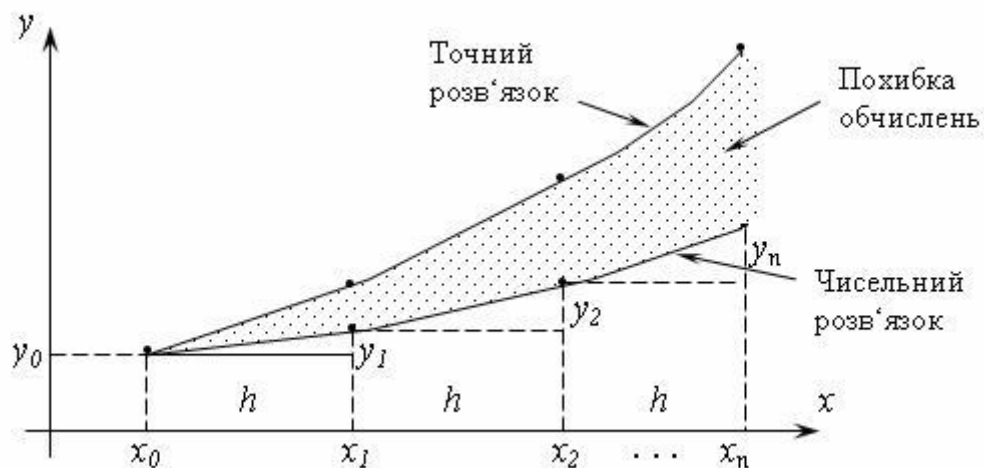


Рисунок 9.2 – Геометричне представлення накопичування похибки в процесі обчислень

Це звичайний для чисельних методів прийом, що являється джерелом похибки, цілком зумовлених використаним методом і не залежать від характеристик ЕОМ.

3. **Похибка поширення** являється результатом накопичення похибок, що з'явилися у попередніх результатах розрахунку. Так як ні один з наближених методів не може

дати зовсім точних результатів, то будь-яка виникла в процесі обчислень похибка зберігається і на наступних стадіях розрахунку (рис. 9.2).

Вказані три джерела похибок є причиною помилок двох типів:

Локальна помилка – сума похибок, що вносяться у розрахунковий процес на кожному етапі обчислення.

Глобальна помилка – різниця між розрахованим та точним значеннями величини на кожному етапі реалізації чисельного алгоритму, що визначає сумарну похибку, що накопичується з моменту початку розрахунку.

9.2. Класифікація чисельних методів розв'язання задачі Коші

На протязі багатьох років чисельний розв'язок задачі Коші був об'єктом пильної уваги науковців оскільки він широко застосовується в різних галузях науки і техніки. Тому і кількість розроблених для нього методів дуже велика.

Чисельні методи розв'язання задачі Коші розділяються на 3 групи:

- одноточкові;
- багатоточкові (методи прогнозу та корекції);
- методи з автоматичним вибором кроку інтегрування.

На рис. 9.3 представлена класифікація найбільш відомих чисельних методів розв'язання диференціальних рівнянь (ДР) на ЕОМ.



Рисунок 9.3– Класифікація чисельних методів розв'язання задачі Коші.

До одноточкових методів відносять методи, які мають певні загальні риси, такі як:

1. В основі усіх одноточкових методів лежить розклад функції в ряд Тейлора, в якому зберігаються члени, що мають h в степені до k включно. Ціле число k називається порядком метода. Похибка на кроці має порядок $k+1$.

2. Всі одноточкові методи не потребують дійсного обчислення похідних, тому що обчислюється лише сама функція, однак можуть потребуватися її значення в деяких проміжних точках. Це тягне за собою, звичайно, додаткові затрати часу і зусиль.

3. Для отримання інформації у новій точці, потрібно мати дані лише в попередній точці. Цю властивість можна назвати „самостартуванням”. Властивість „самостартування” дозволяє легко змінювати величину кроку h .

4. В порівнянні з одноточковими методами методи прогнозу і корекції мають ряд особливостей:

1. Для реалізації методів прогнозу і корекції необхідно мати інформацію про декілька попередніх точок (вони не відносяться до „самостартуючих” методів), тому для отримання додаткової інформації доводиться застосовувати одноточковий метод. Якщо в процесі розв’язку диференціальних рівнянь методом прогнозу і корекції змінюється крок, то звичайно тимчасово доводиться переходити до одноточкового методу.

2. Одноточкові методи і методи прогнозу і корекції забезпечують приблизно однакову точність результатів. Однак другі на відміну від перших дозволяють лише оцінити похибку на кроці. З цієї причини, користуючись одноточковими методами, величину кроку h звичайно обирають трохи менше, ніж це необхідно, тому методи прогнозу і корекції виявляються найбільш ефективними.

3. Використовуючи метод Рунге-Кутта четвертого порядку точності, на кожному кроці доводиться обчислювати чотири значення функції, але для збіжності методу прогнозу і корекції того ж порядку точності часто достатньо двох значень функції. Тому методи прогнозу і корекції вимагають майже вдвічі менше машинного часу, ніж методи Рунге-Кутта порівнюваної точності.

9.3 Одноточкові методи розв’язання задачі Коші на ЕОМ

Розв’язати диференціальне рівняння $y' = f(x, y)$ чисельним методом - це значить для заданої послідовності аргументів x_0, x_1, \dots, x_n і y_0 знайти такі значення y_0, y_1, \dots, y_n що $y_i = F(x_i)$ ($i = 1, 2, \dots, n$) і $F(x_0) = y_0$. Таким чином, чисельні методи дозволяють замість отримання функції $y = F(x)$ одержати таблицю значень цієї функції для заданої послідовності аргументів. Величина $h = x_i - x_{i-1}$ називається *кроком інтегрування*.

Графічно чисельний розв’язок уявляє собою послідовність коротких прямолінійних відрізків, якими апроксимується аналітичний розв’язок $y = F(x)$ рівняння (кусочно-лінійна апроксимація).

Розглянемо алгоритми найбільш відомих чисельних методів.

Метод Ейлера є порівняно грубим і застосовується в основному для орієнтованих розрахунків. Однак ідеї, покладені в основу методу Ейлера, є базовими для інших методів.

Нехай дано диференціальне рівняння першого порядку

$$y' = f(x, y) \quad (9.3)$$

$$\text{з початковими умовами } x = x_0, \quad y(x_0) = y_0. \quad (9.4)$$

Потрібно знайти розв’язок рівняння на відрізку $[x_0, x_n]$.

Розіб’ємо відрізок $[x_0, x_n]$ на n рівних частин і одержимо послідовність $x_0, x_1, x_2, \dots, x_n$, де $x_i = x_0 + ih$ ($i = 0, 1, 2, \dots, n$), а $h = (x_n - x_0)/n$ - крок інтегрування.

Виберемо k -й відрізок $[x_k, x_{k+1}]$ і проінтегруємо рівняння (9.3):

$$\int_{x_k}^{x_{k+1}} f(x, y) dx = \int_{x_k}^{x_{k+1}} y' dx = y(x) \Big|_{x_k}^{x_{k+1}} = y(x_{k+1}) - y(x_k) = y_{k+1} - y_k.$$

або

$$y_{k+1} = y_k + \int_{x_k}^{x_{k+1}} f(x, y) dx. \quad (9.5)$$

Якщо в останньому інтегралі підінтегральну функцію на відрізку $[x_k, x_{k+1}]$ прийняти постійною і рівною початковому значенню в точці $x = x_k$, то одержимо

$$\int_{x_k}^{x_{k+1}} f(x_k, y_k) dx = f(x_k, y_k) \cdot x \Big|_{x_k}^{x_{k+1}} = f(x_k, y_k)(x_{k+1} - x_k) = y'_k h$$

Тоді формула (6.5) прийме вигляд

$$y_{k+1} = y_k + y'_k h \quad (9.6)$$

Позначивши $y_{k+1} - y_k = \Delta y_k$, тобто $y'_k h = \Delta y_k$, отримаємо

$$y'_k h = \Delta y_k \quad (6.7)$$

Продовжуючи цей процес, і щораз приймають, що на відрізку $[x_k, x_{k+1}]$ інтегральна крива $y = F(x)$ приблизно замінюється прямолінійним відрізком, що виходить із точки $M_k(x_k, y_k)$ кутовим коефіцієнтом $f(x_k, y_k)$. Тому в якості наближення шуканої інтегральної кривої одержуємо ламану лінію з вершинами в точках $M_0(x_0, y_0), M_1(x_1, y_1), \dots, M_n(x_n, y_n)$ (рис. 9.4).

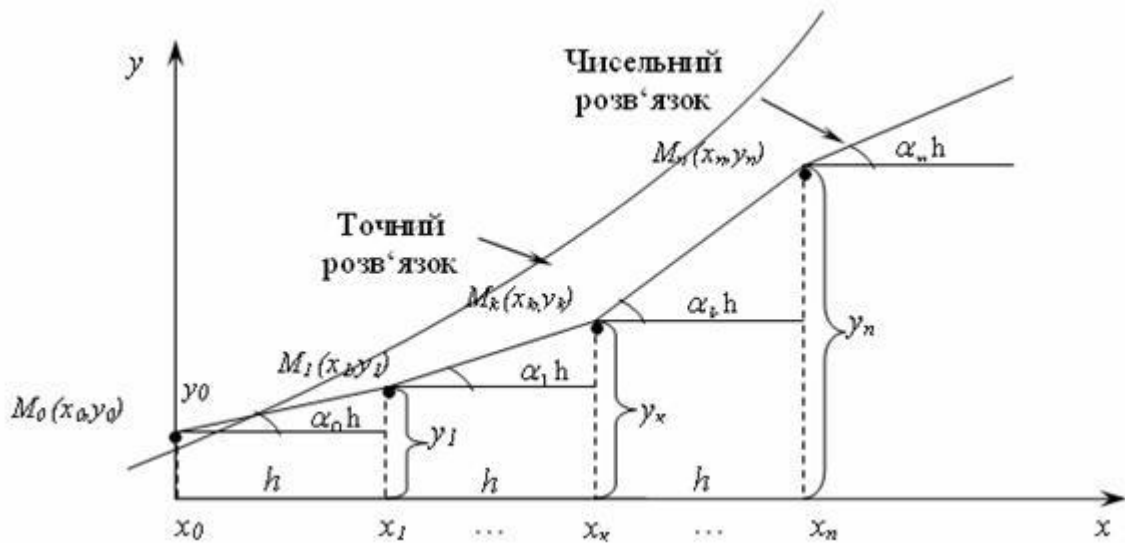


Рисунок 9.4 – Геометричне розв'язання методом Ейлера

Якщо функція $f(x, y)$ у деякій прямокутній області $R\{|x - x_0| \leq a, |y - y_0| \leq b\}$

задовольняє умові

$$|f(x, y_1) - f(x, y_2)| \leq N |y_1 - y_2| \quad (N = \text{const}) \quad (9.8)$$

і, крім того,

$$\left| \frac{df}{dx} \right| = \left| \frac{df}{dx} + f \frac{df}{dy} \right| < M \quad (M = \text{const}), \quad (9.9)$$

то має місце наступна оцінка похибки:

$$|y(x_n) - y_n| \leq \frac{hM}{2N} [(1 + hN)^n - 1], \quad (9.10)$$

де $y(x_n)$ - значення точного розв'язку рівняння (9.3) при $x = x_n$, а y_n - наближене значення, отримане на n -м кроці.

Формула (9.10) має в основному теоретичне застосування. На практиці, як правило, застосовують "подвійний прорахунок". Спочатку чисельне розв'язання рівняння ведеться з кроком h , потім крок дроблять і повторний розрахунок ведеться з кроком $h/2$. Похибка більш точного значення оцінюється формулою

$$|y_n^* - y(x_n)| \approx |y_n^* - y_n|. \quad (9.11)$$

Метод Ейлера може бути застосований до розв'язку систем диференціальних рівнянь вищих порядків. Однак в останньому випадку диференціальні рівняння повинні бути приведені до системи диференціальних рівнянь першого порядку.

Нехай задана система двох рівнянь першого порядку

$$\begin{cases} y' = f_1(x, y, z), \\ z' = f_2(x, y, z) \end{cases} \quad (9.12)$$

з початковими умовами

$$y(x_0) = y_0, \quad z(x_0) = z_0. \quad (9.13)$$

Наближені значення $y(x_i)$ y_i та $z(x_i)$ z_i знаходяться по формулах

$$\begin{cases} y_{i+1} = y_i + \Delta y_i, \\ z_{i+1} = z_i + \Delta z_i, \end{cases} \quad (9.14)$$

$$\Delta y_i = h f_1(x_i, y_i, z_i), (i = 0, 1, 2, \dots)$$

$$\Delta z_i = h f_2(x_i, y_i, z_i) \quad (9.15)$$

Схема алгоритму метода Ейлера представлена на рисунку 9.5.

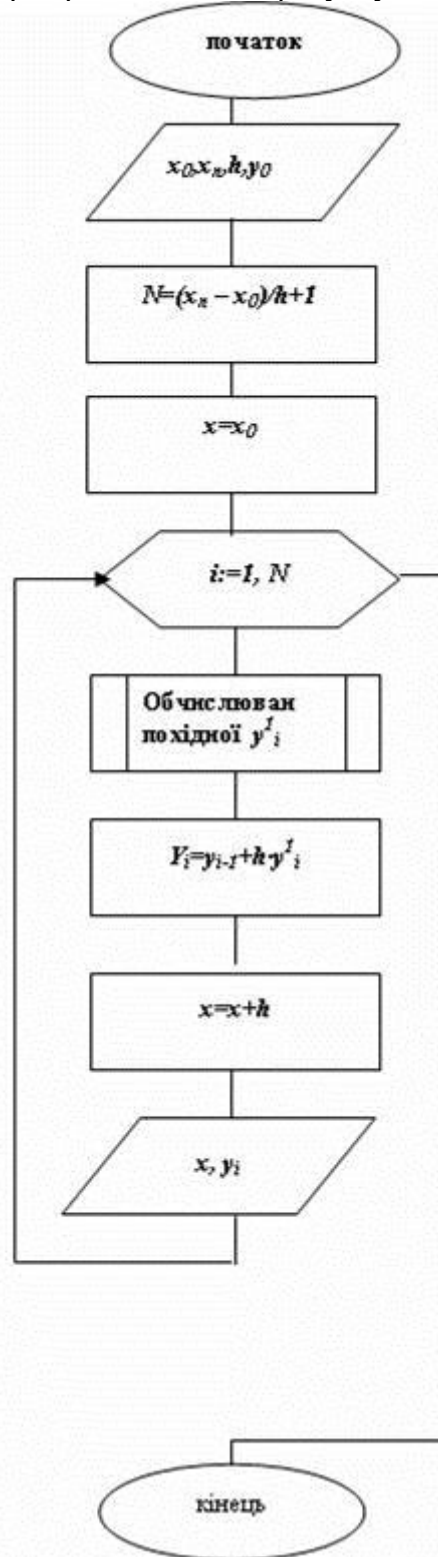


Рисунок 9.5 – Схема алгоритму метода Ейлера

Схема алгоритму метода Ейлера для розв'язання системи звичайних диференціальних рівнянь наведена на рисунку 9.6.

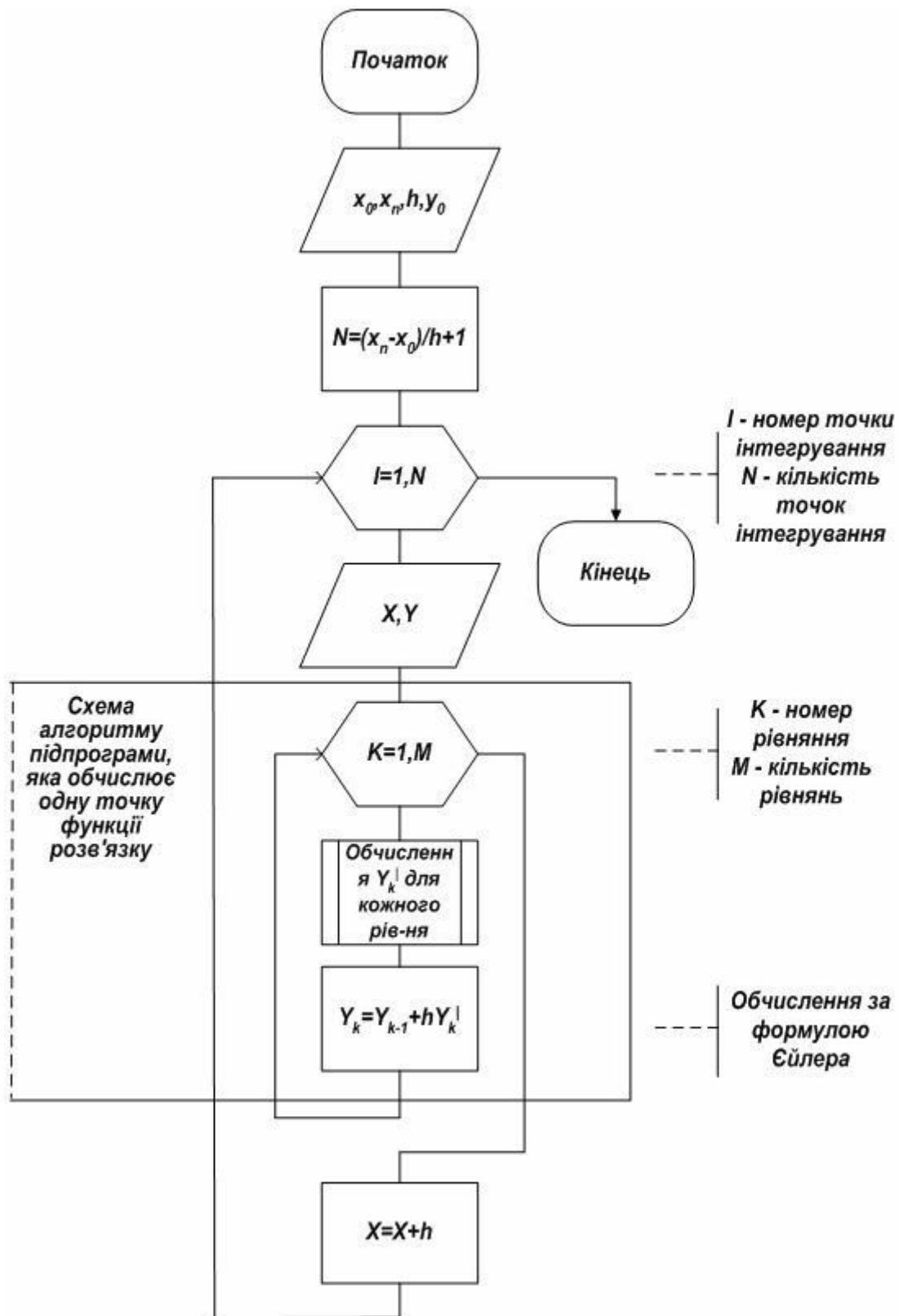


Рисунок 9.6 – Схема алгоритму розв'язання системи ЗДР методом Ейлера

9.1.1 Модифікації методу Ейлера

З метою підвищення точності методу Ейлера використовують різні його модифікації. Суть удосконаленого методу Ейлера полягає в використанні ітераційної формули виду:

$$y_{i+1}^{(0)} = y_i + hf(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}}) \quad (9.16)$$

де $x_{i+\frac{1}{2}}$ - значення аргументу x в точці $x_i + \frac{h}{2}$, а $y_{i+\frac{1}{2}}$ - значення функції в точці $x_i + \frac{h}{2}$. Розглянемо диференціальне рівняння

$$y' = f(x, y) \quad (9.17)$$

з початковою умовою $y(x_0) = y_0$.

Потрібно знайти розв'язок рівняння (9.13) на відрізку $[a, b]$.

Розіб'ємо відрізок $[a, b]$ на n рівних частин точками $x_i = x_0 + ih$ ($i = 0, 1, 2, \dots, n$), де $h = (b - a) / n$ – крок інтегрування

Алгоритм методу складається з:

1. визначення похідної y'_0 в точці (x_0, y_0) :

$$y'_0 = f(x_0, y_0)$$

2. змінна x за формулою:

$$x_{0+\frac{1}{2}} = x_0 + \frac{h}{2}$$

3. визначення значення $y_{0+\frac{1}{2}}$ при $x_{0+\frac{1}{2}}$

$$y_{0+\frac{1}{2}} = y_0 + \frac{h}{2} y'_0$$

4. визначення похідної в точці $(x_{0+\frac{1}{2}}, y_{0+\frac{1}{2}})$

$$y'_{0+\frac{1}{2}} = f(x_{0+\frac{1}{2}}, y_{0+\frac{1}{2}})$$

5. використовуємо отримане значення $y'_{0+\frac{1}{2}}$ для визначення y'_1 за формулою:

$$y'_1 = y_0 + h y'_{0+\frac{1}{2}} = y_0 + hf(x_{0+\frac{1}{2}}, y_{0+\frac{1}{2}})$$

$$x_1 = x_{0+\frac{1}{2}} + \frac{h}{2}$$

6. змінюємо

7. повторюємо всі кроки алгоритму, починаючи з першого.

Графічна інтерпретація методу представлена на рисунку 9.7.

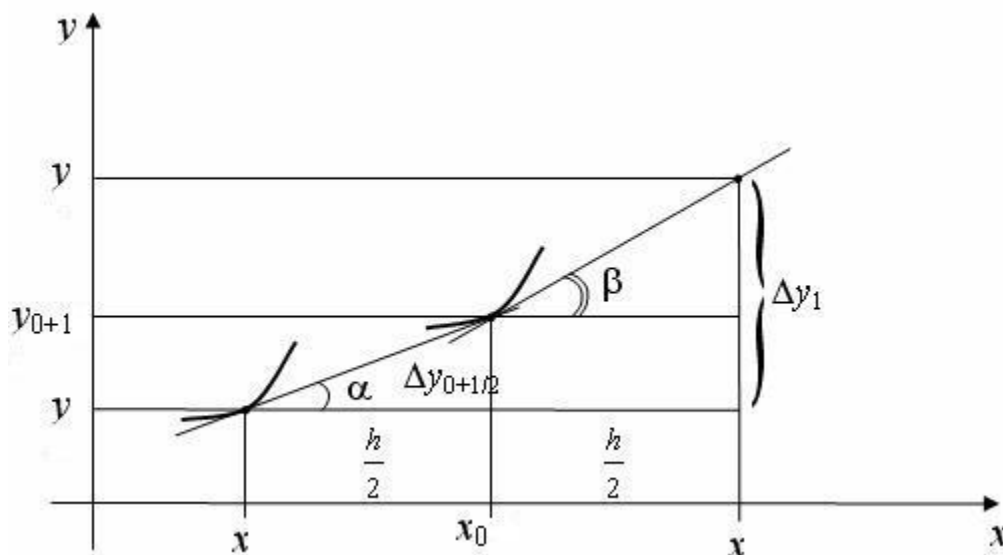


Рисунок 9.7 – Графічна інтерпретація удосконаленого методу Ейлера

y'_i **Зуваження.** Оцінка похибки в точці x_i може бути отримана за допомогою "подвійного прорахунку": розрахунок повторюють із кроком $h/2$ похибка більш точного значення (при кроці $h/2$) оцінюють приблизно в такий спосіб:

$$|y_i^* - y(x_i)| \approx \frac{1}{3} |y_i^* - y_i|, \quad (9.18)$$

де $y(x)$ - точний розв'язок диференціального рівняння. Удосконалений метод Ейлера є більш точним у порівнянні з методом Ейлера та відноситься до методів 3-го порядку точності.

Схема алгоритму представлена на рисунку 9.8.

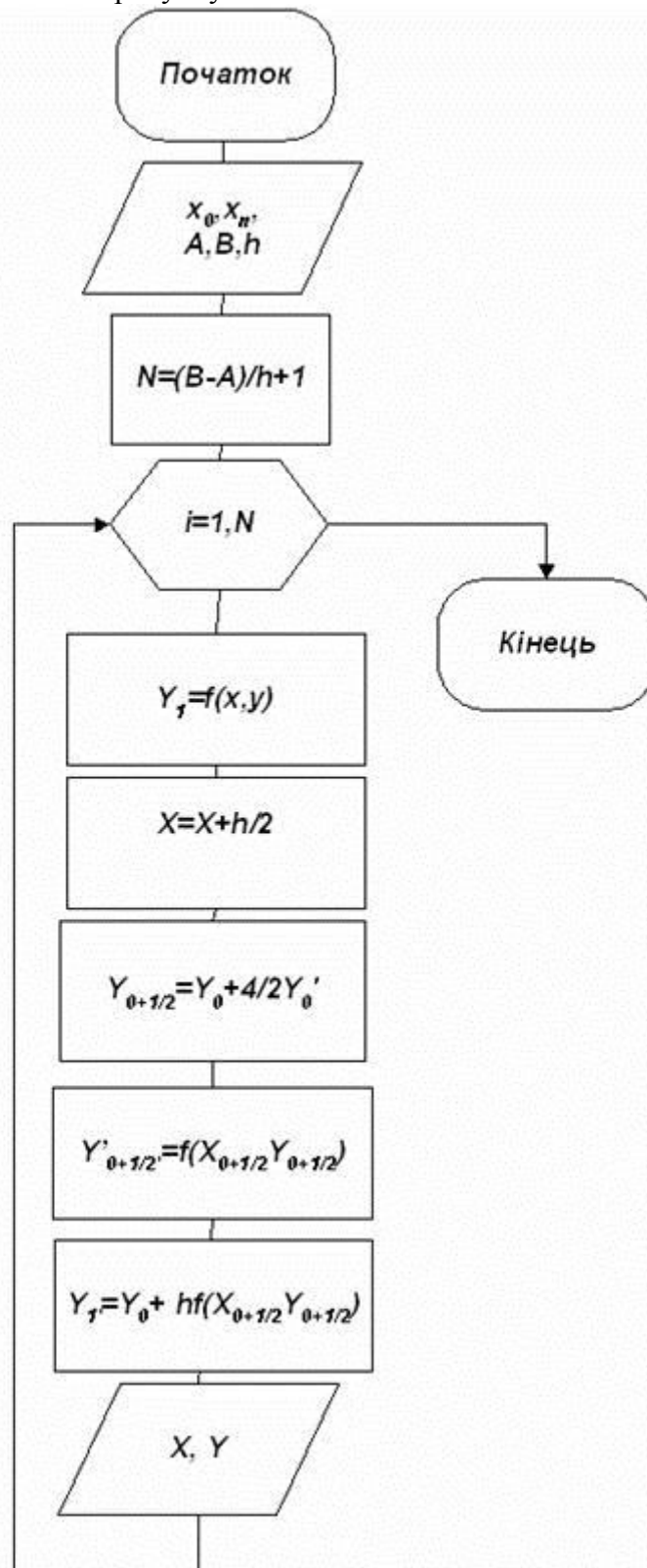


Рисунок 9.8 - Схема алгоритму удосконаленого метода Ейлера

Модифікований метод Ейлера заснований на використанні ітераційної формули виду:

$$y_{i+1} = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1})]. \quad (9.19)$$

Геометрична інтерпретація представлена на рисунку 9.9.

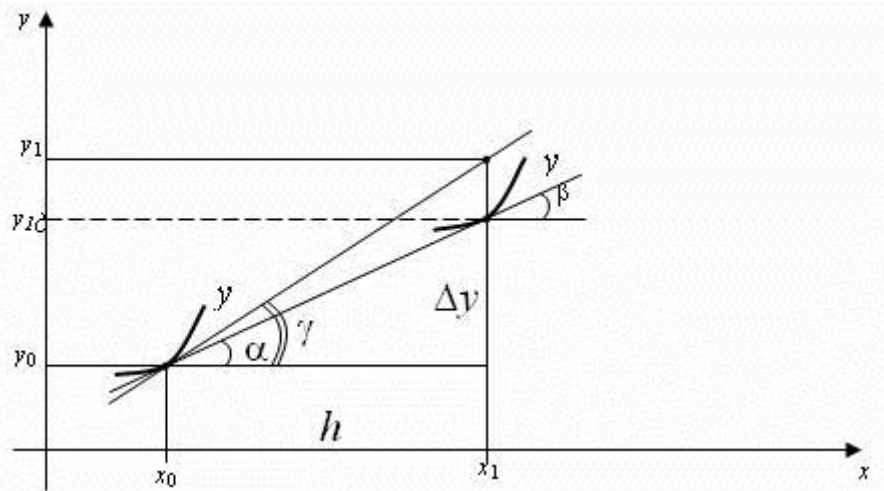


Рисунок 9.9 – Графічна інтерпретація

Алгоритм методу включає наступні кроки:

1. визначення похідної y'_0 в точці (x_0, y_0) :
 $y'_0 = f(x_0, y_0)$
2. зміна незалежної змінної x за формулою:
 $x_1 = x_0 + h$
3. визначення проміжного значення $y_{1\pi}$ за формулою методу Ейлера
 $y_{1\pi} = y_0 + h y'_0$
4. визначення проміжної похідної в точці $(x_1, y_{1\pi})$
 $y'_{1\pi} = f(x_1, y_{1\pi})$
5. визначення середньо арифметичного значення двох похідних
 $\Delta y = \frac{y'_0 + y'_{1\pi}}{2}$
6. визначення y_1 за формулою
 $y_1 = y_0 + h \cdot \Delta y = \frac{h}{2} (y'_0 + y'_{1\pi})$

ітераційний процес повторюється, починаючи з першого кроку.

Схема алгоритму методу представлена на рисунку 9.10.

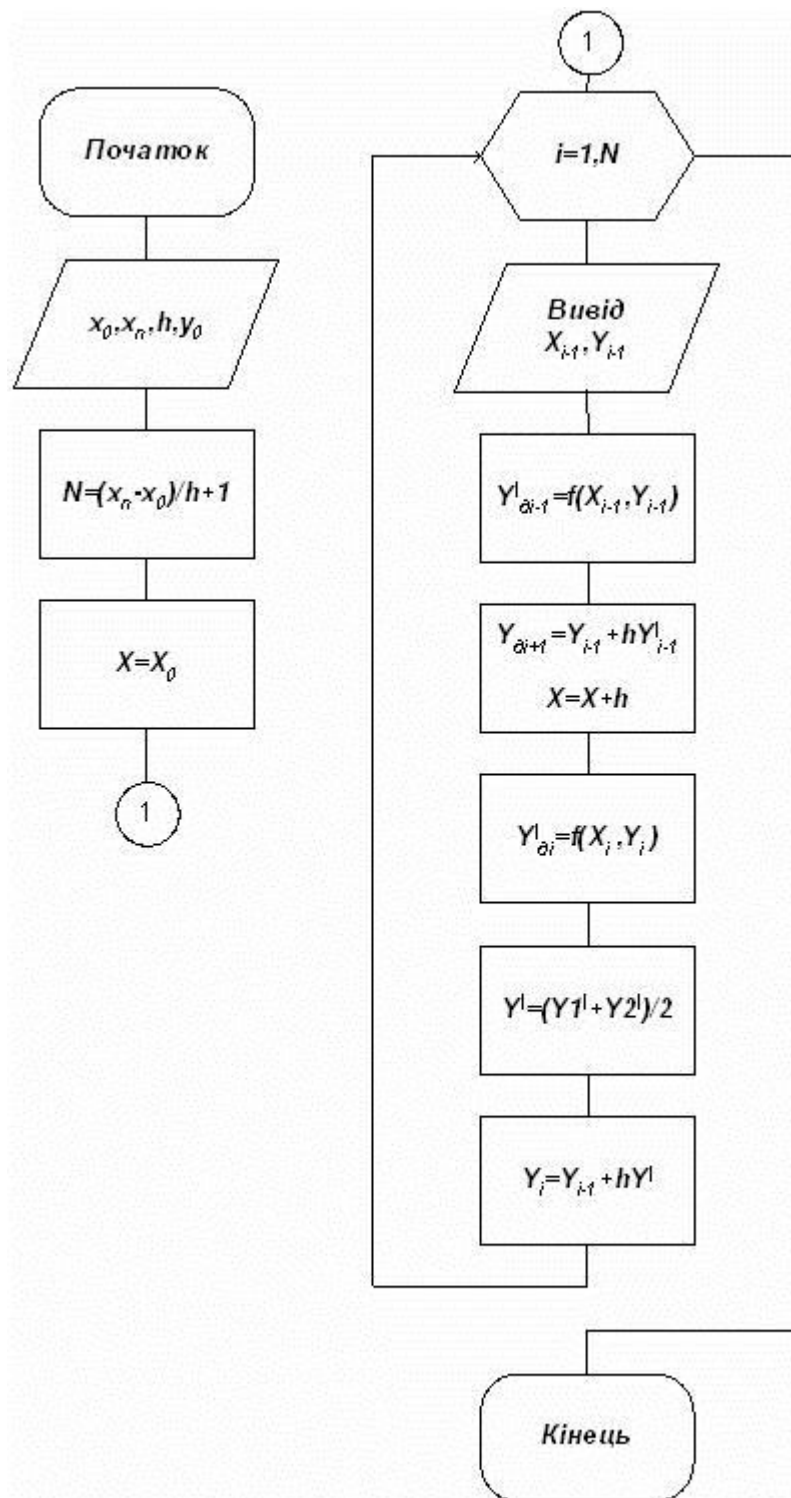


Рисунок 9.10 - Схема алгоритму модифікованого метода Ейлера

Удосконалений метод Ейлера - Коші з наступною ітераційною обробкою. Метод Ейлера - Коші з ітераційною обробкою є більш точним, чим, раніше розглянутий метод Ейлера - Коші. Сутність його полягає в тім, що виробляється ітераційна обробка кожного знайденого значення y_i . Спочатку вибирається грубе наближення потім будується ітераційний процес:

$$y_{i+n}^{(k)} = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+n}, y_{i+n}^{(k-1)})] \quad (9.20)$$

Ітерації продовжуються доти, поки два послідовних наближення не збіжаться до заданої похибки. Після цього приймається $y_{i+n}^{(k)} \approx y_{i+n}^{(k-1)}$. Якщо після трьох-чотирьох ітерацій, при обраному значенні h , збігу потрібних знаків не відбувається, то варто зменшити крок розрахунку h .

9.1.2 Метод Рунге-Кутта

Метод Рунге-Кутта є одним з методів підвищеної точності, але має багато загального з методом Ейлера.

Нехай на відрізку $[a, b]$ потрібно знайти чисельний розв'язок рівняння

$$y' = f(x, y) \quad (9.21)$$

з початковою умовою $y(x_0) = y_0$.

Розіб'ємо відрізок $[a, b]$ на n рівних частин точками $x_i = x_0 + ih (i = 0, 1, \dots, n)$, де $h = (b - a)/n$ - крок інтегрування. В методі Рунге-Кутта, так само як у методі Ейлера, послідовні значення y_i шуканої функції y визначаються по формулі

$$y_{i+1} = y_i + \Delta y_i. \quad (9.22)$$

Якщо розкласти функцію y у ряд Тейлора й обмежитися членами до h^4 включно, то збільшення функції y можна представити у вигляді

$$\Delta y = y(x+h) - y(x) = hy'(x) + \frac{h^2}{2} y''(x) + \frac{h^3}{6} y'''(x) + \frac{h^4}{24} y^{IV}(x), \quad (9.23)$$

де похідні $y''(x), y'''(x), y^{IV}(x)$ визначаються послідовним диференціюванням з рівняння (9.4).

Замість безпосередніх обчислень по формулі (9.23) у методі Рунге-Кутта визначаються чотири числа:

$$\begin{aligned} k_1 &= hf(x, y), \\ k_2 &= hf\left(x + \frac{h}{2}, y + \frac{k_1}{2}\right), \\ k_3 &= hf\left(x + \frac{h}{2}, y + \frac{k_2}{2}\right), \\ k_4 &= hf(x+h, y+k_3). \end{aligned} \quad (9.24)$$

Можна довести, що якщо числам k_1, k_2, k_3, k_4 додати відповідно вагу $1/6; 1/3; 1/3; 1/6$, то середньозважене цих чисел, тобто

$$\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 \quad (9.25)$$

з точністю до четвертих ступенів дорівнює значенню y , обчисленому по формулі (9.23):

$$\Delta y = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4). \quad (9.26)$$

Таким чином, для кожної пари поточних значень x_i і y_i по формулах (9.24) визначаються значення

По формулі (9.26) знаходиться

$$\begin{aligned} k_1^{(i)} &= hf(x_i, y_i), \\ k_2^{(i)} &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2}\right), \\ k_3^{(i)} &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2}\right), \\ k_4^{(i)} &= hf(x_i+h, y_i+k_3^{(i)}). \end{aligned} \quad (9.27)$$

і потім

$$y_{i+1} = y_i + \Delta y_i$$

Числа k_1, k_2, k_3, k_4 мають простий геометричний зміст. Нехай крива M_0CM_1 (рис. 9.11) являє собою розв'язок диференціального рівняння (9.21) з початковою умовою (9.22). Точка C цієї кривої лежить на прямій, рівнобіжній осі Oy і розділяючою відрізок $[x_i, x_{i+h}]$ навпіл, B і G - точки перетинання дотичної, проведеної до кривої в точці M_0 , з ординатами AC і N_1M_1 . Тоді число k_1 з точністю до множника h (де $h = x_{i+h} - x_i$) є кутівий

$$\Delta y_i = \frac{1}{6}(k_1^{(i)} + 2k_2^{(i)} + 2k_3^{(i)} + k_4^{(i)})$$

коефіцієнт дотичної в точці M_0 до інтегральної кривої M_0CM_1 тобто

$$k_1 = hy'_i = hf(x_i, y_i).$$

Точка B має координати $x = x_i + \frac{h}{2}$, $y = y_i + \frac{k_1}{2}$. Отже число k_2 з точністю до множника h є кутовий коефіцієнт дотичної, проведеної до інтегральної кривої в точці B (BF - відрізок цієї дотичної).

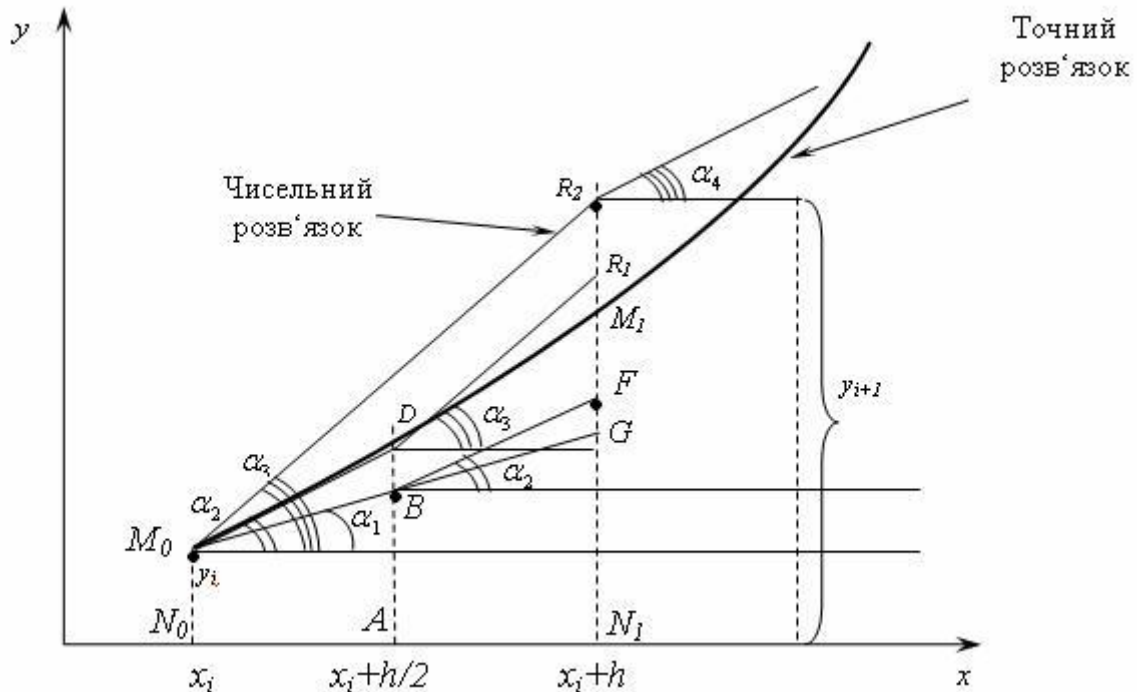


Рисунок 9.11 – Геометрична інтерпретація методу Рунге-Кутта

Через точку M_0 проведемо пряму, рівнобіжну відрітку BF . Тоді точка D має координати і k_3 з точністю до множника h - кутовий коефіцієнт дотичної, перевіреної до інтегральної кривої в точці D (DR_1 - відрізок цієї дотичної). Нарешті, через точку M_0 проведемо пряму, рівнобіжну DR_1 , що перетне продовження M_1N_1 у точці $R_2(x_i + h, y_i + k_3)$.

Тоді k_4 з точністю до множника h є кутовий коефіцієнт дотичної, проведеної до інтегральної кривої в точці R_2 .

Обчислення по методу Рунге-Кутта зручно розташовувати за алгоритмом,

1. Значення x_0 и y_0 підставляють у праву частину диференціального рівняння (9.21), визначають $f(x, y)$.

2. Отримане значення $f(x, y)$ множать на крок інтегрування h , обчислюють $k_1 = hf(x, y)$.

3. Змінюють значення x_0

$$x_{0+\frac{1}{2}} = x_0 + \frac{h}{2}$$

4. визначають допоміжне значення $y_{0\partial}$

$$y_{0\partial} = y_0 + \frac{k_1}{2}$$

5. визначення похідної в точці $(x_{0+\frac{1}{2}}, y_{0\partial})$

$$y'_{0\partial} = f(x_{0+\frac{1}{2}}, y_{0\partial})$$

6. визначають значення k_2

$$k_2 = h f(x_{0+\frac{1}{2}}, y_{0\partial})$$

7. визначають нове допоміжне значення $y_{0\partial}$

$$y_{0\partial} = y_0 + \frac{k_2}{2}$$

8. визначення похідної в точці $(x_{0+\frac{1}{2}}, y_{0\partial})$

$$y'_{0\partial} = f(x_{0+\frac{1}{2}}, y_{0\partial})$$

9. визначають значення k_3

$$k_3 = h \cdot y'_{0\partial} = h \cdot f(x_{0+\frac{1}{2}}, y_{0\partial})$$

10. визначають нове значення допоміжного $y_{1\partial}$

$$y_{1\partial} = y_0 + k_3$$

11. змінюють значення $x_{0+\frac{1}{2}}$

$$x_1 = x_{0+\frac{1}{2}} + \frac{h}{2}$$

12. визначають допоміжну похідну в точці $(x_1, y_{1\partial})$

$$y'_{1\partial} = f(x_1, y_{1\partial})$$

13. визначають значення k_4

$$k_4 = h \cdot y'_{1\partial} = h \cdot f(x_0 + h, y_0 + k_3)$$

14. визначають нове значення y_1 за формулою

$$y_1 = y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

Для визначення y_2, y_3, \dots, y_n повторюють ітераційний процес, починаючи з першого кроку. Схема алгоритму метода Рунге-Кутта представлена на рисунку 9.12.

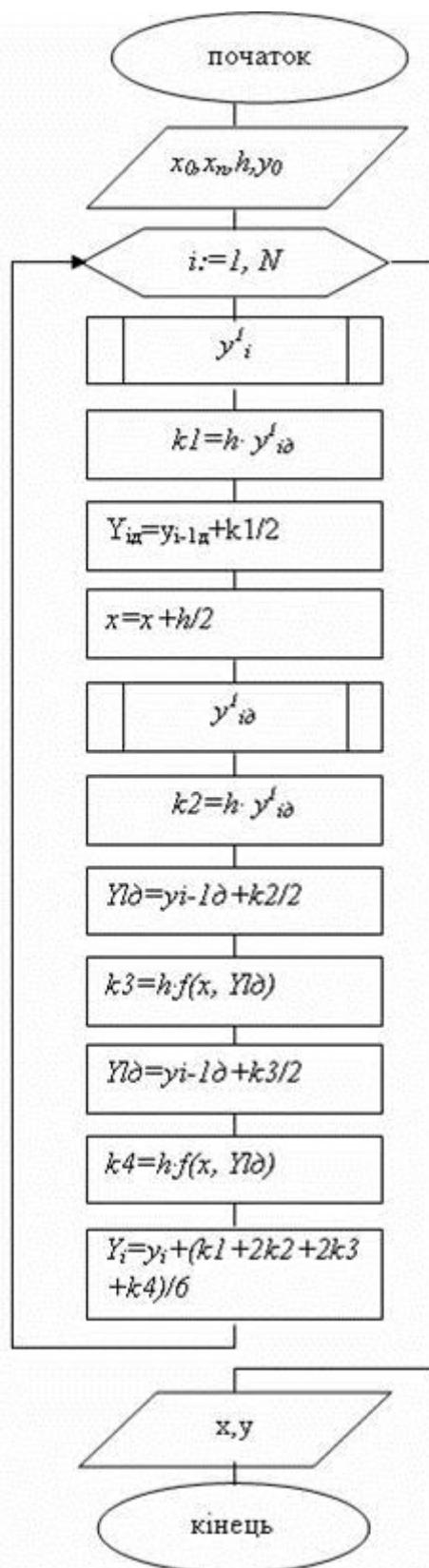


Рисунок 9.12 – Схема алгоритму метода Рунге – Кутта

Потім всі обчислення повторюють починаючи з I кроку, доти, поки не буде пройдений весь відрізок $[a, b]$.

Метод Рунге - Кутта має порядок точності h^4 на усьому відрізку $[a, b]$. Оцінка точності методу цього дуже складна. Грубу оцінку погрішності можна одержати за допомогою "подвійного прорахунку" по формулі

$$|y_i^* - y(x_i)| \approx \frac{y_i^* - y_i}{15}, \quad (9.28)$$

де $y^{(x)}$ - значення точного розв'язку рівняння (9.21) у точці x_i , а $y^*, i y_i$ наближені значення, отримані з кроком $h/2$ і h . Якщо ϵ - задана точність розв'язку, то число n (число розподілів) для визначення кроку інтегрування $h = (b - a) / n$ вибирається таким чином, щоб

$$h^4 < \epsilon \quad (9.29)$$

Однак крок розрахунку можна змінювати при переході від однієї точки до іншої.

Для оцінки правильності вибору кроку h використовують рівність

$$q = \frac{|k_2^{(i)} - k_3^{(i)}|}{|k_1^{(i)} - k_2^{(i)}|}, \quad (9.30)$$

де q повинно бути дорівнює декільком сотим, у противному випадку крок h зменшують.

Нехай задана система диференціальних рівнянь першого порядку:

$$\begin{cases} y' = f(x, y, z), \\ z' = g(x, y, z) \end{cases} \quad (9.31)$$

У цьому випадку паралельно визначаються числа Δy_i і Δz_i :

$$\begin{aligned} \Delta y_i &= \frac{1}{6} (k_1^{(i)} + 2k_2^{(i)} + 2k_3^{(i)} + k_4^{(i)}), \\ \Delta z_i &= \frac{1}{6} (l_1^{(i)} + 2l_2^{(i)} + 2l_3^{(i)} + l_4^{(i)}), \end{aligned} \quad (9.33)$$

$$k_1^{(i)} = hf(x_i, y_i, z_i),$$

$$l_1^{(i)} = hg(x_i, y_i, z_i);$$

$$k_2^{(i)} = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2}, z_i + \frac{l_1^{(i)}}{2}\right);$$

$$l_2^{(i)} = hg\left(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2}, z_i + \frac{l_1^{(i)}}{2}\right);$$

$$k_3^{(i)} = hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2}, z_i + \frac{l_2^{(i)}}{2}\right);$$

$$l_3^{(i)} = hg\left(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2}, z_i + \frac{l_2^{(i)}}{2}\right);$$

$$k_4^{(i)} = hf(x_i + h, y_i + k_3^{(i)}, z_i + l_3^{(i)});$$

$$l_4^{(i)} = hg(x_i + h, y_i + k_3^{(i)}, z_i + l_3^{(i)}). \quad (9.34)$$

Тоді одержимо розв'язок системи диференційних рівнянь другого порядку:

$$y_{i+1} = y_i + \Delta y_i, \quad z_{i+1} = z_i + \Delta z_i.$$

$$l_2 = h \frac{\left[\frac{2(y_0 + \frac{k_1}{2})}{(z_0 + \frac{l_1}{2}) + (x_0 + \frac{h}{2})} \right]}{=} = 0,1 \cdot \frac{2 \cdot 1,075}{1,06667 + 0,55} = 0,13299;$$

$$k_3 = h \frac{\left[\frac{2(y_0 + \frac{k_2}{2}) - (x_0 + \frac{h}{2})}{z_0 + \frac{l_2}{2}} \right]}{=} = 0,1 \cdot \frac{2 \cdot 1,07050 - 0,55}{1,06650} = 0,14918;$$

$$l_3 = h \frac{\left[\frac{2(y_0 + \frac{k_2}{2})}{(z_0 + \frac{l_2}{2}) + (x_0 + \frac{h}{2})} \right]}{=} = 0,1 \cdot \frac{2 \cdot 1,07050}{1,06650 + 0,55} = 0,13245;$$

$$k_4 = h \left[\frac{2(y_0 + k_3) - (x_0 + h)}{z_0 + l_3} \right] = 0,1 \cdot \frac{2 \cdot 1,14918 - 0,6}{1,13245} = 0,14998$$

$$l_4 = h \left[\frac{2(y_0 + k_3)}{(z_0 + l_3) + (x_0 + h)} \right] = 0,1 \cdot \frac{2 \cdot 1,14918}{1,13245 + 0,6} = 0,13266$$

Отже,

$$\Delta y_0 = \frac{1}{6}(0,15 + 2 \cdot 0,14100 + 2 \cdot 0,14918 + 0,14998) = 0,14672;$$

$$\Delta z_0 = \frac{1}{6}(0,13333 + 2 \cdot 0,13299 + 2 \cdot 0,13245 + 0,13266) = 0,13281$$

і остаточно одержуємо значення шуканих функцій у точці $x = 0,6$:

$$y_1 = 1 + 0,14672 = 1,14672; \quad z_1 = 1 + 0,13281 = 1,13281.$$

Приклад 9.1. Розв'язати на ЕОМ систему рівнянь виду:

$$\begin{cases} \frac{dy_1}{dx} = F_1(x, y_1, y_2) = y_2 \\ \frac{dy_2}{dx} = F_2(x, y_1, y_2) = \left(\frac{y_1}{x} - y_2 \right) \frac{1}{x} - y_1 \end{cases}$$

При розробці програми для розв'язку системи диференціальних рівнянь, створюють: 1) підпрограму, яка містить праві частини рівнянь; 2) підпрограму, яка реалізує алгоритм метода Рунге-Кутта для обчислення значень функцій системи в одній точці; 3) підпрограму, яка дозволяє отримати розв'язок системи на заданому відрізку дослідження. Наприклад, підпрограму, яка містить дану систему можна записати на алгоритмічній мові С у вигляді:

```
void systema (float *Y, float *F, float X)
{ F[1] = Y[2];
  F[2] = (Y[1] / X - Y[2]) / X - Y[1];
}
```

де $F[1], F[2]$ - значення правої частини рівняння заданої системи.

Розглянемо особливості алгоритмів методів прогнозу і корекції.

Так як в методах, що розглядаються використовується інформація про декілька раніше отриманих точок, то на відміну від однокрокових методів вони не володіють властивістю „самостартування”. Тому, перед тим як застосовувати метод прогнозу і корекції, необхідно обчислювати вихідні данні за допомогою якого-небудь однокрокового методу. Часто для цього використовують метод Рунге-Кутта. Обчислення проводять наступним чином. Спочатку по формулі прогнозу і початковим значенням змінних знаходять значення $y_{n+1}^{(0)}$. Верхній індекс (0) означає, що прогнозоване значення є одним з послідовності значень y_{n+1} , що розташовані в порядку зростання точності. По прогнозованому значенню $y_{n+1}^{(0)}$ за допомогою приведеного вище диференціального рівняння знаходять похідну

$$y_{n+1}^{(0)'} = f(x_{n+1}, y_{n+1}^{(0)}) \quad (9.36)$$

яка потім підставляється у формулу корекції для обчислення уточненого значення $y_{n+1}^{(j+1)}$.

В свою чергу $y_{n+1}^{(j+1)}$ використовується для отримання більш точного значення похідної за допомогою диференціального рівняння

$$y_{n+1}^{(j+1)'} = f(x_{n+1}, y_{n+1}^{(j+1)}) \quad (9.37)$$

Якщо це значення похідної недостатньо близьке до попереднього, то воно вводиться у формулу корекції й ітераційний процес продовжується. Якщо ж похідна змінюється в

допустимих границях, то значення $y_{n+1}^{(j+1)'}$ використовується для обчислення остаточного значення y_{n+1} , яке і видається на друк. Після цього процес повторюється – здійснюється наступний крок, на якому обчислюється y_{n+2} .

Звичайно при вводі формул прогнозу і корекції розв'язок рівняння розглядають як процес наближеного інтегрування, а самі формули отримують за допомогою кінцево-різницевого методів.

Якщо диференціальне рівняння $y'(x) = f(x, y)$ проінтегровано в інтервалі значень від x_n до x_{n+k} , то результат прийме вигляд

$$y(x_{n+k}) - y(x_n) = \int_{x_n}^{x_{n+k}} f(x, y) dx. \quad (9.38)$$

Цей інтеграл не можна обчислити безпосередньо, так як залежність $y(x)$ заздалегідь невідома. Наближене значення інтегралу можна знайти за допомогою одного з кінцево-різницевого методів. Вибір методу і буде визначати метод розв'язку диференціальних рівнянь. На етапі прогнозу можна використовувати будь-яку формулу чисельного інтегрування, якщо до неї не входить попереднє значення $y'(x_{n+1})$.

9.4.1 Метод Мілна

В цьому методі на етапі прогнозу використовується формула Мілна

$$y_{n+1} = y_{n-3} + \frac{4}{3}h(2y'_n - y'_{n-1} + 2y'_{n-2}) + \frac{28}{90}h^5 y^{(5)}, \quad (9.39)$$

а на етапі корекції - формула Сімпсона

$$y_{n+1} = y_{n-1} + \frac{1}{3}h(y'_{n+1} + 4y'_n + y'_{n-1}) - \frac{1}{90}h^5 y^{(5)}. \quad (9.40)$$

Останні члени в обох формулах в дійсності в ітераційному процесі не використовуються і слугують лише для оцінки помилки відсічення. Метод Мілна відносять до методів четвертого порядку точності, так як в ньому відкидаються члени, які містять h п'ятій степені і більш високих степенях. Може виникнути питання навіщо потрібна корекція, якщо прогноз дає четвертий порядок точності. Відповідь на це питання дає оцінка відносної величини членів, що виражають похибку. В даному випадку похибка відсічення при корекції в 28 разів менше і тому представляє великий інтерес. Незважаючи на те що формула Мілна містить менший числовий коефіцієнт (1/90) перед членом, що відкидається, її використовують рідше, ніж інші (з більшими відкидуваними членами), так як їй притаманна нестійкість. Це означає, що похибка поширення може рости експоненціально, при чому цей висновок справедливий для всіх формул корекції, оснований на правилі Сімпсона.

9.4.2 Метод Адамса - Башфорта

Цей метод також має четвертий порядок точності. Формула, що використовується в ньому отримана інтегруванням оберненої інтерполяційної формули Ньютона і має вид

$$y_{n+1} = y_n + \frac{1}{24}h(55y'_n - 59y'_{n-1} + 37y'_{n-2} - 9y'_{n-3}) + \frac{251}{720}h^5 y^{(5)}. \quad (9.41)$$

На етапі корекції використовується формула

$$y_{n+1} = y_n + \frac{1}{24}h(9y'_{n+1} - 19y'_n - 37y'_{n-1} - y'_{n-2}) + \frac{19}{720}h^5 y^{(5)}. \quad (9.42)$$

Розрахунки по методу Адамса - Башфорта виконуються так як і по методу Мілна, але на відміну від останнього похибка, внесена на якому-небудь кроці, не має тенденції до експоненціального росту.

Можна припустити, що так як величина відкиненого члену відома, то її можна використовувати для уточнення скоректованого значення залежної змінної. Так як внесення поправок в корегуючий член може негативно вплинути на стійкість рахунку, то для підвищення точності рахунку слід використовувати методи більш високих порядків точності.

9.4.3 Метод Хемінга

В методі Хемінга використовуються наступні формули:

$$y_{n+1}^{(0)} = y_{n-3} + \frac{4}{3}h(2y'_n - y'_{n-1} + 2y'_{n-2}), \quad (9.43)$$

уточнення прогнозу

$$\bar{y}_{n+1}^{(0)} = y_{n+1}^{(0)} + \frac{112}{121}(y_n - y_n^{(0)}), \quad (9.44)$$

$$[\bar{y}_{n+1}^{(0)}]' = f(x_{n+1}, \bar{y}_{n+1}^{(0)}), \quad (9.45)$$

корекції

$$y_{n+1}^{(j+1)} = \frac{1}{8}(9y_n - y_{n-2}) + \frac{3}{8}h([\bar{y}_{n+1}^{(j)}]') + 2y'_n - y'_{n-1}. \quad (9.46)$$

Це стійкий метод четвертого порядку точності, в основі якого лежать наступні формули прогнозу:

$$y_{n+1} = y_{n-3} + \frac{4}{3}h(2y'_n - y'_{n-1} + 2y'_{n-2}) + \frac{28}{90}h^5 y^{(5)}, \quad (9.47)$$

і корекції

$$y_{n+1} = \frac{1}{8}[9y_n - y_{n-2} + 3h(y'_{n+1} + 2y'_n - y'_{n-1})] - \frac{1}{40}h^5 y^{(5)}. \quad (9.48)$$

Особливістю методу Хемінга є те, що він дозволяє оцінювати похибки, що вносяться на стадіях прогнозу і корекції і усувати їх. Завдяки простоті та стійкості цей метод є одним з найбільш поширених методів прогнозу і корекції.

ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ КРАЙОВИХ ЗАДАЧ

10.1 Чисельні методи розв'язання крайових задач

Як зазначалося вище, задача, що заключається в розв'язку звичайного диференційного рівняння при додаткових умовах, що поставлені при декількох значеннях незалежної змінної, називається *крайовою*.

Розглянемо звичайне диференційне рівняння другого порядку

$$\frac{d^2 y}{dx^2} = f(x, y, y') \quad (10.1)$$

при граничних умовах

$$\begin{aligned} y(a) &= A, \\ y(b) &= B. \end{aligned} \quad (10.2)$$

Рівняння більш високих порядків можна вирішувати тими ж методами, які поділимо на дві групи:

1. Методи, основані на заміні розв'язку крайової задачі розв'язком декількох задач Коши.

2. Методи, в яких використовується кінцево-різницева форма диференціального рівняння.

10.1.1 Методи стрільби

Якщо звичайне диференційне рівняння другого порядку є лінійним, тобто має вигляд

$$y'' = f_1(x)y' + f_2(x)y + f_3(x) \quad (10.3)$$

при додаткових умовах

$$y(a) = A \text{ і } y(b) = B, \quad (10.4)$$

то крайову задачу можна звести до задачі Коши за допомогою початкових умов

$$y(a) = A \text{ і } y'(a) = \alpha_1.$$

Знайшовши розв'язок $y_1(x)$ можна поставити другі граничні умови

$$y(a) = A \text{ і } y'(a) = \alpha_2$$

і отримати другий розв'язок $y_2(x)$. Якщо $y_1(b) = \beta_1$, а $y_2(b) = \beta_2$, причому $\beta_1 \neq \beta_2$, то розв'язок

$$y(x) = \frac{1}{\beta_1 - \beta_2} [(B - \beta_2)y_1(x) + (\beta_1 - B)y_2(x)] \quad (10.5)$$

задовольняє обом початковим граничним умовам.

Якщо вирішується нелінійне звичайне диференційне рівняння, то розв'язок крайової задачі можна звести до розв'язку декількох задач Коши, послідовно вводячи в початкові умови значення α :

$$y(a) = A \text{ і } y'(a) = \alpha \quad (10.6)$$

і прагнучи знайти розв'язок, який задовольняє умову $y(b) = B$. При цьому може допомогти інтерполяція, що дозволяє побудувати впорядковану послідовність α і звести до мінімуму об'єм обчислень.

10.1.2 Кінцево-різницеві методи

Перевага кінцево-різницевих методів в тому, що вони дозволяють звести розв'язок крайової задачі до розв'язку системи алгебраїчних рівнянь. При розв'язку двоточкової крайової задачі

$$y'' = f(x, y, y'), \quad (10.7)$$

при $y(a) = A$ і $y(b) = B$ інтервал $[a, b]$ можна розділити на n рівних частин:

$$x_i = x_0 + ih, i = 1, 2, \dots, n, \quad (10.8)$$

де $x_0 = a$, $x_n = b$, а $h = \frac{b-a}{n}$.

В точках x_i , що називаються вузлами, намагаються знайти значення розв'язок y_i . Знаючи координати вузлів користуючись кінцево-різницевиими виразами для похідних

$$y'(x_i) = \frac{1}{2h}(y_{i+1} - y_{i-1}), \quad (10.9)$$

$$y''(x_i) = \frac{1}{h^2}(y_{i+1} - 2y_i + y_{i-1}), \quad (10.10)$$

можна представити диференціальне рівняння у виді різницевого рівняння.

а) якщо початкове ЗДР лінійне, то задача буде складатися з розв'язання системи лінійних алгебраїчних рівнянь;

б) якщо початкове ЗДР нелінійне, то задача зводиться до розв'язування системи нелінійних алгебраїчних або трансцендентних рівнянь;

в) методи, алгоритми та програми розв'язування таких лінійних та нелінійних рівнянь відомі, але звести розв'язування крайової задачі методом кінцевих різниць до стандартної програми важко, оскільки формулювання кожної задачі залежить від вигляду рівняння, що розглядається.

Приклад 10.1. Нехай потрібно вирішити диференціальне рівняння $y'' = 2x + y$ при умовах $y(0) = 0$ і $y(1) = 1$ і кроці $h = 0.2$. В різницевій формі це рівняння має вид

$$y''(x_i) = \frac{1}{0.04}(y_{i+1} - 2y_i + y_{i-1}) = 2x_i + 3y_i.$$

Використовуючи цю формулу і граничні умови можна виписати наступну систему чотирьох лінійних рівнянь з чотирма невідомими:

$$\begin{cases} -2.12y_1 + y_2 = 0.016, \\ y_3 - 2.12y_2 + y_1 = 0.032, \\ y_4 - 2.12y_3 + y_2 = -0.064, \\ -2.12y_4 + y_3 = 0.936. \end{cases}$$

$$y(x) = \frac{5 \operatorname{sh} \sqrt{3}x}{3 \operatorname{sh} \sqrt{3}} - \frac{2}{3}.$$

10.1.3 Метод прогонки

Якщо звичайне диференціальне рівняння (ЗДР) 2-го порядку розв'язати методом скінченних різниць, то початкове ЗДР зводиться до системи лінійних алгебраїчних рівнянь з тридіагональною матрицею коефіцієнтів, тобто такою, що кожне рівняння системи має три сусідні невідомі:

$$\begin{aligned} a_{11}y_0 + a_{12}y_1 + 0 \cdot y_2 + \dots + 0 \cdot y_n &= f_1; \\ a_{21}y_0 + a_{22}y_1 + a_{23}y_2 + 0 \cdot y_3 + \dots + 0 \cdot y_n &= f_2; \\ \dots & \dots \\ 0 \cdot y_0 + 0 \cdot y_1 + 0 \cdot y_2 + \dots + a_{m-1}y_{n-2} + a_{m-1}y_{n-1} + a_{mn}y_n &= f_n. \end{aligned} \quad (10.11)$$

Для розв'язання такої системи розроблено спеціальний метод – прогонки. Розглянемо лінійне диференціальне рівняння другого порядку

$$y'' + p(x)y' + q(x)y = f(x) \quad (10.12)$$

з двоточковими граничними умовами:

$$\begin{aligned} \alpha_0 y(a) + \alpha_1 y'(a) &= A, \\ \beta_0 y(b) + \beta_1 y'(b) &= B, \end{aligned} \quad (10.13)$$

$$|\alpha_0| + |\alpha_1| \neq 0; \quad |\beta_0| + |\beta_1| \neq 0, \quad (10.14)$$

де функції $p(x)$, $q(x)$, $f(x)$ неперервні на $[A, B]$.

Прийдемо до скінченнорізницевого рівняння

$$\begin{aligned} h &= (B - A) / N; \quad x_i = x_0 + ih; \quad x_0 = A; \quad x_n = B; \\ q(x_i) &= q_i; \quad p(x_i) = p_i; \quad f(x_i) = f_i; \quad y(x_i) = y_i. \end{aligned} \quad (10.15)$$

Підставляючи вираз (10.15) у початкове ЗДР (10.12), отримаємо для внутрішніх точок $x_i (i = 1, 2, \dots, n-1)$ систему скінченнорізницевих рівнянь:

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + p_i \frac{y_{i+1} - y_{i-1}}{2h} + q_i y_i = f_i \quad (10.16)$$

Після деяких перетворень (10.16) матимемо

$$y_{i+1} + m_i y_i + n_i y_{i-1} = \hat{f}_i h^2; \quad i = 1, 2, \dots, n-1, \quad (10.17)$$

де

$$m_i = \frac{2 - q_i h^2}{1 + \frac{p_i}{2} h}; \quad n_i = \frac{1 - \frac{p_i}{2} h}{1 + \frac{p_i}{2} h}; \quad \hat{f}_i = \frac{f_i}{1 + \frac{p_i}{2} h} \quad (10.18)$$

Для похідних на кінцях відрізка інтегрування $x_0 = A$ і $x_n = B$ знаходимо скінченнорізницеві рівняння виду

$$y_0' = (y_1 - y_0) / h; \quad y_n' = (y_{n-1} - y_n) / (-h) \quad (10.19)$$

і, підставляючи їх у граничні умови (10.13) і (10.14), отримаємо два рівняння:

$$\begin{aligned} \alpha_0 y_0 + \frac{\alpha_1}{h} (y_1 - y_0) &= A, \\ \beta_0 y_n + \frac{\beta_1}{-h} (y_{n-1} - y_n) &= B. \end{aligned} \quad (10.20)$$

Система $N+1$ рівнянь (10.17), (10.20) відносно невідомих y_0, y_1, \dots, y_n являє собою систему лінійних алгебраїчних рівнянь з три діагональною матрицею коефіцієнтів. Цю систему зручно розв'язати методом прогонки.

Розглянемо суть цього методу.

1. Представимо рівняння (10.17) відносно y_i :

$$y_i = \frac{\hat{f}_i}{m_i} h^2 - \frac{1}{m_i} y_{i+1} - \frac{n_i}{m_i} y_{i-1} \quad (10.21)$$

2. Допустимо, що за допомогою рівнянь (10.20) і (10.21) виключена складова y_{i-1} , тоді рівняння будуть мати вигляд

$$y_i = c_i (d_i - y_{i+1}), \quad (10.22)$$

де $c_i, d_i (i = 1, 2, \dots, N-1)$ – деякі коефіцієнти.

3. За аналогією з (10.22) представимо y_{i-1} у вигляді

$$y_{i-1} = c_{i-1}(d_{i-1} - y_i) \quad (10.23)$$

і, підставляючи його в (10.17), матимемо

$$y_{i+1} + m_i y_i + n_i c_{i-1}(d_{i-1} - y_i) = \hat{f}_i h^2 \quad (10.24)$$

і, отже,

$$y_i = \frac{(\hat{f}_i h^2 - n_i c_{i-1} d_{i-1}) - y_{i+1}}{m_i - n_i c_{i-1}} \quad (10.25)$$

4. Порівнюючи вирази (10.22) та (10.25), отримуємо рекурентні формули для визначення коефіцієнтів c_i, d_i вигляду

$$c_i = \frac{1}{m_i - n_i c_{i-1}}; \quad d_i = \hat{f}_i h^2 - n_i c_{i-1} d_{i-1}, \quad (i = 1, 2, \dots, N-1) \quad (10.26)$$

5. Для визначення виразу для c_0 та d_0 використовуємо рівняння (10.20) та (10.22):

$$y_0 = \frac{Ah - \alpha_1 y_1}{\alpha_0 h - \alpha_1}; \quad (10.27)$$

$$\text{та } y_0 = c_0(d_0 - y_1). \quad (10.28)$$

Порівнюючи останні дві нерівності, знаходимо:

$$c_0 = \frac{\alpha_1}{\alpha_0 h - \alpha_1}; \quad d_0 = \frac{Ah}{\alpha_1} \quad (10.29)$$

На основі формул (10.17) та (10.29) послідовно визначаємо коефіцієнти c_i, d_i ($i = 1, 2, \dots, N-1$) до c_{n-1} і d_{n-1} включно. Процес знаходження коефіцієнтів c_i, d_i називають прямим ходом методу прогонки.

Зворотній хід методу прогонки починається з визначення y_n . Використовуючи другу граничну умову (10.20) та формулу (10.22), отримуємо систему з двох рівнянь:

$$\begin{cases} \beta_0 y_n + \beta_1 \frac{y_{n+1} - y_n}{-h} = B, \\ y_{n+1} = c_{n+1}(d_{n+1} - y_n) \end{cases} \quad (10.30)$$

Розв'язуючи цю систему відносно y_n , отримаємо:

$$y_n = \frac{Bh + \beta_1 c_{n+1} d_{n+1}}{\beta_0 h + \beta_1 c_{n+1} + 1} \quad (10.31)$$

Тепер за формулою (10.22) послідовно знаходимо розв'язок початкового ЗДР: y_{n+1}, y_n, \dots, y_1 .

Таким чином, метод прогонки достатньо простий, легко алгоритмізується та включає прямий та зворотній хід. Причому:

1) прямий хід полягає в знаходженні коефіцієнтів за рекурентними формулами (10.26);

2) зворотній хід полягає у знаходженні розв'язку початкового ЗДР y_{n+1}, y_n, \dots, y_1 за формулою (10.22).

Схему алгоритму методу прогонки показано на рис. 10.1.

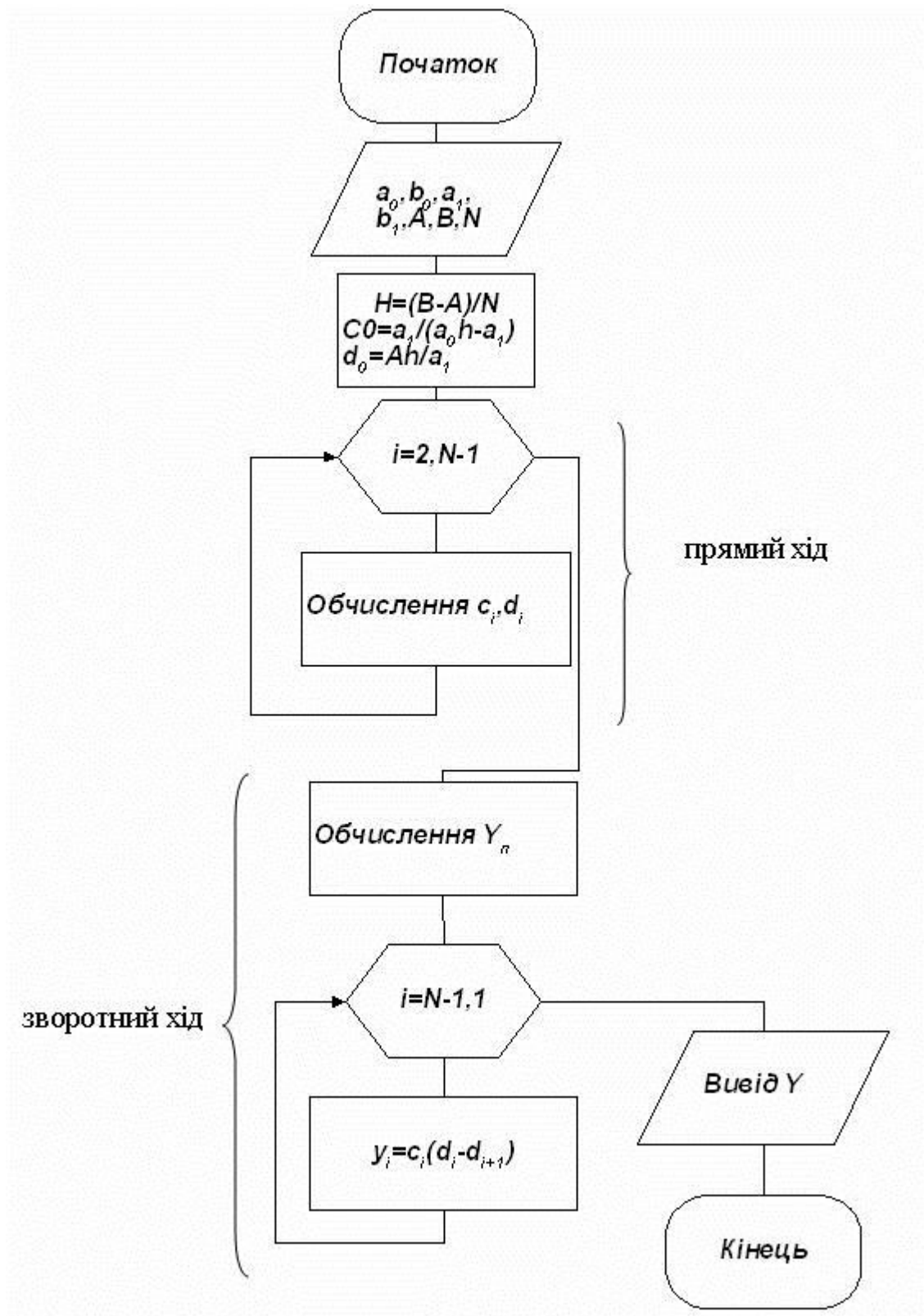


Рисунок 10.1 - Схему алгоритму методу прогонки

ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ У ЧАСТИННИХ ПОХІДНИХ

Диференціальні рівняння (ДР) в частинних похідних складають у даний час одну з галузей чисельного аналізу, які набувають дуже швидких темпів розвитку. Области науки і техніки, де розглядаються рівняння в частинних похідних, численні і важливі. До них відносяться, наприклад, ядерна фізика, метрологія, аеро-, гідро-, теплодинаміка, теорія управління і т.п. При цьому існує достатньо невелика кількість задач, які вирішуються у явному вигляді.

Диференційні рівняння в частинних похідних класифікуються:

- залежно від їх математичної природи (наприклад, еліптичні, параболічні тощо);
- залежно від фізичного змісту задач, які розв'язуються з їх допомогою (рівняння дифузії, хвильове рівняння тощо).

Розглянемо рівняння 2-го порядку з двома невідомими незалежними змінними:

$$F(x, y, u, u_x, u_y, u_{xx}, u_{yy}, u_{xy}) = 0, \quad (11.1)$$

де x, y – незалежні змінні; u – шукана функція; u_x, u_y, u_{xx}, u_{yy} – її 1-і та 2-і частинні похідні за аргументами x та y .

Розв'язком рівняння (11.1) називається функція $u = u(x, y)$, яка перетворює це рівняння в тотожність. Графік функції розв'язку представляє поверхню у просторі OXY (інтегральна поверхня) (рис. 11.1).

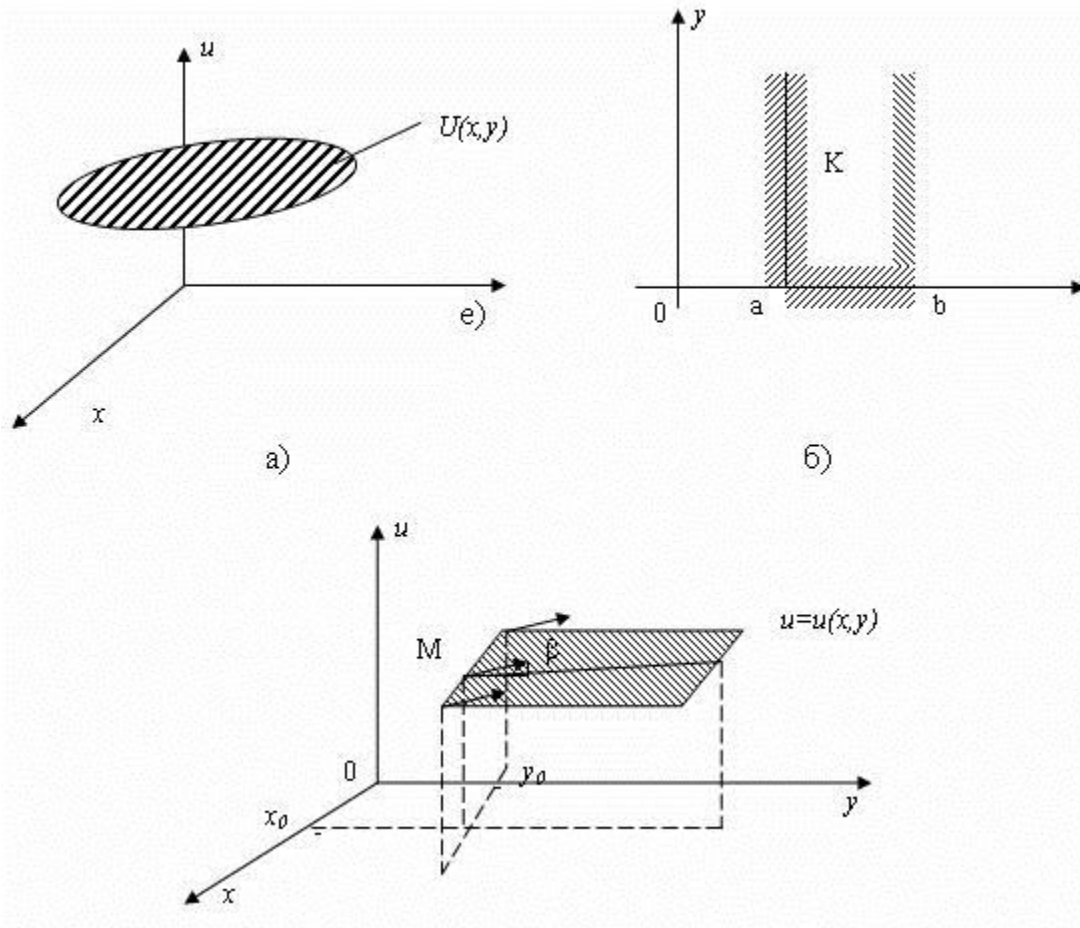


Рисунок 11.1 – Приклади інтегральних поверхонь

Рівняння (11.1) називається *лінійним*, якщо воно відносно 1-го степеня шуканої функції і всіх її похідних, і не стримує їхні добутки, та може мати вигляд:

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + E \left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) = 0, \quad (11.2)$$

де коефіцієнти A, B, C, E можуть залежати від x, y в окремих випадках можуть бути константами. Для класифікації ДР в частинних похідних вводиться термін *дискримінант рівняння*: $D = AC - B^2$. Залежно від знака дискримінанта D , лінійне диференційне рівняння (11.2) можна віднести до одного з таких типів:

- якщо $D > 0$, то рівняння (11.2) відноситься до еліптичного;
- якщо $D = 0$, то рівняння (11.2) відноситься до параболічного;
- якщо $D < 0$, то рівняння (11.2) відноситься до гіперболічного.

У наступних розділах розглянемо приклади кожного з трьох типів рівнянь з відповідними граничними умовами і розглянемо для них чисельні методи розв'язання. Слідє відмітити, що еліптичні рівняння описують установлені процеси (стаціонарні), а гіперболічні та параболічні рівняння описують еволюційні процеси.

11.1 Різницеві методи розв'язування диференційних рівнянь у частинних похідних

Класичне визначення похідної функції однієї змінної записується у вигляді

$$\frac{dy}{dx} = \lim_{h \rightarrow 0} \frac{y(x+h) - y(x)}{h}$$

Природно, в ЕОМ не можна зробити граничного переходу. З іншого боку, можна додати деяке мале, хоча і ненульове, значення h і спробувати перевірити, що наближення виходить досить точним (проблема точності) і що помилка не зростає в ході процесу обчислень (проблема стійкості), тобто цей метод зводиться до того, що похідну заміняємо різницею.

Оскільки тепер є дві незалежні змінні, то обидві вони повинні брати участь у різницевому рівнянні. Розглянемо спочатку різниці в напрямку x .

Розклад функції $u(x, y_0)$ в ряд Тейлора в околі точки (x_0, y_0) можна записати у вигляді:

$$u(x, y_0) = u(x_0, y_0) + (x - x_0)u_x(x_0, y_0) + \frac{(x - x_0)^2}{2}u_{xx}(\xi, y_0),$$

де ξ лежить між x і x_0 . Якщо тепер представити $x = x_0 + h$, то після деяких перетворень одержуємо

$$u_x(x_0, y_0) - \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} = -\frac{h}{2}u_{xx}(\xi, y_0).$$

Класичне визначення похідної функції однієї змінної можна записати:

$$\frac{dy}{dx} = \lim_{h \rightarrow 0} \frac{y(x+h) - y(x)}{h}$$

Очевидно, на ЕОМ неможливо провести граничний перехід. А тому із розкладу функції в ряд Тейлора в околі точки (x_0, y_0) можна наближено записати

$$u_x(x_0, y_0) = \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} \quad (11.3)$$

з помилкою обмежування $\xi_T = \frac{h}{2}u_{xx}(\xi, y_0)$; $x < \xi < x_0 + h$.

Вираз (11.3) називають *правою різницею*, а вираз

$$u_x(x_0 + h, y_0) = \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} \quad (11.4)$$

лівою різницею.

Користуючись (11.3) та (11.4), можна отримати різницевий вираз для 2-ї похідної

$$u_{xx}(x_0, y_0) = \frac{u(x_0 + h, y_0) - 2u(x_0, y_0) + u(x_0 - h, y_0)}{h^2} \quad (11.5)$$

і відповідно

$$u_{yy}(x_0, y_0) = \frac{u(x_0, y_0 + k) - 2u(x_0, y_0) + u(x_0, y_0 - k))}{k^2}, \quad (11.6)$$

де h, k – величина кроку відповідно за координатами x та y .

Помилка обмеження дорівнює

$$E_T = -\frac{k^2}{12} u_{yyy}(x_0, \eta); \quad y_0 - k \leq \eta \leq y_0 + k.$$

Аналогічно (11.5) і (11.6), можна вивести вираз для u_{xx} . Використовуючи ці вирази, загальновідоме рівняння Лапласа $u_{xx} + u_{yy} = 0$, наприклад, можна записати в кінцевих різницях у вигляді:

$$\frac{u(x_0 + h, y_0) - 2u(x_0, y_0) + u(x_0 - h, y_0))}{h^2} + \frac{u(x_0, y_0 + k) - 2u(x_0, y_0) + u(x_0, y_0 - k))}{k^2} = 0. \quad (11.7)$$

11.2 Етапи чисельного розв'язування диференціальних рівнянь (ДР) у частинних похідних на ЕОМ

1. У область розв'язування рівняння, яке розглядається, вводять рівномірну сітку $N \times M$ вузлових точок відповідно до характеру задачі і граничних умов.

2. Диференціальне рівняння, що розв'язується, записують у найзручнішій системі координат, замінюючи похідні кінцевими різницями (із таб. 11.1), приводять його до вигляду різницевого рівняння. Отримане різницеве рівняння використовують далі для описування функціонального зв'язку між сусідніми вузлами сітки. Тобто різницеве рівняння записують для усіх вузлів сітки і отримують внаслідок цього систему $(N-1)(M-1)$ рівнянь з $(N-1)(M-1)$ невідомими.

3. Систему $(N-1)(M-1)$ рівнянь з $(N-1)(M-1)$ невідомими розв'язують одним із відомих швидких і ефективних чисельних методів, вибір яких визначається деякими властивостями системи рівнянь (наприклад, система лінійних або нелінійних рівнянь та ін.).

Помітимо, що рівняння (11.7) можна представити схематично, накресливши п'ять вузлів різницевого рівняння (11.7) і позначивши біля кожного з них відповідний коефіцієнт. Рисунок (наприклад, рис. 11.2) називається *різницевою схемою*.

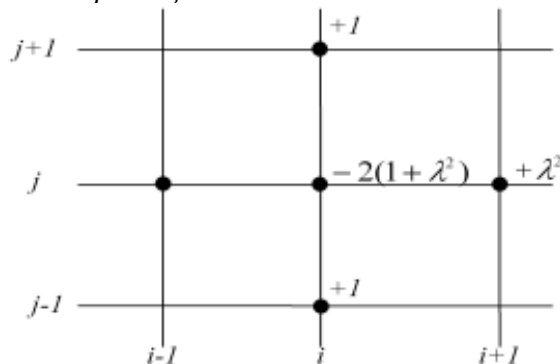


Рисунок 7.2 - Різницева схема

Рисунок 11.2 - Різницева схема

Різницева схема геометрично ілюструє різницеву апроксимацію диференціального рівняння. Відомо, що при $h \rightarrow 0$ і $k \rightarrow 0$ різницеве рівняння наближується до диференціального рівняння. Однак цікавить зовсім інше питання: чи наближається при $h \rightarrow 0$ і $k \rightarrow 0$ розв'язок різницевого рівняння до розв'язку диференціального рівняння? Для еліптичних рівнянь на це питання можна дати позитивну відповідь. Далі побачимо що у випадку параболічних і гіперболічних рівнянь прийдеться дотримуватись ряду обмежень, що забезпечать такого роду збіжність.

11.3 Еліптичні рівняння

До еліптичних відносять рівняння виду:

$$\text{рівняння Пуассона} \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2 \quad (11.8)$$

рівняння Лапласа $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ (11.9)

Розглянемо класичну задачу Діріхле

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

у деякій області R і $\psi = f(x, y)$ на границі цієї області, якою є крива G . Звичайно, рівняння (11.9) являє собою рівняння Лапласа, окремий випадок рівняння Пуассона.

Припустимо що область R має вигляд прямокутника ширини A і висоти B (рис. 11.3, 11.4.б).

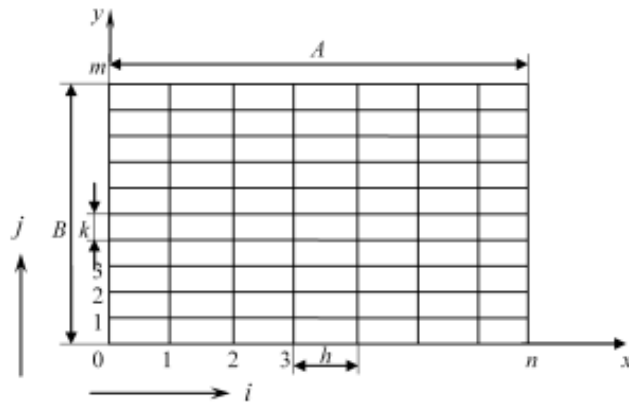


Рисунок 11.3 – Побудова сітки в прямокутній області

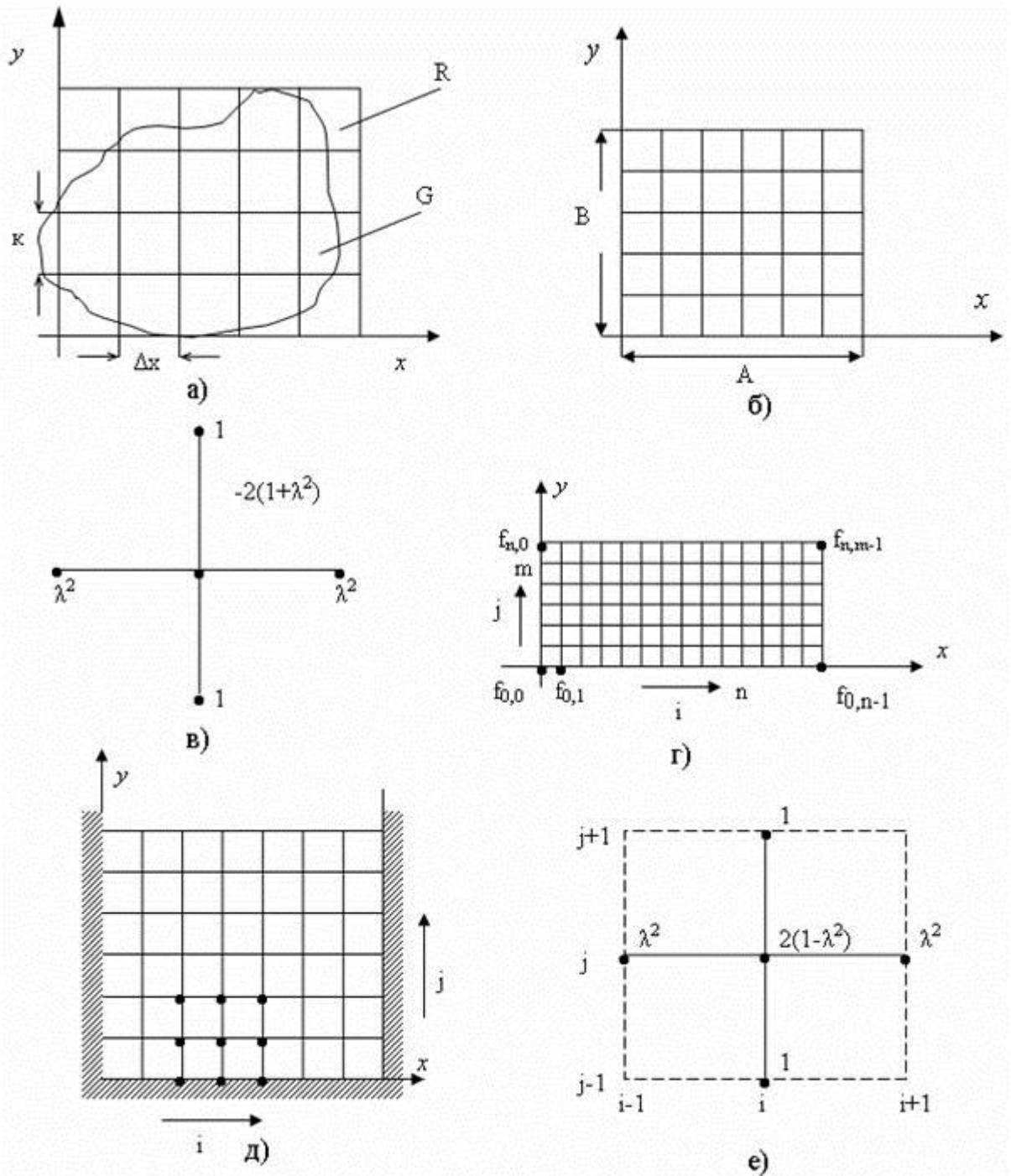


Рисунок 11.4 – Графічне представлення переходу до кінцево-різницевої формул

Розділимо спочатку ширину прямокутника A на n інтервалів, кожен розміром $h = A/n$; а висоту B – на m частин розмірами $k = B/m$. У середині області виходять при цьому $(n-1)(m-1)$ перерізів сітки. Визначимо різницеve співвідношення для кожної внутрішньої точки i розв'яжемо отриману систему рівнянь.

Таблиця 7.1 – Відношення між похідними та значеннями функції у вузлах сітки

Похідна	Схема	Наближена формула	Порядок похибки
---------	-------	-------------------	-----------------

$\frac{\partial u}{\partial x}$		$\frac{\partial u_{i,k}}{\partial x} \approx \frac{u_{i+1,k} - u_{i-1,k}}{2h}$ $\frac{\partial u_{i,k}}{\partial x} \approx \frac{u_{i,k} - u_{i-1,k}}{h}$ $\frac{\partial u_{i,k}}{\partial x} \approx \frac{u_{i+1,k} - u_{i,k}}{h}$	h^2 h h
$\frac{\partial u}{\partial y}$		$\frac{\partial u_{i,k}}{\partial y} \approx \frac{u_{i,k+1} - u_{i,k-1}}{2e}$ $\frac{\partial u_{i,k}}{\partial y} \approx \frac{u_{i,k+1} - u_{i,k}}{e}$ $\frac{\partial u_{i,k}}{\partial y} \approx \frac{u_{i,k} - u_{i,k-1}}{e}$	e^2 e e
$\frac{\partial^2 u}{\partial x^2}$		$\frac{\partial^2 u_{i,k}}{\partial x^2} \approx \frac{u_{i+1,k} - 2u_{i,k} + u_{i-1,k}}{h^2}$ $\frac{\partial^2 u_{i,k}}{\partial x^2} \approx \frac{1}{12h^2} (-u_{i+2,k} + 16u_{i+1,k} - 30u_{i,k} + 16u_{i-1,k} - u_{i-2,k})$ $\frac{\partial^2 u_{i,k}}{\partial x^2} \approx \frac{1}{h^2} (u_{i+2,k} - 2u_{i+1,k} + u_{i,k} + u_{i,k} - 2u_{i,k} + u_{i-1,k} - 2u_{i-1,k} + u_{i-2,k})$	h^2 h^4 h^4
$\frac{\partial^2 u}{\partial y^2}$		$\frac{\partial^2 u_{i,k}}{\partial y^2} \approx \frac{u_{i,k+1} - 2u_{i,k} + u_{i,k-1}}{e^2}$ $\frac{\partial^2 u_{i,k}}{\partial y^2} \approx \frac{1}{12e^2} (-u_{i,k+2} + 16u_{i,k+1} - 30u_{i,k} + 16u_{i,k-1} - u_{i,k-2})$ $\frac{\partial^2 u_{i,k}}{\partial y^2} \approx \frac{1}{3e^2} (u_{i,k+2} - 2u_{i,k+1} + u_{i,k} + u_{i,k} - 2u_{i,k} + u_{i,k-1} - 2u_{i,k-1} + u_{i,k-2})$	e^2 e^4 e^4

Нумерація точок перетинання сітки починається в горизонтальному напрямку зліво-направо, від нуля, крайня права точка буде при цьому n -й. Аналогічно у вертикальному напрямку нумеруємо точки знизу-вверх від нуля до m . Вузол з індексами (i, j) буде i -м ліворуч і j -м знизу.

Нехай початок координат збігається з точкою, координати якої $(0,0)$ позначимо

$$u(ih, jk) = u_{i,j}$$

аналогічно записується

$$f(ih, jk) = f_{i,j}$$

Використовуючи таку систему позначень граничну умову (11.10) можна записати у вигляді:

$$\begin{aligned}
u_{i,0} &= f_{i,0}; \quad i=0, 1, 2, \dots, n, \\
u_{i,m} &= f_{i,m}; \quad i=0, 1, 2, \dots, n, \\
u_{0,j} &= f_{0,j}; \quad j=0, 1, 2, \dots, m, \\
u_{n,j} &= f_{n,j}; \quad j=0, 1, 2, \dots, m.
\end{aligned} \quad (11.11)$$

Якщо позначити $\lambda = k/h$, то диференціальне рівняння (11.9) зведеться до різницевого рівняння вигляду

$$\lambda^2 u_{i+1,j} + \lambda^2 u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 2(1 + \lambda^2) u_{i,j} = 0 \quad (11.12)$$

для $i=1, 2, 3, \dots, n-1$ та $j=1, 2, 3, \dots, m-1$.

При $\lambda=1$, тобто при однакових величинах інтервалів розбивки в горизонтальному і вертикальному напрямках, це співвідношення означає, що значення $u_{i,j}$ є середнім арифметичним з чотирьох сусідніх з ним:

$$u_{i,j} = \frac{1}{4} (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) \quad (11.13)$$

Важливо те, що одержана система – це система лінійних алгебраїчних рівнянь. Усього є $(m-1)(n-1)$ рівнянь щодо $(m+1)(n+1)$ невідомих. Після того як $2(m+n)$ невідомих будуть виключені за допомогою граничної умови (11.11), залишається точно $(m-1)(n-1)$ рівнянь щодо $(m-1)(n-1)$ невідомих. Вираз (11.7) являє собою систему лінійних алгебраїчних $(n-1)(m-1)$ рівнянь відносно $(n+1)(m+1)$ невідомих. Після того, як $2(n+m)$ невідомі будуть виключені за допомогою граничних умов, залишається $(n-1)(m-1)$ рівняння відносно $(n-1)(m-1)$ невідомих.

У (11.11) і (11.12) кожна пара значень (i, j) визначає вузол, в якому рівняння (11.6) розв'язується відносно $u_{i,j}$. Рівняння (11.12) можна легко запам'ятати, якщо знати різницеву схему п'яти вузлів (рис. 11.2, 11.4.д,е).

Розпишемо докладно деякі з рівнянь (11.12). Для зручності прийнявши $\lambda=1$, тобто $h=k$, але спільність висновків не важко перевірити при кожному $\lambda > 0$. Починаючи з $i=1, j=1$ і при незмінному j пройдемо значення $i=1, 2, \dots, n-1$:

$$\begin{aligned}
4u_{1,1} - u_{2,1} - u_{1,2} &= f_{1,0} + f_{0,1}; \\
-u_{1,1} + 4u_{2,1} - u_{3,1} - u_{2,2} &= f_{2,0}; \\
-u_{2,1} + 4u_{3,1} - u_{4,1} - u_{3,2} &= f_{3,0}; \\
&\dots\dots\dots \\
-u_{n-2,1} + 4u_{n-1,1} - u_{n,1,2} &= f_{n-1,0} + f_{n,1}.
\end{aligned} \quad (11.14)$$

Тепер збільшуючи j до 2, пройдемо значення при $i=1, 2, \dots, n-1$:

$$\begin{aligned}
-u_{1,1} + 4u_{1,2} - u_{2,2} - u_{1,3} &= f_{0,2}; \\
-u_{2,1} - u_{1,2} + 4u_{2,2} - u_{3,3} - u_{2,3} &= 0; \\
-u_{3,1} - u_{2,2} + 4u_{3,2} - u_{4,2} - u_{3,3} &= 0; \\
&\dots\dots\dots \\
-u_{n-2,1} - u_{n-3,2} + 4u_{n-2,2} - u_{n-1,2} - u_{n-2,3} &= 0; \\
-u_{n-1,1} - u_{n-2,2} + 4u_{n-1,2} - u_{n-1,3} - u_{n-1,3} &= f_{n,2}.
\end{aligned} \quad (11.15)$$

В такий же спосіб, збільшуючи щоразу j і проходячи значення $i=1, 2, \dots, n-1$. Кінцеве значення j буде дорівнювати $m-1$.

Система (11.14), (11.15) має дві важливі властивості, що допомагають вибрати для неї метод розв'язку. Властивості ці наступні:

1. Переважаюча частина коефіцієнтів системи дорівнює нулю.

2. У кожному рівнянні один з коефіцієнтів дорівнює + 4. Якщо в рівнянні є п'ять коефіцієнтів, відмінних від нуля, то сума інших чотирьох коефіцієнтів дорівнює - 4, якщо ж кількість ненульових коефіцієнтів менше п'яти, то сума інших дорівнює - 2 чи - 3.

Таким чином, у цій системі виконані умови збіжності ітераційного методу Гаусса - Зейделя (на основі її другої властивості). Перша ж властивість системи робить рішення методом виключення дуже незручним: вихідна система з великою кількістю рівних нулю коефіцієнтів перетвориться після виключення невідомих у повну трикутну систему.

Розглянемо деякі рівняння в тому вигляді, у якому з ними будуть проводитися ітерації. Позначаючи верхніми індексами порядковий номер ітерації і припускаючи, $u_{i,j}^{(0)} = 0$ для усіх i, j , одержимо наступний порядок розв'язання системи рівнянь:

$$\begin{aligned}
 u_{1,1}^{(1)} &= \frac{1}{4} [f_{1,0} + f_{0,1} + u_{2,1}^{(0)} + u_{1,2}^{(0)}] \\
 u_{2,1}^{(1)} &= \frac{1}{4} [f_{2,0} + u_{1,1}^{(1)} + u_{3,1}^{(0)} + u_{2,2}^{(0)}] \\
 u_{3,1}^{(1)} &= \frac{1}{4} [f_{3,0} + u_{2,1}^{(1)} + u_{4,1}^{(0)} + u_{3,2}^{(0)}] \\
 &\dots\dots\dots \\
 u_{n-1,1}^{(1)} &= \frac{1}{4} [f_{n-1,0} + f_{n,1} + u_{n-2,1}^{(1)} + u_{n-1,2}^{(0)}] \\
 u_{1,2}^{(1)} &= \frac{1}{4} [f_{0,2} + u_{1,1}^{(1)} + u_{2,2}^{(0)} + u_{1,3}^{(0)}] \\
 u_{2,1}^{(1)} &= \frac{1}{4} [u_{2,1}^{(1)} + u_{1,2}^{(1)} + u_{3,2}^{(0)} + u_{2,3}^{(0)}] \\
 &\dots\dots\dots
 \end{aligned} \tag{11.16}$$

Для цієї системи рівнянь виконується умова зближення ітераційного методу Гауса - Зейделя. Щоб розв'язати цю систему, розроблено алгоритм (рис. 11.5), який використовує *метод Гауса-Зейделя*, який застосовують до еліптичних рівнянь різницевих, називають *методом Лібмана* (або методом послідовних зсувів).

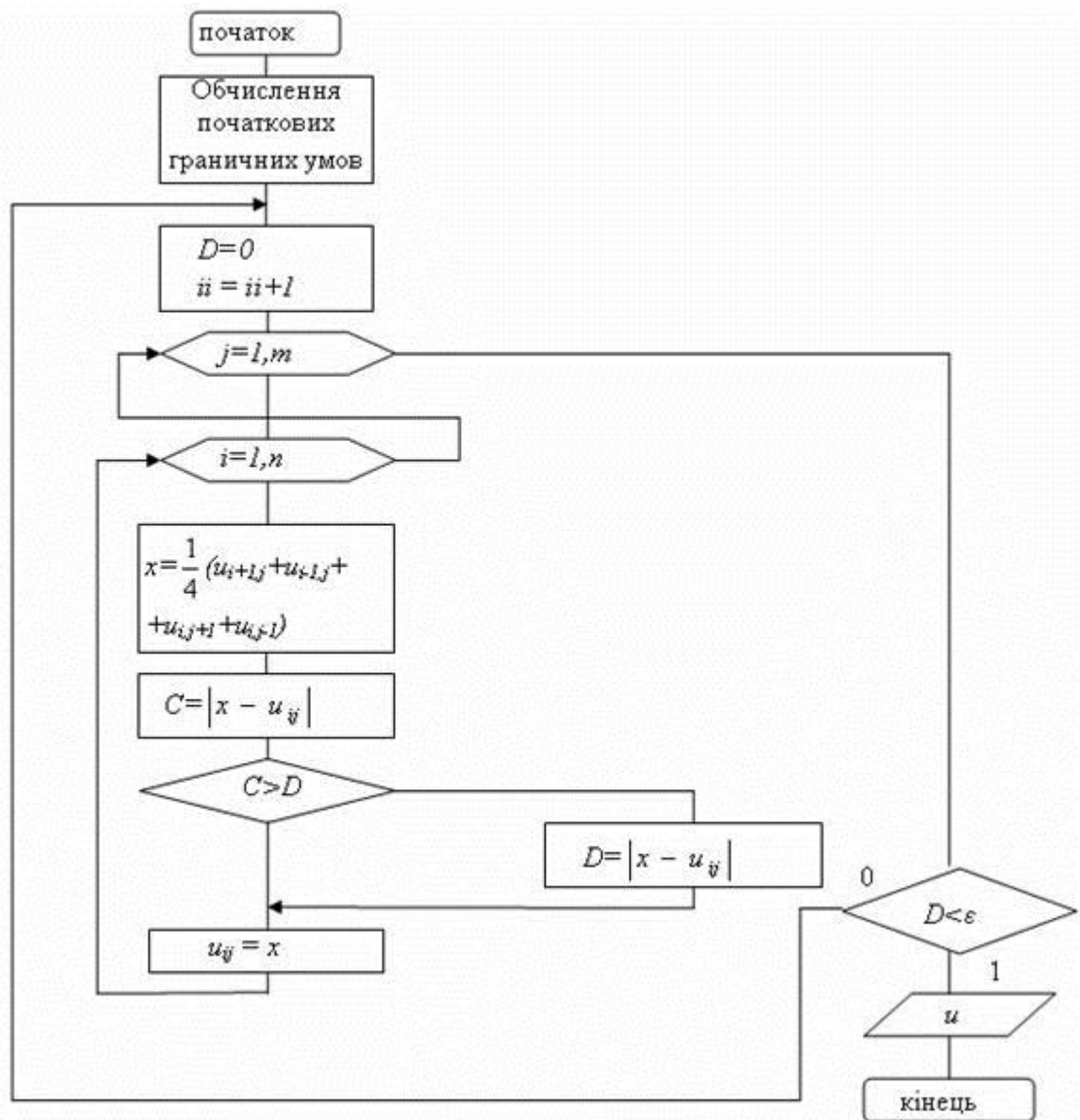


Рисунок 11.5 – Схема алгоритму розв'язання еліптичного рівняння на ЕОМ

При порівнянні реальних систем ДР буває корисно збільшити чи зменшити чергове виправлення при розрахунку чергового наближення до кореня рівняння (аналогічний метод може виявитися корисним при розв'язанні систем лінійних алгебраїчних рівнянь). У дійсності це справедливо і для методу Лібмана. Щоб можна було збільшувати чи зменшувати чергові виправлення, замінимо в алгоритмі (рис. 11.5) визначення $u_{i,j}$ за формулою:

$$u_{i,j} = \omega x + (1 - \omega) u_{i,j}$$

де x – значення $u_{i,j}$, обчислене з рівняння (11.12). Параметр ω називається *прискорюючим множником*. У загальному випадку вимагається, щоб значення ω лежало в межах від 1 до 2. Якщо $\omega = 1$, то є звичайний метод Лібмана, при $\omega > 1$ є прискорений метод Лібмана.

Метод Лібмана було розглянуто тільки для випадку рівняння Лапласа. Взагалі кажучи, будь-яке еліптичне рівняння без членів, що містять u_{xy} , зводиться до системи різницевих рівнянь, що задовільняє умовам збіжності.

Усе сказане дотепер відносилось до лінійних рівнянь. Питання про збіжність для нелінійних рівнянь розроблено дуже слабо. Потрібно відзначити, однак, що багаторазово робилися успішні спроби вирішення квазілінійних (квазілінійними називаються рівняння, лінійні по u_{xx} , u_{yy} і u_{xy} , але нелінійні по u_x , u_y , i , x , y) рівнянь за допомогою екстрапольованого методу Лібмана. Загальний підхід у цьому випадку зводиться до того, що задається навмання деяке значення і спостерігається швидкість збіжності ітераційного процесу. Потім, змінюючи ω і спостерігаючи викликані цим зміни швидкості збіжності, можна підібрати таке її значення, при якому

ітераційний процес сходиться швидше всього. Вдавалося одержати збіжність навіть для рівнянь із членом типу u_{xy} , хоча в цьому випадку немає ніяких теоретичних основ очікування збіжності.

Є багато інших способів вирішення різницевих рівнянь (11.12). Найбільш часто використовуються *лінійна релаксація, блокова релаксація і метод напрямків*, що чергуються. Часто вони виявляються більш ефективними, ніж метод Лібмана. Нарешті, можна використовувати екстраполяційний перехід до границі.

11.4 Гіперболічні рівняння

Знайомство з даним типом диференційних рівнянь у частинних похідних почнемо з розгляду відповідної фізичної задачі.

Припустимо, що струна довжиною L , натягнута між двома точками осі x , точкою $x = 0$ і точкою $x = L$. Натяг струни дорівнює T . Якщо відхилити струну від положення рівноваги і відпустити, то вона почне коливатися. Зсув кожної точки струни щодо положення рівноваги залежить не тільки від координати x цієї точки, але і від часу t . Відхилення струни від положення рівноваги описується рівнянням гіперболічного типу, що наводиться без виведення

$$u_{xx} - a^2 u_{tt} = 0$$

для $0 < x < L$ і $t > 0$. Коефіцієнт a враховує фізичні характеристики Тут? – вага струни на одиницю довжини, T – натяг, g – прискорення сили ваги. Це рівняння звичайне називається хвильовим рівнянням. Для простоти $a = 1$, так що задача виразиться в такий вид:

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2} \quad (11.17)$$

$$\left. \begin{array}{l} 0 \leq x \leq L, \\ t > 0, \end{array} \right\} \quad (11.18)$$

$$\left. \begin{array}{l} u(0, t) = u(L, t) = 0, \quad t \geq 0, \\ u(x, 0) = f(x), \quad 0 \leq x \leq L, \\ u_t(x, 0) = g(x), \quad 0 \leq x \leq L, \end{array} \right\} \quad (11.19)$$

$$a^2 = \frac{\omega}{Tg}$$

Таке формулювання задачі зовсім не означає втрати спільності, тому що проста заміна змінних зводить будь-яке хвильове рівняння до виду (11.17).

Оскільки кінці струни закріплені, то

$$u(0, t) = u(L, t) = 0, \quad t \geq 0, \quad (11.20)$$

початковими умовами є початковий зсув

$$u(x, 0) = f(x), \quad 0 \leq x \leq L \quad (11.21)$$

та початкова швидкість

$$u_t(x, 0) = g(x); \quad 0 < x < L \quad (11.22)$$

Наприклад, якщо відтягнути струну за середину, як показано на рис. 11.6 і відпустити її без додавання їй початкової швидкості, то початкові умови запишуться у вигляді

$$f(x) = \begin{cases} \frac{2\delta}{L}x, & 0 \leq x \leq \frac{L}{2}, \\ \frac{2\delta}{L}(L-x), & \frac{L}{2} \leq x \leq L, \end{cases}$$

$$g(x) = 0, \quad 0 < x < L.$$

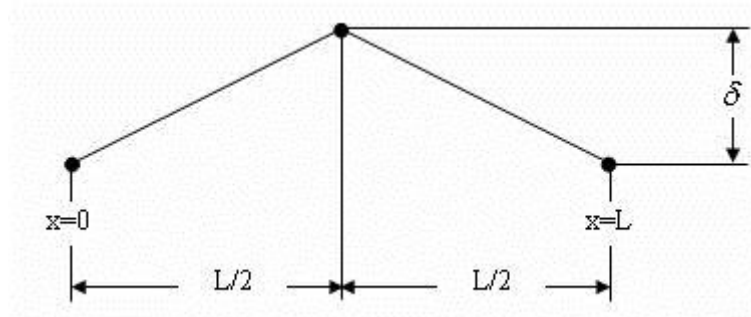


Рис. 11.6 – Можливі початкові умови для коливальної струни, описуваної хвильовим рівнянням.

Помітимо, що ці умови називаються *початковими*, а не *граничними*. У дійсності, якщо задати для гіперболічного рівняння граничні умови, то отримана задача не буде мати єдиного розв'язку. Подібні задачі називаються *некоректно поставленими*. Дуже важливо, щоб додаткові умови належним чином відповідали кожному типу рівнянь.

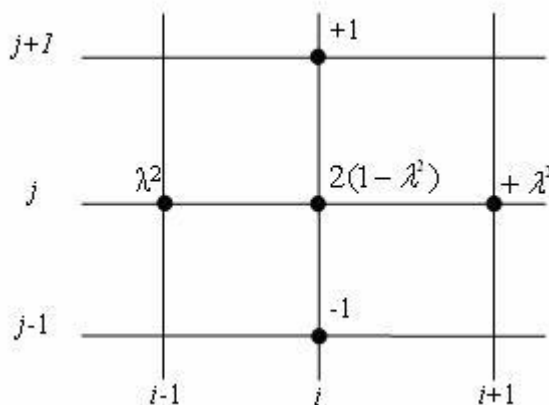


Рисунок 11.7 – Різницева схема для гіперболічного рівняння

Щоб знайти різницеві рівняння, що відповідають (11.17), , скористаємося знову рівностями (11.18) і (11.19), причому замість u записується t . Знову накреслимо сітку, але тепер ця сітка простирається нескінченно в напрямку позитивних значень t . Можна шукати розв'язок для як завгодно далекого моменту часу. У напрямку x приймаємо крок сітки рівним h , у напрямку t – рівним k . Тому інтервал L розділяється на $n = L/h$ малих інтервалів h , а в напрямку t може бути як завгодно багато інтервалів k .

Якщо позначити

$$u_{i,j} = u(ih, jk) \quad \lambda = \frac{k}{h},$$

то різницеве рівняння запишеться у вигляді

$$u_{i,j+1} = 2(1 - \lambda^2)u_{i,j} + \lambda^2(u_{i+1,j} + u_{i-1,j}) - u_{i,j-1} \quad (11.23)$$

для $i = 1, 2, \dots, n$ і для $j = 1, 2, \dots$.

Різницеву схему при цьому виглядає так (рис.11.7):

Граничну умову (11.20) легко записати у вигляді

$$u_{0,j} = u_{n,j} = 0; \quad j = 1, 2, 3, \dots \quad (11.24)$$

Початкову умову (11.21) можна написати у вигляді

$$u_{i,0} = f(ih); \quad i = 1, 2, \dots, n-1 \dots \quad (11.25)$$

Щоб записати в різницевому вигляді початкову умову (11.22), можна використовувати рівність (11.3), тоді отримаємо:

$$\frac{u_{i,1} - u_{i,0}}{k} = g(ih)$$

Після цього, використовуючи (11.25), одержуємо

$$u_{i,1} = f(ih) + kg(ih) \quad (11.26)$$

Помітимо тепер, що (11.25) і (11.26) дають значення $u_{i,j}$ для перших двох рядків: ($j = 0$ і $j = 1$). Підставляючи $j = 1$ у (11.23), одержимо

$$u_{i,2} = 2(1 - \lambda^2)u_{i,1} + \lambda^2(u_{i+1,1} + u_{i-1,1}) - u_{i,0} \quad (11.27)$$

Усі доданки в правій частині цього рівняння включають значення i тільки з перших двох рядків сітки; всі ці значення відомі з початкових умов. Тому в останньому рівнянні є тільки одне невідоме і всі значення функції, що відповідають третьому рядку, можна обчислити в явному вигляді. Після цього можна обчислювати значення функції в четвертому рядку, виходячи зі значень у другому і в третьому рядках, і т.д. стільки разів, скільки буде потрібно. При цьому навіть не приходиться вирішувати систему рівнянь.

Таким чином, (11.27) являє собою явну схему рішення хвильового рівняння. Схема алгоритму представлена на рис. 11.8.

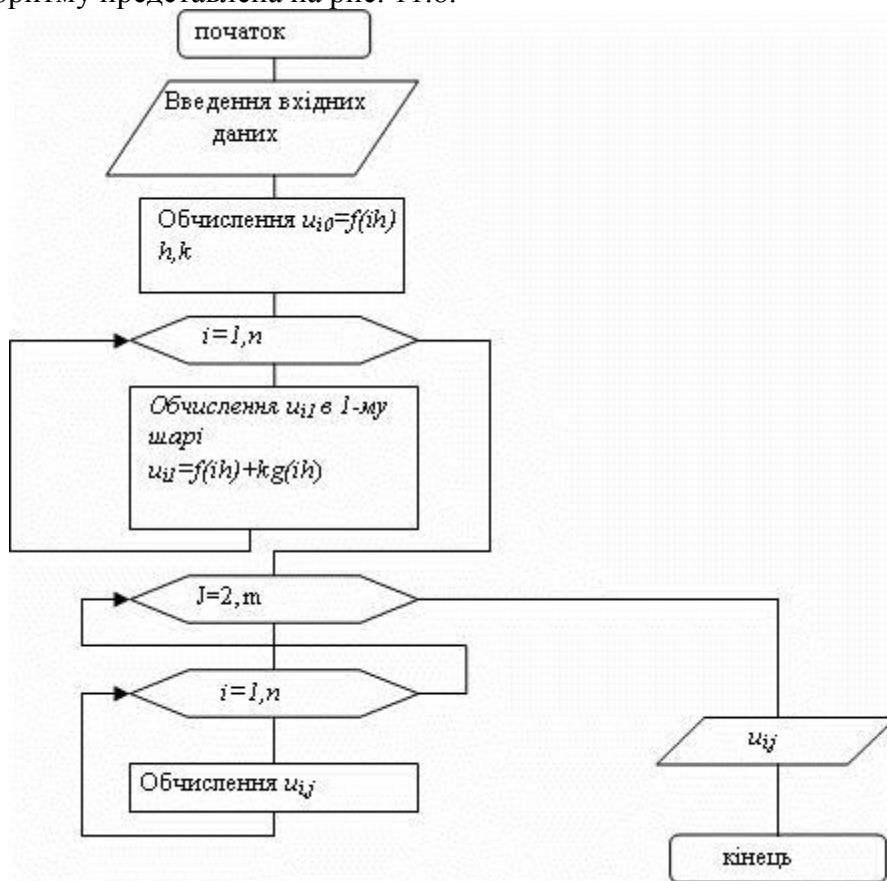


Рисунок 11.8 – Схема алгоритму розв'язування гіперболічного рівняння на ЕОМ

При роз'язанні еліптичного рівняння (11.13) було використано схему, в якій в кожному рівнянні було більш ніж по одному невідомому; таким чином, цей метод можна було назвати *неявним*.

Розглянемо тепер питання *збіжності і стійкості методу* (не приводячи доведень, а обмежуючись тільки аналізом остаточних результатів). Можна стверджувати, що рішення (11.27) сходиться (мається на увазі, що при $h > 0$ і $k > 0$ розв'язання різницевого рівняння асимптотично наближається до розв'язання диференціального рівняння), якщо

$$\lambda < 1, \text{ чи, що те ж саме, } k < h < i = \dots < /h >$$

Ця умова є достатньою для збіжності; але вона не є необхідною. Іншими словами, існують рівняння і величини інтервалів, при яких ця умова не виконується, але усе-таки виходить вірний результат. Уся справа в тім, що тоді не можна гарантувати збіжність чисельного розв'язку.

Таким чином, як тільки обрана величина інтервалу розбивки h у напрямку x , то з'являється обмеження на величину інтервалу по t . Якщо необхідно зробити обчислення для великого відрізка t , то може знадобитися велика кількість кроків.

Іншою настільки ж важливою обставиною є те, що при $\lambda > 1$, метод стає нестійким, як в абсолютному, так і у відносному змісті. Це означає, як і для звичайних диференціальних рівнянь, що будь-які помилки зростають у ході обчислення розв'язку.

МЕТОДИ ОПТИМІЗАЦІЇ ОДНОВИМІРНИХ ЗАДАЧ

Процес оптимізації лежить в основі всієї інженерної діяльності, оскільки класичні функції інженера полягають у тому, щоб, з одного боку, проектувати нові, ефективніші та менш дорогі технічні системи, а з другого - розробляти методи підвищення якості функціонування існуючих систем. Ефективність методів оптимізації, які дозволяють вибирати найкращий варіант без перевірки всіх можливих варіантів, тісно пов'язана із широким використанням досягнень у галузі математики шляхом реалізації ітераційних обчислювальних схем, що опираються на досить обґрунтовані методи та алгоритми із застосуванням обчислювальної техніки.

В даний час, в зв'язку з доступністю персональних комп'ютерів, велика увага приділяється використанню чисельних методів оптимізації в інженерній практиці, які можливо розділити на дві великі групи: *методи безумовної і умовної оптимізації*. Цей розподіл пов'язаний з різним описом простору проектування. Область дослідження, тобто область, в якій інженер намагається визначити оптимум певної задачі, називаються *простором проектування*.

Простір проектування, який визначається проектними параметрами, зазвичай обмежений низкою умов, пов'язаних з фізичною сутністю задачі і розглядається в у вигляді двох варіантів:

- 1) якщо проектний параметр один, то зазвичай обмеження пов'язані з його значеннями, тобто область проектування звужується до відрізка дослідження $[a, b]$;
- 2) якщо проектних параметрів декілька, то обмеження можуть накладатись або на їх значення, обмежуючи область дослідження, або в вигляді взаємозалежності між проектними параметрами, які повинні враховуватись при пошуку рішення (ці залежності в реальних задачах можуть відображати закони природи, економіки, права, наявність необхідних матеріалів і т. ін.).

В даному розділі розглядаються методи безумовного одновимірного пошуку оптимуму цільової функції, які базуються на використанні першого варіанту представлення простору проектування, де *цільова функція* – це вираз (функція), значення якого інженер намагається мінімізувати або максимізувати. При цьому передбачається, що цільові функції, які досліджуються є *унімодальними*, тобто мають на інтервалі дослідження, який розглядається тільки один оптимум (рис. 13.1). Таке обмеження на характер цільової функції не є таким жорстким, як може здатися, так як багато задач, з якими інженер стикається в своїй практиці, виявляються унімодальними.

Чисельні методи, які орієнтуються на розв'язання задач безумовної оптимізації, можна розділити на три класи:

- *методи прямого пошуку*, що базуються на обчисленні тільки значень цільової функції;

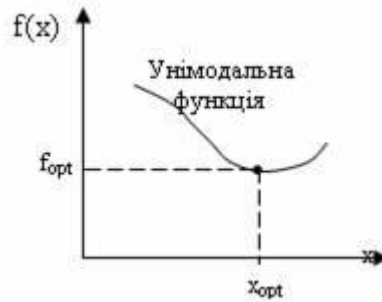


Рисунок 13.1 – Унімодальна цільова функція

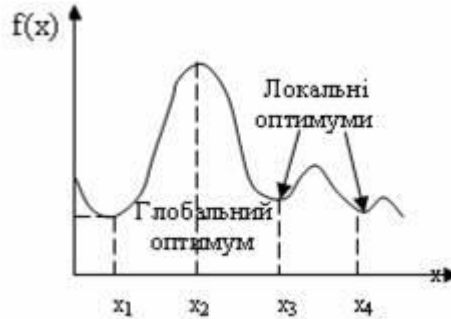


Рисунок 13.2 – Цільова функція з локальним та глобальним оптимумом

- *градієнтні методи*, в яких використовуються точні значення перших похідних $f'(x)$;

- *методи другого порядку*, в яких поряд з першими похідними використовуються також другі похідні функції $f(x)$.

Задача одновимірної оптимізації ставиться таким чином: значення параметра X цільової функції $f(x)$, який називають *проектним параметром*, знаходяться на інтервалі дослідження $[a, b]$. В процесі пошуку оптимуму цільової функції цей інтервал, який називається *інтервалом невизначеності*, постійно зменшується (звужується), тому методи одновимірної оптимізації іноді називають методами звуження інтервалу невизначеності.

Вибір чисельного методу в першу чергу залежить від виду цільової функції, яка може бути однопараметричною і багатопараметричною (рис. 13.3, 13.4).

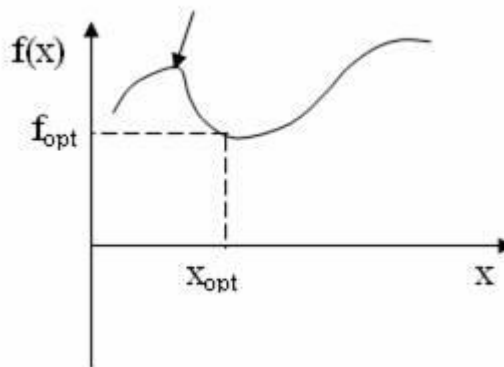


Рисунок 13.3 – Однопараметрична цільова функція

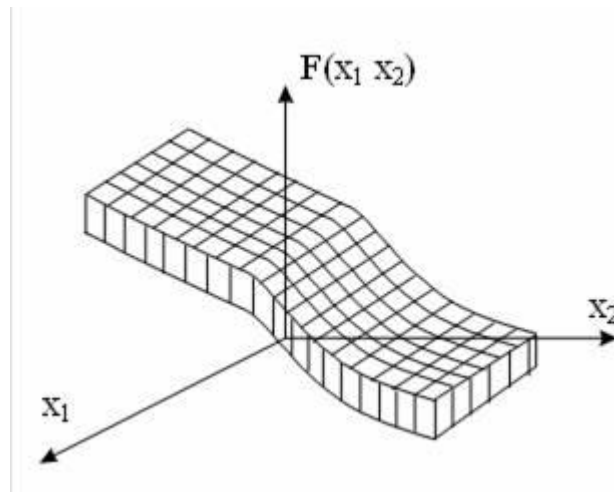


Рисунок 13.4 – Двопараметрична цільова функція

Деякі алгоритми оптимізації пристосовані до пошуку максимуму, а інші – для пошуку мінімуму.

Однак, незалежно від типу задачі, яка розв’язується на екстремум (оптимум) можливо користуватись одним і тим же алгоритмом, так як задачу мінімізації можливо легко переробити в задачу на пошук мінімуму, змінивши знак цільової функції на протилежний (рис. 13.5).

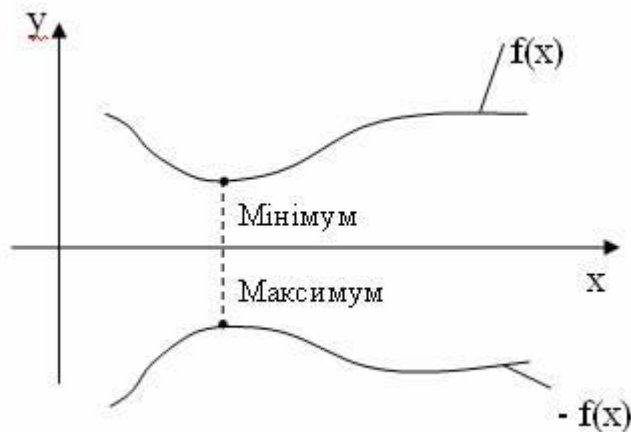


Рисунок 13.5 – Зміною знаку цільової функції на протилежний задача на мінімум перетворюється в задачу на максимум

Загальна постановка задачі для методів одновимірної оптимізації ставиться наступним чином: нехай значення параметра X знаходиться на відрізку $[a, b]$, а цільова функція унімодальна в області, яку досліджуємо. Більшість чисельних методів одновимірної оптимізації - це методи звуження відрізка, а саме: метод розділення відрізка навпіл, метод дихотомії, метод золотого перерізу, метод Фібоначчі.

В процесі одновимірної оптимізації цільової функції на ЕОМ можна виділити два етапи:

- 1) встановлення меж відрізка, на якому реалізується процедура пошуку оптимуму;
- 2) зменшення відрізка до заданої похибки обчислення ϵ .

Перший етап реалізується за допомогою евристичних методів пошуку і є дуже складним. Другий - називають правилом виключення відрізків, реалізують алгоритм пошуку, що дозволяє знайти точку оптимуму шляхом послідовного виключення частин початкового обмеженого відрізка $[a, b]$, тобто за допомогою ітераційних алгоритмів. В якості умови закінчення ітераційного процесу використовується момент, коли підінтервал, що залишився, зменшиться до достатньо малих розмірів (зазвичай для цього задають значення заданої похибки обчислення ϵ).

13.1 Метод загального пошуку

Очевидно, найбільш природнім способом звуження інтервалу невизначеності для одновимірної унімодальної функції є ділення його на декілька рівних частин з наступним обчисленням значень цільової функції в вузлах отриманої сітки (рис. 13.6).

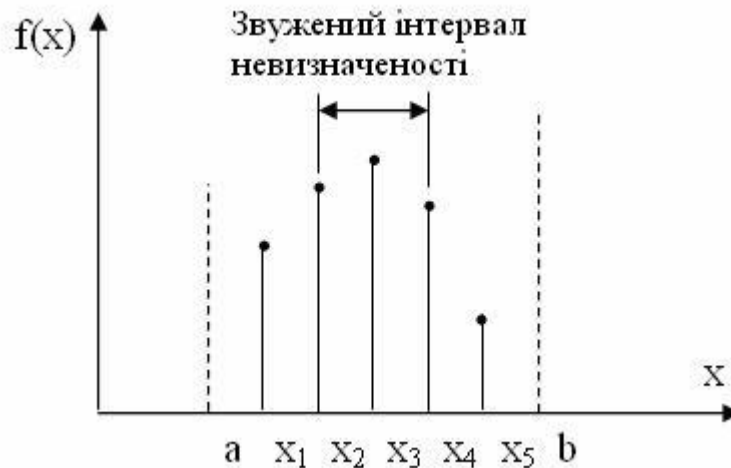


Рисунок 13.6 – Метод загального пошуку

В результаті інтервал невизначеності звужується до двох кроків сітки. Звичайно говорять про дроблення інтервалу невизначеності, яке характеризується коефіцієнтом f . Розділивши інтервал невизначеності на N частин, отримаємо $N+1$ вузол, і тоді

$$f = \frac{2}{N+1}.$$

Щоб отримати значення $f = 0,01$, необхідно обчислити цільову функцію в 199 точках, а при $f = 0,001$ $N=1999$. Звідси видно, що ефективність цього методу при зменшенні інтервалу невизначеності швидко падає. Напрошується інший варіант: щоб отримати $f=0,01$, необхідно обчислити спочатку функцію в 19 точках і отримати $f = 0,1$, а потім обчислити ще 19 значень функції на скороченому інтервалі невизначеності, отримати $f = 0,01$, зробивши при цьому всього 38, а не 199 обчислень. Таким чином, при деякій винахідливості ефективність пошуку можна різко збільшити.

13.2 Метод половинного ділення (розділення відрізка навпіл)

Суть метода полягає в постійному діленні відрізка дослідження цільової функції $[a, b]$ навпіл і визначенні на ньому координат трьох точок x_1, x_2, x_m . При чому значення їх визначаються як:

$$x_m = \frac{a+b}{2}, \quad L=b-a, \quad x_1 = a + \frac{L}{4}, \quad x_2 = b - \frac{L}{4}.$$

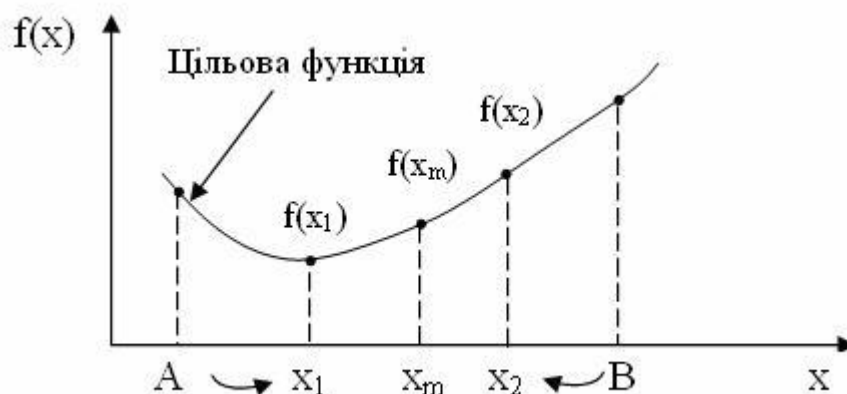


Рисунок 13.7 - Геометрична інтерпретація методу половинного ділення

Точки x_1, x_m, x_2 , поділяють відрізок $[a, b]$ на чотири рівні частини (рис. 13.6), обчислюємо значення цільової функції $f(x_1), f(x_2)$. Потім порівнюємо значення $f(x_1)$ і $f(x_m)$, якщо $f(x_1) < f(x_m)$, то виключаємо з дослідження відрізок $[x_m, b]$ та покладемо $b = x_m$. Тоді середньою точкою нового відрізка $[a, b]$ стає x_1 ($x_m = x_1$). Але якщо $f(x_1) \geq f(x_m)$, то порівнюємо значення цільової функції $f(x_2)$ і $f(x_m)$; якщо $f(x_2) < f(x_m)$, то виключаємо відрізок $[a, x_m]$, покладемо $a = x_m, x_m = x_2$; якщо $f(x_2) \geq f(x_m)$, то виключаємо відрізок $[a, x_1]$ та $[x_2, b]$, покладемо $a = x_1, b = x_2$, тобто формуємо новий відрізок дослідження. Обчислюємо $L = b - a$, якщо $|L| < \varepsilon$, якщо немає, то знову повертаємося до початку.

Даний алгоритм ітераційний, тому зазвичай в якості умови закінчення ітераційного процесу обирають умову $|a - b| < \varepsilon$, тобто звуження відрізка виконується до тих пір, поки його величина не зменшиться до заданої обчислювальної похибки ε . Алгоритм методу представлений на рис. 13.8.

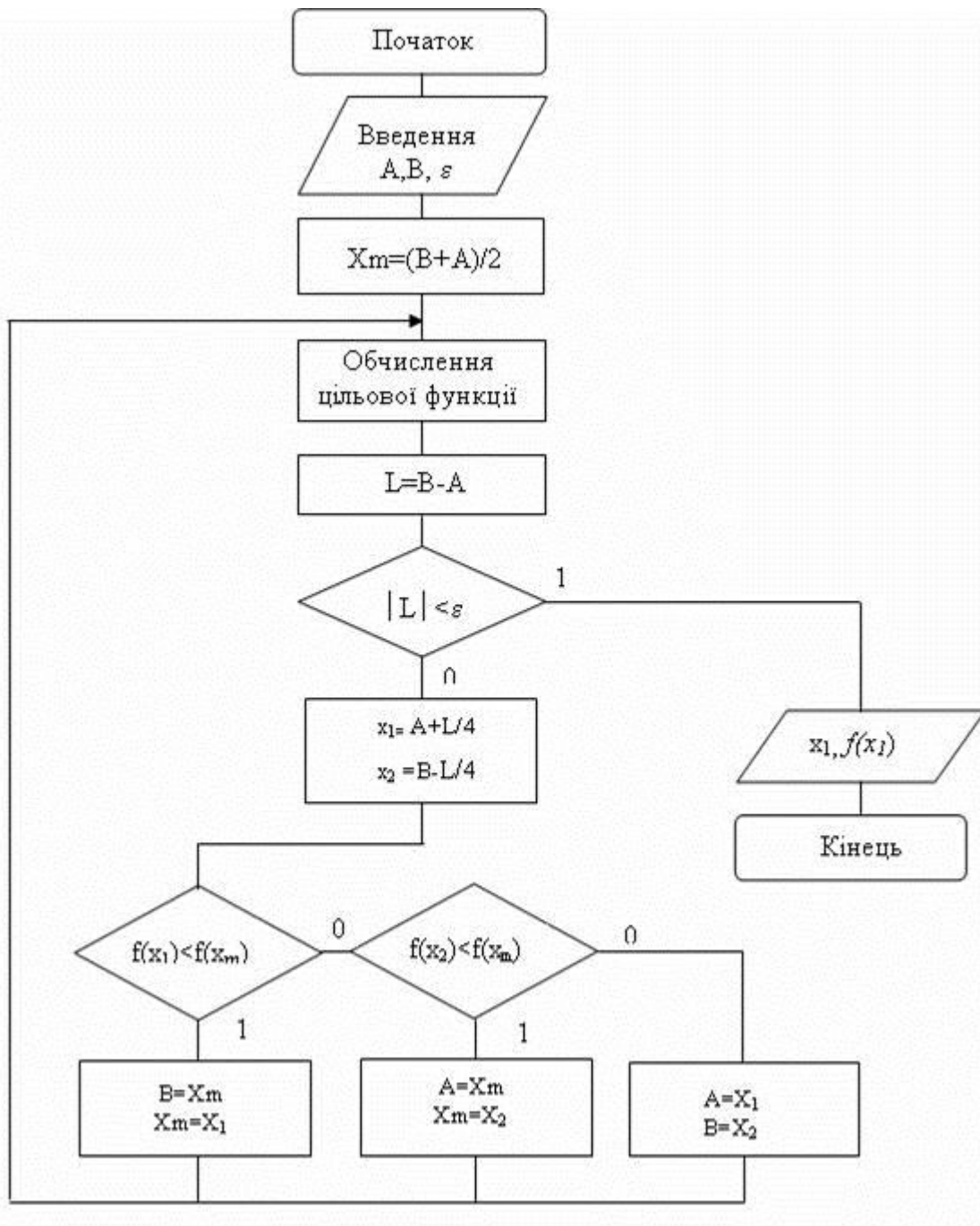


Рисунок 13.8 – Схема алгоритму метода половинного ділення

Цей метод відрізняється великою ефективністю.

13.3 Метод дихотомії

Обчислення цільової функції в двох точках інтервалу невизначеності дозволяє його звужити. Можна таким чином обрати ці точки, що інтервал невизначеності буде мінімальним. На рис. 13.9 показані позначення, які використовуються в цій схемі.

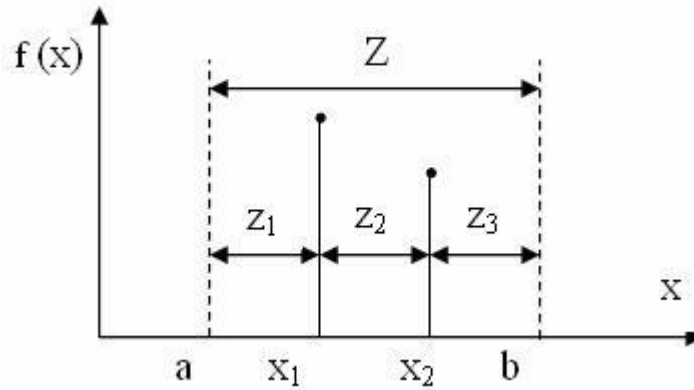


Рисунок 13.9 – Позначення, які використовуються в методі дихотомії

Якщо значення цільової функції при x_1 більше, ніж при x_2 , то новий інтервал невизначеності дорівнює $Z_1 = z_1 + z_2$. В протилежному випадку він визначається виразом $Z_2 = z_2 + z_3$. Задача полягає в тому, щоб одночасно мінімізувати Z_1 і Z_2 , задовольнивши умовам

$$z_1 + z_2 + z_3 = Z,$$

$$z_1 > 0,$$

$$z_2 > 0,$$

$$z_3 > 0.$$

Із рівності можна виключити z_2 . Тоді

$$Z - z_3 = \min, \quad Z - z_1 = \min.$$

Так як величина Z задана, то праві частини цих рівнянь будуть тим менше, чим більше z_1 і z_3 . Отже, оптимум відповідає умові

$$z_1 = z_3 = 0,5Z.$$

Але тоді $z_2 = 0$, що суперечить умові $z_2 > 0$.

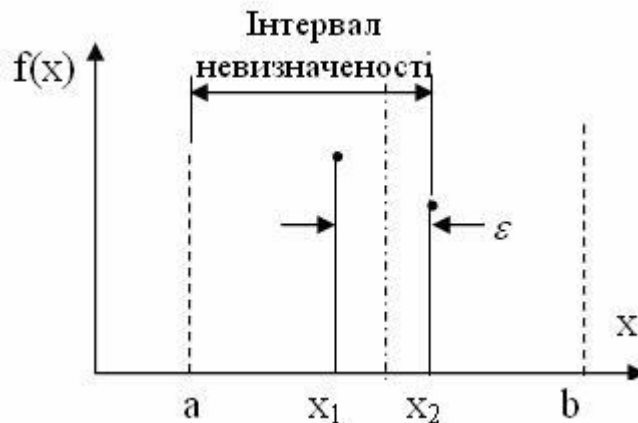


Рисунок 13.10 – Метод дихотомії

Нехай z_2 має деяке дуже маленьке значення ε . Тоді із z_1 і z_3 віднімемо по $\frac{\varepsilon}{2}$. В результаті після обчислення першої пари значень цільової функції при близьких значеннях x інтервал невизначеності звужиться, як показано на рис. 13.10, і коефіцієнт ділення буде дорівнювати

$$f = \frac{l}{2} + \frac{\varepsilon}{2}.$$

В межах, при $\varepsilon \rightarrow 0$, $f \rightarrow \frac{l}{2}$. Надалі при використанні метода дихотомії виконуються ті операції, що і при використанні методу ділення відрізка навпіл.

13.4 Метод “золотого перетину”

З кожних трьох значень цільової функції, які були обчислені в інтервалі невизначеності в подальшому використовуються лише два, а третє не дає додаткової інформації і в подальшому не використовується. В методі золотого перетину цільова функція обчислюється в точках інтервалу невизначеності, які розташовані таким чином, щоб кожне обчислене значення цільової функції давало нову корисну інформацію.

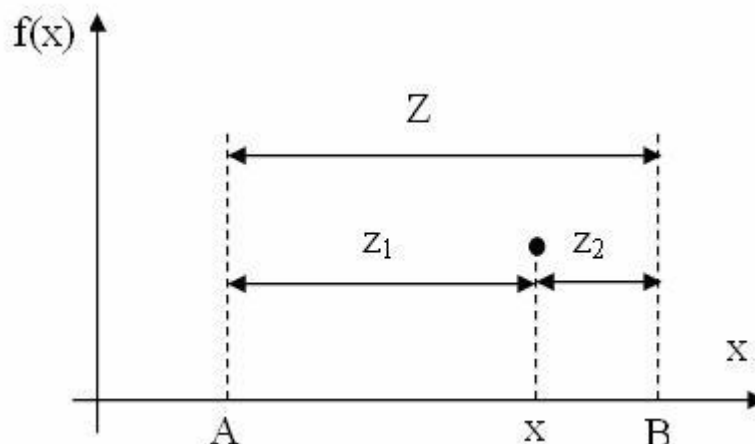


Рисунок 13.11 – Позначення, які використовуються в методі золотого перетину

Суть цього методу полягає в наступному. Інтервал невизначеності ділиться на дві нерівні частини, відношення довжини більшого відрізка до довжини всього інтервалу дорівнює відношенню довжини меншого відрізка до довжини більшого відрізка. На рис. 13.11 показаний інтервал невизначеності Z , який складається з відрізків z_1 і z_2 , відношення довжин яких визначається правилом золотого перетину.

$$\frac{z_1}{Z} = \frac{z_2}{z_1}.$$

Крім того, $z_1 + z_2 = Z$. Із першого рівняння витікає $z_1^2 = Z \cdot z_2$. Підставивши значення Z з другого рівняння і поділивши обидві частини на z_1^2 , отримаємо

$$\left(\frac{z_2}{z_1}\right)^2 + \frac{z_2}{z_1} - 1 = 0.$$

Розв'язуючи це квадратне рівняння, знаходимо для додатнього кореня значення

$$\frac{z_2}{z_1} = \frac{-1 + \sqrt{5}}{2} = 0,618.$$

На рис. 13.12 показано ділення інтервалу невизначеності в цьому відношенні і нанесені відповідні значення цільової функції, які дозволяють зменшити інтервал невизначеності в $1/0,618$ раз.

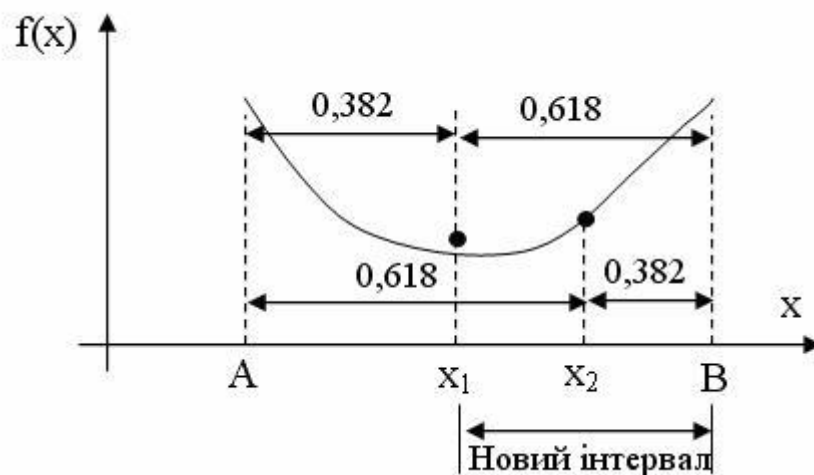


Рисунок 13.12 – Метод золотого перетину

На цій стадії ще не видно переваг методу золотого перетину в порівнянні з методом дихотомії, однак їх добре видно при подальшому діленні інтервалу, так як виявляється, що одне із значень цільової функції, яке необхідно обчислити на наступному кроці, вже відомо. Тому, щоб зменшити невизначеність ще в $1/0,618$ раз, потрібно додатково обчислити тільки одне значення цільової функції в точці, яка визначається правилом золотого перетину.

При $n > 2$ ефективність методу золотого перетину вища, ніж у метода дихотомії, так як при кожному наступному обчисленні цільової функції інтервал невизначеності скорочується в $1/0,618$ раз. Після обчислення N значень цільової функції коефіцієнт дроблення інтервалу невизначеності складає

$$f = 0,618 N - 1.$$

Метод золотого перетину дозволяє відмітити цікаву закономірність: найбільше скорочення наступних інтервалів невизначеності досягається при обчисленні цільової функції в точках, рівновіддалених від його центру. Якщо продовжувати таким чином і кожного разу, обчислюючи цільову функцію, скорочувати інтервал невизначеності, то будуть справедливими наступні відношення:

$$Z_{j-2} = Z_{j-1} + Z_j, \quad 1 \leq i \leq n, \quad 1 \leq j \leq n-1$$

де Z_j довжина інтервалу невизначеності після обчислення J -го значення цільової функції.

На рис. 13.13 представлений алгоритм вибору наступної точки пошуку. Задана точність може, звичайно, змінюватися вибором значення. Для функції $f(x) = -e^{-x} \ln(x)$, пошук відбувався в інтервалі $(0, 2)$.

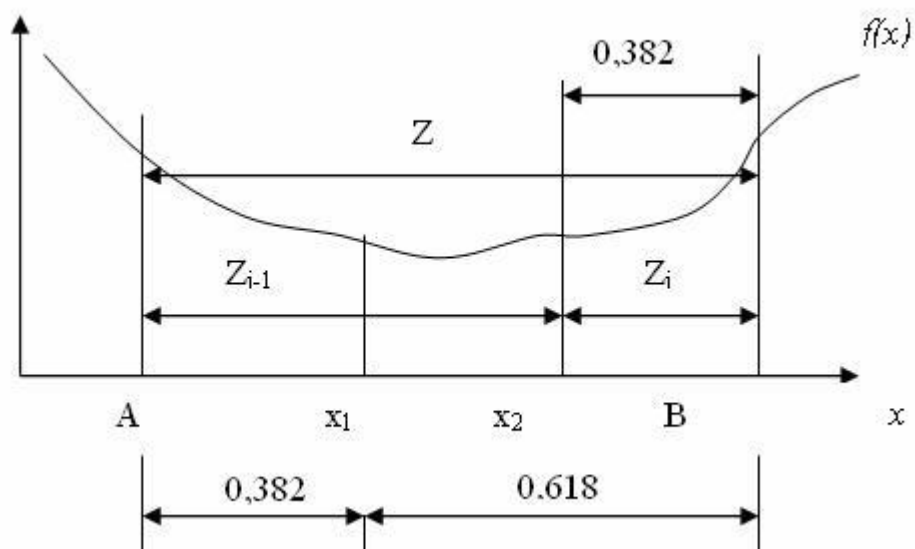


Рисунок 13.13 - Послідовність етапів вибору наступної точки пошуку

Істинний мінімум знаходиться в точці $1,76322211$, де значення функції дорівнює $-0,0972601313$.

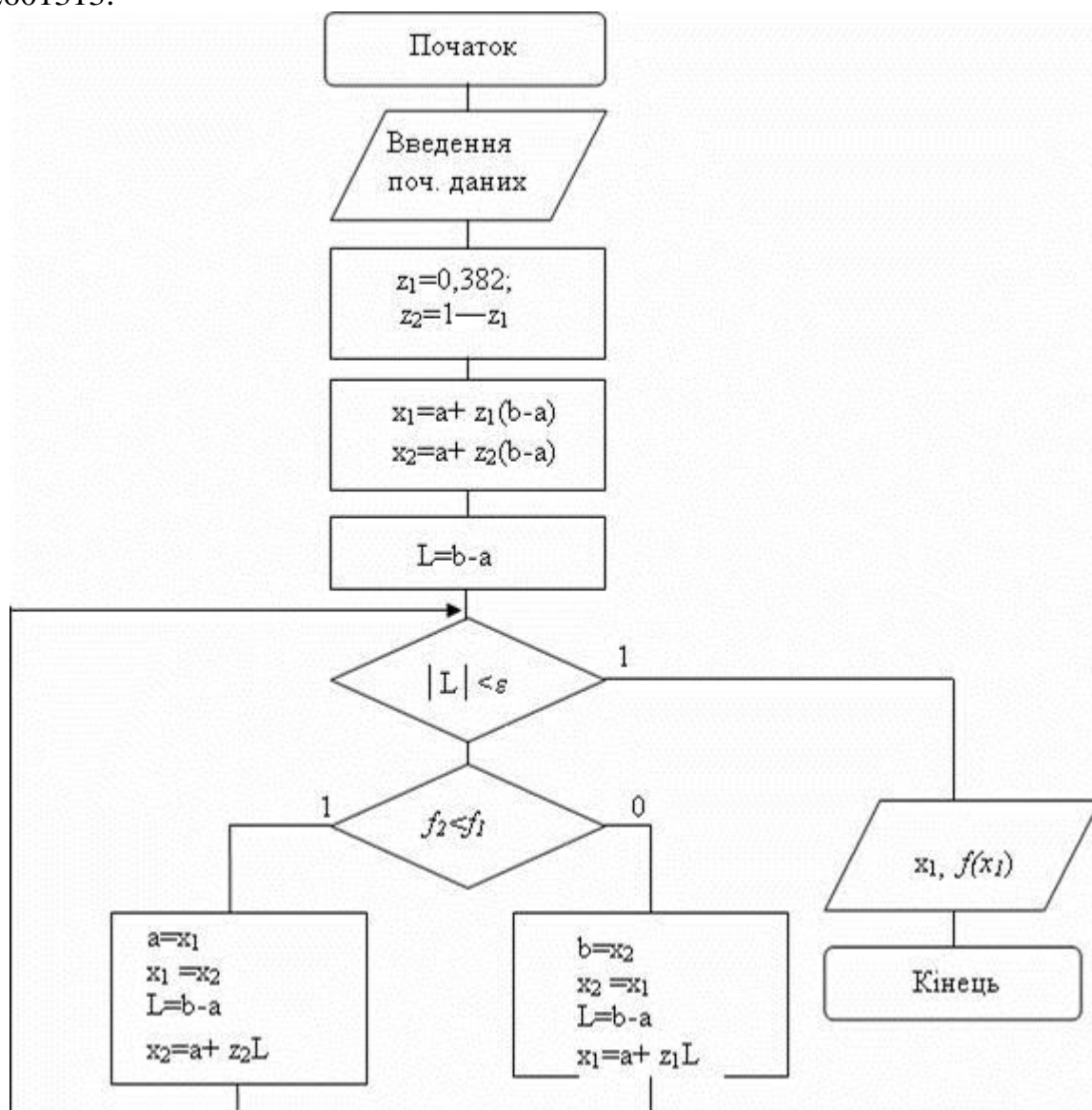


Рисунок 13.14 - Схема алгоритму метода „золотого перетину”

При розробці програм для рішення задач однопараметричної оптимізації використовують наступні співвідношення:

$$z_1 = 0.382$$

$$z_2 = 1 - z_1 = 0.618$$

$$x_1 = A + z_1(b - a)$$

$$x_2 = A + z_2(b - a)$$

13.5 Метод Фібоначчі

Припустимо, що потрібно визначити мінімум цільової функції як можна точніше, тобто з найменшим можливим інтервалом невизначеності, але при цьому можна виконати тільки n обчислень функції. Як слід вибрати n точок, в яких обчислюється функція? З першого погляду здається ясным, що не слід шукати рішення для всіх точок, які одержані в результаті експерименту. Навпаки, треба спробувати зробити так, щоб значення функції, отримані в попередніх експериментах, визначали положення наступних точок. Справді, знаючи значення функції, ми тим самим маємо інформацію про саму функцію і положення її мінімуму і використаємо цю інформацію в подальшому пошуку.

Припустимо, що є інтервал невизначеності (x_1, x_3) і відомо значення функції $f(x_2)$ всередині цього інтервалу (див. рис. 13.15).

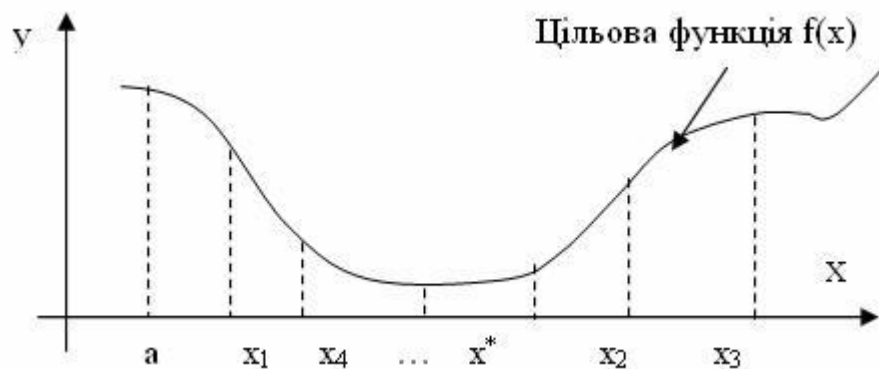


Рисунок 13.15 – Унімодальна цільова функція. Геометрична інтерпретація метода Фібоначчі

Якщо можна обчислити функцію всього один раз в точці x_4 , то де слід помістити точку x_4 , для того щоб отримати найменший можливий інтервал невизначеності?

Припустимо $x_2 - x_1 = L$ і $x_3 - x_2 = R$, причому $L > R$, як показано на рис. 13.15, і ці значення будуть фіксовані, якщо відомі x_1, x_2 і x_3 . Якщо x_4 знаходиться в інтервалі (x_1, x_2) , то:

1. Якщо $f(x_4) < f(x_2)$, то новим інтервалом невизначеності буде (x_1, x_2) довжиною $x_2 - x_1 = L$.

2. Якщо $f(x_4) > f(x_2)$, то новим інтервалом невизначеності буде (x_4, x_3) довжиною $x_3 - x_4$.

Оскільки невідомо, яка з цих ситуацій буде мати місце, виберемо x_4 таким чином, щоб зробити мінімальною найбільшу з довжин $x_3 - x_4$ і $x_2 - x_1$. Досягнути цього можна, зробивши довжини $x_3 - x_4$ і $x_2 - x_1$ рівними, тобто помістивши x_4 всередині інтервалу симетрично відносно точки x_2 , що вже лежить всередині інтервалу. Будь-яке інше положення точки x_4 може призвести до того, що отриманий інтервал буде більший L . Помістивши x_4 симетрично відносно x_2 , ми нічим не ризикуємо в будь-якому випадку.

Якщо виявиться, що можна виконати ще одне обчислення функції, то слід застосувати описану процедуру до інтервалу (x_1, x_2) , в якому вже є значення функції, обчислене в точці x_2 . Отже, стратегія зрозуміла з самого початку. Потрібно помістити наступну точку всередині інтервалу невизначеності симетрично відносно точки, яка вже знаходиться там. Парадоксально, але, щоб зрозуміти, як потрібно починати обчислення, необхідно розібратися в тому, як його потрібно закінчувати.

На n -ому обчисленні n -у точку потрібно помістити симетрично по відношенню до $(n-1)$ -ї точки. Положення цієї останньої точки в принципі залежить від нас. Для того щоб отримати найбільше зменшення інтервалу на даному етапі, слід поділити навпіл попередній інтервал. Тоді точка x_n буде співпадати з точкою x_{n-1} . Однак при цьому ми не одержуємо жодної нової інформації. Звичайно точки x_{n-1} і x_n знаходяться одна від одної на достатній відстані, щоб визначити, в якій половині, лівій чи правій, знаходиться інтервал невизначеності. Вони розміщуються на відстані $\varepsilon/2$ по обидві сторони від середини відрізка L_{n-1} ; можна самостійно задати величину ε або вибрати цю величину рівній мінімально можливій відстані між двома точками. (Припустимо, що в нашому прикладі інженер може регулювати температуру з інтервалом в 1° , тому $\varepsilon = 1$.)

Інтервал невизначеності буде мати довжину L_n , отже, $L_{n-1} = 2L_n - \varepsilon$ (рис. 13.16, нижня частина).

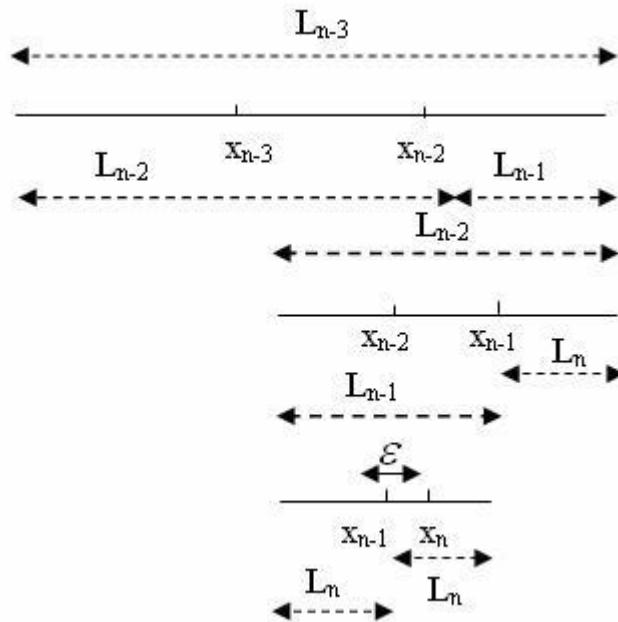


Рисунок 13.16 – Геометрична інтерпретація ітераційного процесу Фібоначчі

На попередньому етапі точки x_{n-1} і x_{n-2} повинні бути поміщені симетрично всередині інтервалу з L_{n-2} на відстані L_{n-1} від кінців цього інтервалу. Отже, $L_{n-1} = L_{n-2} + L_{n-1}$ (рис. 13.16, середня частина).

Зауваження. З рисунку зрозуміло, що на передостанньому етапі x_{n-2} залишається в якості внутрішньої точки.

Аналогічно

$$L_{n-2} = L_{n-3} + L_{n-2}. \text{ (рис. 13.16, верхня частина)}$$

В загальному випадку

$$L_{j-1} = L_j + L_{j+1} \quad \text{при } 1 < j \leq n-1$$

Таким чином,

$$\begin{aligned}
L_{n-1} &= 2L_n - \varepsilon, & (13.2) \\
L_{n-2} &= L_{n-1} + L_n = 3L_n - \varepsilon, \\
L_{n-3} &= L_{n-2} + L_{n-1} = 5L_n - 2\varepsilon, \\
L_{n-4} &= L_{n-3} + L_{n-2} = 8L_n - 3\varepsilon \text{ і т.д.}
\end{aligned}$$

Якщо визначити послідовність чисел Фібоначчі наступним чином:

$F_0 = 1, F_1 = 1$, та $F_k = F_{k-1} + F_{k-2}$ для $k = 2, 3, \dots$, тоді

$$L_{n-j} = F_{j+1} L_n - F_j \varepsilon, \quad j = 1, 2, \dots, n-1 \quad (13.3)$$

Якщо початковий інтервал (a, b) має довжину $L_1 = (b - a)$, то

$$L_1 = F_n L_n - \varepsilon \cdot F_{n-1}.$$

Тобто

$$L_n = \frac{L_1}{F_n} + \varepsilon \frac{F_{n-1}}{F_n}. \quad (13.4)$$

Отже, зробивши n обчислень функції, ми зменшимо початковий інтервал невизначеності в $1/F_n$ раз у порівнянні з його початковою довжиною (нехтуючи ε), і це буде найкращий результат.

Якщо ми вже почали пошук, то його нескладно продовжити, використовуючи описане вище правило симетрії. Отже, необхідно знайти положення першої точки, що розміщується на відстані L_2 від одного з кінців початкового інтервалу, причому не важливо, від якого кінця, оскільки друга точка розміщується згідно правила симетрії на відстані L_2 від другого кінця інтервалу:

$$L_2 = F_{n-1} L_n - \varepsilon F_{n-3} = F_{n-1} \frac{L_1}{F_n} + \varepsilon \frac{(F_{n-1} F_{n-2} - F_n F_{n-3})}{F_n} = \frac{F_{n-1}}{F_n} \cdot L_1 + \frac{(-1)^n \varepsilon}{F_n}. \quad (13.5)$$

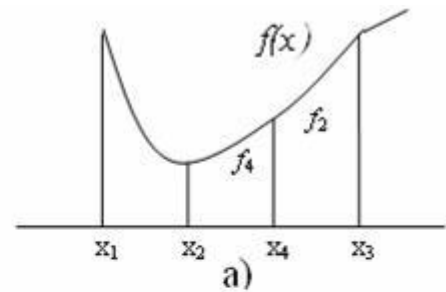
Після того як знайдене положення першої точки, числа Фібоначчі більше не потрібні. Використане значення ε може бути визначене з практичних міркувань. Воно повинне бути менше L_1 / F_{n+1} , в протилежному випадку ми будемо марно витратити час на обчислення функції.

Таким чином, пошук методом Фібоначчі, названий так через появу при пошуку чисел Фібоначчі, є ітераційною процедурою. В процесі пошуку інтервалу (x_1, x_2) з точкою x_2 , що вже лежить в цьому інтервалі, наступна точка x_4 завжди вибирається такою, що $x_3 - x_4 = x_2 - x_1$ або $x_4 - x_1 = x_3 - x_2$, тобто

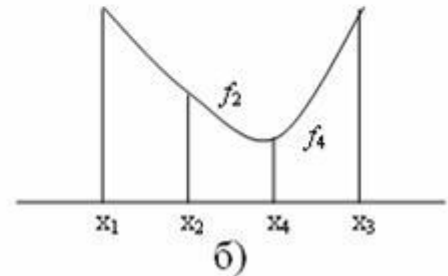
$$x_4 = x_1 - x_2 + x_3. \quad (13.6)$$

Якщо $f(x_2) = f_2$ та $f(x_4) = f_4$, тоді можна розглянути чотири випадки (рис. 13.17).

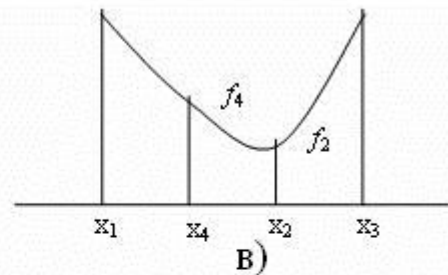
а) $x_4 < x_2$
 $f_4 < f_2$
 Новий інтервал (x_1, x_2) ,
 що містить точку x_4



б) $x_4 > x_2$
 $f_4 < f_2$
 Новий інтервал (x_2, x_3) ,
 що містить точку x_4



в) $x_4 < x_2$
 $f_4 > f_2$
 Новий інтервал (x_4, x_3)
 що містить точку x_2



г) $x_4 > x_2$
 $f_4 > f_2$
 Новий інтервал (x_1, x_4)
 що містить точку x_2

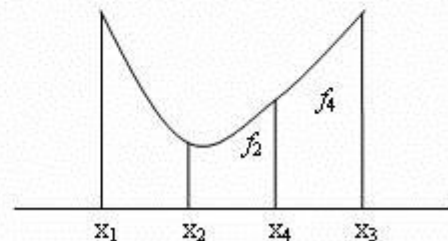


Рисунок 13.17 – Геометрична інтерпретація алгоритму визначення цільової функції на i -му кроці ітерації

Приклад 1. Використати метод Фібоначчі для пошуку мінімуму функції $f(x) = 2x^2 - e^x$ в інтервалі $(0,1)$ при 10-кратному обчисленні функції.

Розв'язок. *Остаточний інтервал невизначеності має довжину*

$$0,359550563 - 0,348314606 = 0,011235957 \approx \frac{1}{89} = \frac{1}{F_{10}}$$

З точністю до шостого знаку після коми мінімум досягається в точці $x = 0.357403$, і в цій точці $f(x^*) = -1.174138$.

13.6 Порівняння методів одновимірного пошуку

Найкращими критеріями порівняння методів пошуку, які були описані вище, є їх ефективність і універсальність. Під ефективністю алгоритму розуміють число обчислень функції, необхідне для досягнення необхідного звуження інтервалу невизначеності. Із табл. 13.1 видно, що найкращим в цьому відношенні є метод Фібоначчі, а найгіршим – метод загального пошуку. Конструктор не з великим задоволенням використовує метод Фібоначчі, так як при його застосуванні необхідно заздалегідь задавати число обчислень значень функції. Однак він може скористатися

методом золотого перетину. Як правило, методи Фібоначчі і золотого перетину, володіють високою ефективністю, найбільш підходять для розв'язку одновимірних унімодальних задач оптимізації.

Універсальність алгоритму означає, що його можна легко застосувати для розв'язку самих різноманітних задач. В цьому відношенні метод Фібоначчі, поступається іншим, так як потребує окремого обчислення положення точок, в яких будуть визначатися значення цільової функції на кожному новому кроці. Цим приходиться розплачуватися за підвищення ефективності метода. З точки зору універсальності малоефективний метод загального пошуку має по крайній мірі одну перевагу – його можна з успіхом застосовувати і для неунімодальних функцій, якщо вони достатньо гладкі. Нерідко заздалегідь не відомо, чи є розглянута цільова функція унімодальною. В таких випадках слід використати декілька різних алгоритмів і подивитись, чи дають вони усі один і той самий оптимум. Звідси витікає важливий висновок, який слід мати на увазі, розв'язуючи задачі оптимізації: не існує універсального алгоритму, який дозволяв би розв'язувати будь-які задачі. Вирішуючи складні задачі оптимізації, слід користуватися різними методами, так як це дозволяє збільшити долю вигідних розв'язків.

МЕТОДИ ОПТИМІЗАЦІЇ БАГАТОВИМІРНИХ ЗАДАЧ

14.1 Основні поняття та визначення

У цій частині розглядаються фундаментальні поняття і конкретні методи, які використовують при пошуку безумовних мінімумів функцій декількох змінних. Викладене ґрунтується на матеріалі розділу 3, оскільки одномірні методи відіграють дуже важливу роль при дослідженні функції декількох змінних.

На перший погляд може показатися, що відмінність між методами багатовимірного і одномірного пошуку полягає лише в тому, що перші вимагають більшого об'єму обчислень, і, що в принципі методи, які придатні для функцій однієї змінної, можна застосовувати і для функцій багатьох змінних. Однак це не так, оскільки багатовимірний простір якісно відрізняється від одномірного. Передусім зі збільшенням числа вимірів зменшується ймовірність унімодальності цільової функції. Крім того, безліч елементів, які утворюють багатовимірний простір, значно потужніші множини елементів одномірного простору. Об'єм обчислень, які необхідні для звуження інтервалу невизначеності в багатовимірному просторі, є степеневою функцією, показник якої рівний розмірності простору. Так, якщо в випадку одномірного простору для досягнення $f^* = 0,1$ вимагається обчислити 19 значень цільової функції, то у випадку двомірного простору це число складає 361, тримірного – 6859, чотиримірного – 130321, а п'ятимірного – 2476099! Оскільки при виборі оптимальної конструкції нерідко потрібно мати діло з п'ятьма і більше змінними, серйозність труднощів, зумовлених багатовимірністю, стає очевидною.

Спочатку розглянемо *питання аналізу* (в статистичній літературі з використанням положення лінійної алгебри і диференційного обчислення, а також умови, які (в достатньо загальних можливих випадках) дозволяють ідентифікувати точки оптимуму. Такі умови використовують для перевірки обраних точок і дають можливість з'ясувати, чи є ці точки точками мінімуму чи сідловими точками. При цьому задача вибору вказаних точок залишається за межами цього аналізу; основна увага віддається розв'язанню питання про те, чи відповідають досліджувані точки розв'язку багатовимірної задачі безумовної оптимізації, у якій вимагається мінімізувати $f(x)$, $x \in \mathbb{R}^N$, (14.1) коли обмеження відсутні на x , де x - вектор *керованих змінних* розмірності N , f — скалярна *цільова функція*. Звичайно допускається, що x_i (для всіх значень $i = 1, 2, 3, \dots, N$) можуть приймати будь-які значення, хоча інколи в практичних застосуваннях область значень x обирається в вигляді дискретної множини. Крім того, часто зручно допускати, що функція f і її похідні існують і неперервні скрізь, хоча ми знаємо, що оптимуми можуть досягати в точках розриву f чи її *градієнта*

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \dots, \frac{\partial f}{\partial x_N} \right]^T. \quad (14.1)$$

Потрібно мати на увазі, що функція f може приймати мінімальні значення в точці \bar{x} , в якій f чи ∇f розриваються. Крім того, цієї точки \bar{x} може не існувати. Для того, щоб побудувати систему конструктивних критеріїв оптимальності, необхідно (якнайменше на першій стадії дослідження) виключити із розгляду подібні ситуації, які дуже ускладнюють аналіз. У ряді випадків приходиться обмежуватись лише

ідентифікацією *локальних* оптимумів, оскільки нелінійна цільова функція f не завжди має випуклий характер, може бути мультимодальною. На рис. 14.1 відображенні лінії рівняння функції Хіммельблау

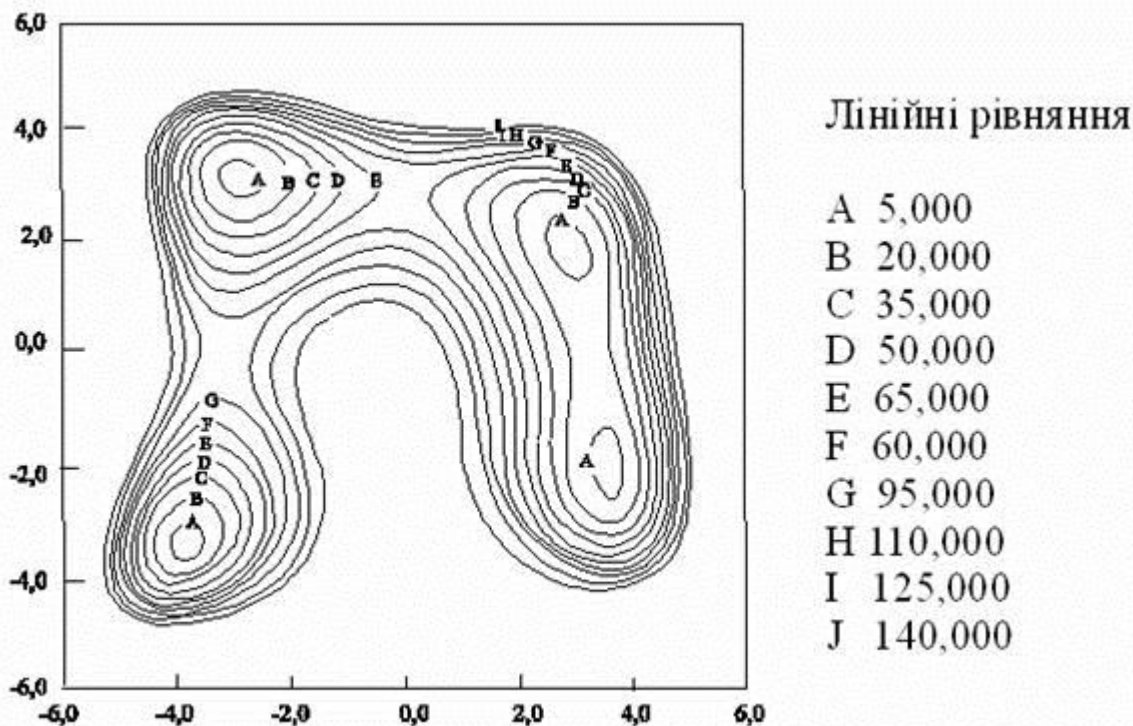


Рисунок 14.1 - Лінії рівняння мультимодальної функції

$$f(x)=[x_1^2+x_2^2-11]^2+[x_1+x_2^2-7]^2, \quad (14.2)$$

неважко побачити, що ця функція має чотири різних мінімуму.

Далі підійдемо до *питання аналізу* (в динаміці), яке формулюється таким чином: якщо точка $x^{(0)}$ не задовольняє вимогам, які

Функція Хіммельблау:

$$f(x)=[x_1^2+x_2^2-11]^2+[x_1+x_2^2-7]^2 \quad (14.3)$$

визначаються критеріями оптимальності, то як отримати (добре) нове наближення $x^{(1)}$ до розв'язку x ? Спроба дати відповідь на це питання приводить до необхідності розгляду ряду методів. Методи, які розглядаються, класифікуються у відповідності з тим, чи використовується інформація про похідні досліджуваної функції.

14.2 Критерії оптимальності

Розглянемо умови, які дозволяють характеризувати (тобто класифікувати) точки простору керованих змінних. Критерії оптимальності необхідні для розпізнавання рішень і, крім того, складають основу більшості методів пошуку рішень, що використовуються.

Для цього розглянемо розклад Тейлора для функції декількох змінних :

$$f(x) = f(\bar{X}) - \nabla f(\bar{X}) \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(\bar{X}) \Delta x + O_3(\Delta x). \quad (14.4)$$

де \bar{X} - точка розкладення із простору R^N ; $\Delta x = x - \bar{X}$ — величина зміни x ; $\Delta f(x)$ — N -мірний вектор - стовпчик перших похідних $f(x)$, обчислених в точці \bar{X} ; $\nabla^2 f(\bar{X}) = H_f(\bar{X})$ — симетрична матриця порядку $N \times N$ других частинних похідних $f(x)$, вирахованих у точці \bar{X} . (Цю матрицю часто називають *матрицею Гессе*. Її елемент, розташований на перетині i -й строчки і j -го стовпчика, дорівнює $\partial^2 f / \partial x_i \partial x_j$.) $O_3(\Delta x)$ - сума всіх членів розкладу, які мають порядок по Δx вище другого. Нехтуючи членами найвищих

порядків (тобто виключаючи $O_3(\Delta x)$), визначимо величину зміни цільової функції $f(x)$, відповідно довільній зміні x :

$$f(x) = f(\bar{x}) - f(\bar{x}) = \Delta f(\bar{x})^T \Delta x + \frac{1}{2} \Delta x^T \nabla^2 f(\bar{x}) \Delta x. \quad (14.5)$$

Нагадаємо, що по визначенню у всіх точках із околу точки мінімуму цільова функція приймає значення, які вище мінімального, тобто має місце нерівність

$$\Delta f = f(x) - f(\bar{x}) \geq 0. \quad (14.6)$$

Точка \bar{x} є точкою *глобального* мінімуму, якщо нерівність (14.6) виконується для всіх $x \in \mathbb{R}^N$; такі точки будемо позначати через x^{**} . Коли формула (14.6) справедлива лише в деякому δ -околі точки \bar{x} , т.б. для всіх x , таких, що $|x - \bar{x}| < \delta$ при заданому $\delta > 0$, то \bar{x} є точкою *локального* мінімуму, чи x^* . Якщо

$$\Delta f = f(x) - f(\bar{x}) \leq 0, \quad (14.7)$$

то \bar{x} є точкою максимуму (локального чи глобального у відповідності з даним вище визначенням). Виключення знака рівності із формул (14.6) і (14.7) дозволяє визначити точку *строгого* мінімуму чи максимуму. В тому випадку, коли Δf приймає як додатне, так і від'ємне, так і нульове значення в залежності від вибору точок із δ -околу, точка \bar{x} являє собою *сідлову точку*.

Повернемося до рівняння (14.5) і згадаємо про висунуте раніше припущення про те, що $f(x)$, $\Delta f(x)$ і $\Delta^2 f(x)$ існують і неперервні для всіх $x \in \mathbb{R}^N$. Як впливає з формули (14.5), для того щоб знак Δf не змінювався при довільній зміні Δx , градієнт $\nabla f(\bar{x})$ повинен дорівнювати нулю, тобто \bar{x} повинна бути *стаціонарною точкою*. В протилежному випадку різниця Δf може приймати додатні чи від'ємні значення в залежності від знаків $\nabla f(\bar{x})$ і Δx . Таким чином, точка \bar{x} повинна задовольняти *умові стаціонарності*

$$\nabla f(\bar{x}) = 0, \quad (14.8)$$

і формула (14.5) приймає такий вигляд:

$$\Delta f(x) = \frac{1}{2} \Delta x^T \nabla^2 f(\bar{x}) \Delta x. \quad (14.9)$$

Звідси видно, що знак $\Delta f(x)$ визначається квадратичною формою

$$Q(x) = \Delta x^T \nabla^2 f(\bar{x}) \Delta x, \quad (14.10)$$

чи $Q(x) = z^T A z$. (14.11)

Із лінійної алгебри ми знаємо, що:

- A-** додатньо визначена матриця, якщо $Q(z) > 0$ для любых z ;
- A-** додатньо напіввизначена матриця, якщо $Q(z) \geq 0$ для любых z ;
- A-** від'ємно визначена матриця, якщо $Q(z) < 0$ для любых z ;
- A-** від'ємно напіввизначена матриця, якщо $Q(z) \leq 0$ для любых z ;
- A-** невизначена матриця, якщо $Q(z) > 0$ для деяких z і $Q(z) < 0$ для останніх z .

Із (14.11) видно, що стаціонарна точка \bar{x} є

- точкою мінімуму, якщо $\nabla^2 f(\bar{x})$ — додатньо напіввизначена матриця;
- точкою максимуму, якщо $\nabla^2 f(\bar{x})$ — від'ємно напіввизначена матриця;
- сідлова точка, якщо $\nabla^2 f(\bar{x}) >$ (або $<$) 0 — невизначена матриця.

Крім того, іноді корисно провести аналіз стаціонарної точки \bar{x} в деякому іншому аспекті. Розглянемо стаціонарну точку \bar{x} разом з оточуючим її δ -колом і векторами, що виходять із точки \bar{x} (рис. 14.2). При цьому

$$\tilde{x} = \bar{x} + \alpha s(\bar{x}). \quad (14.12)$$

Шляхом відповідного вибору α і s можливо побудувати всі точки \bar{x} із кола точки \bar{x} . Підстановка (14.12) в (14.9) дає формулу

$$\Delta f(x) = (\alpha^2 / 2) s^T \nabla^2 f(\bar{x}) s. \quad (14.13)$$

Тепер за допомогою (14.11) і (14.12) можливо класифікувати $s(x)$ як напрямок спуску, напрямок підйому чи напрямок загального вигляду. Якщо напрямок спуску знайти *не вдається*, то \bar{x} є точкою локального мінімуму x^* , що відповідає випадку, коли $\nabla^2 f(\bar{x})$ - додатньо напіввизначена матриця.

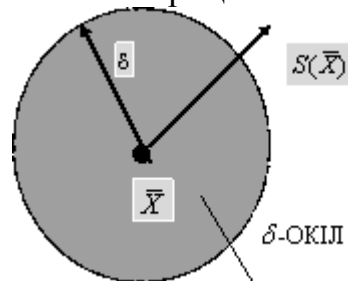


Рисунок 14.2 - Окіл стаціонарної точки

Теорема 1. Необхідні умови

Для наявності в точці x^* локального мінімуму необхідно, щоб виконувалась рівність

$$\nabla f(x^*) = 0 \quad (14.14a)$$

і матриця $\nabla^2 f(x^*)$ була додатньо напіввизначеною:

$$\nabla^2 f(x^*) \geq 0. \quad (14.14b)$$

Теорема 2. Достатні умови

Якщо

$$\nabla f(x^*) = 0 \quad (14.15a)$$

і матриця $\nabla^2 f(x^*)$ додатньо визначена, то

$$\nabla^2 f(x^*) > 0. \quad (14.15b)$$

Доведення цих теорем витікає із наведеного вище судження. Звичайно доводиться обмежуватись знаходженням локального мінімуму; разом з тим, якщо можна показати, що $x^T \nabla^2 f(x) x \geq 0$ для всіх x , то $f(x)$ називають *випуклою* функцією, а локальний мінімум називають глобальним.

Приклад 14.1. Критерії оптимальності

Розглянемо функцію

$$f(x) = 1x_1^2 + 4x_1x_2^3 - 10x_1x_2 + x_2^2,$$

лінії рівня якої зображені на рис. 14.3.

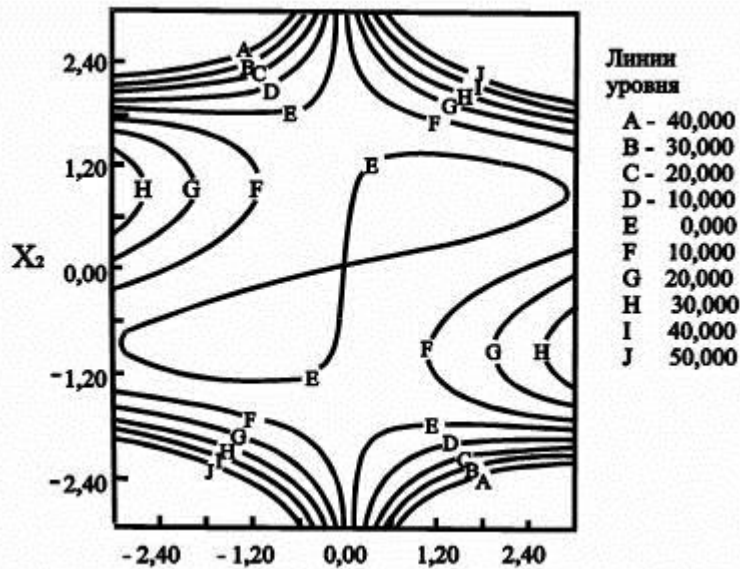


Рисунок 14.3 - Лінії рівня нелінійної функції двох змінних з прикладу 14.1

Треба класифікувати точку $\bar{x} = [0,0]^T$.

$$f(x) = 2(x_1)^2 + 4x_1x_2^3 - 10x_1x_2 + x_2^2$$

Розв'язок.

$$\frac{\partial f}{\partial x_1} = 4x_1 + 4x_2^3 - 10x_2,$$

$$\frac{\partial f}{\partial x_2} = 12x_1x_2^2 - 10x_1 + 2x_2,$$

$$\nabla f(\bar{x}) = [0,0]^T.$$

Звідси, точка \bar{x} — стаціонарна.

$$\frac{\partial^2 f}{\partial x_1^2} = 4,$$

$$\frac{\partial^2 f}{\partial x_2^2} = 24x_1x_2 + 2 = +2,$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 12x_2^2 - 10 = -10.$$

Звідси,

$$\nabla^2 f(\bar{x}) = H_f(\bar{x}) = \begin{bmatrix} +4 & -10 \\ -10 & +2 \end{bmatrix}.$$

Матриця $\nabla^2 f(\bar{x})$ є невизначеною, так як квадратична форма $z^T H_f z$ приймає додатне значення при $z=(0,1)$ і від'ємне значення при $z=(1,1)$. Тому \bar{x} представляє собою *сідлову точку*, що і відображено на рис. 14.3.

14.3 Градієнтні методи

В багатьох алгоритмах багатовимірної оптимізації так чи інакше використовується інформація про градієнти. Проілюструємо це положення наступним простим прикладом. Представимо собі, що альпіністу зав'язали очі і сказали, що він повинен добратися до вершини “унімодальної” гори. Навіть якщо він не буде нічого бачити, він може це зробити, якщо весь час буде рухатися вгору. Хоча будь-яка тропа, яка веде вгору, в кінці-кінців приведе його до вершини, найкоротшою з них буде найкрутіша, якщо, правда, альпініст не нашттовхнеться на вертикальний обрив, який необхідно буде обійти. (Математичним еквівалентом обриву на поверхні, яку створює

цільова функція, являються ті її місця, де поставлені умовні обмеження). Уявімо, що задача оптимізації не містить обмежень. Пізніше ми включимо їх в схему пошуку. Метод оптимізації, в основі якого лежить ідея руху по самій крутій тропі, називається *методом найшвидшого підйому або найшвидшого спуску*. Вектор градієнта перпендикулярний лінії рівня і вказує напрямком до нової точки в просторі проектування. Відмітимо, що градієнтний метод на відміну від методу дотичної до лінії рівня можна використовувати до будь-якої унімодальної функції, а не тільки тих, у яких ця властивість явно виражена.

Щоб краще зрозуміти ідею градієнтних методів, більш конкретно зупинимося на властивостях градієнтів. Розглянемо систему незалежних одиничних векторів $e_1, e_2, e_3, \dots, e_N$, які направлені в здовж вісей координат $x_1, x_2, x_3, \dots, x_N$, які є в той же час проектними параметрами. Вектор градієнта довільної цільової функції $F = (x_1, x_2, x_3, \dots, x_N)$ має вигляд

$$\frac{\partial F}{\partial x_1} e_1 + \frac{\partial F}{\partial x_2} e_2 + \dots + \frac{\partial F}{\partial x_N} e_N,$$

де частинні похідні обчислюються в точці, яка розглядається. Цей вектор направлений вгору, в напрямку підйому; обернений йому вектор вказує напрямком спуска. Одиничний вектор градієнта часто представляють у вигляді

$$s_1 e_1 + s_2 e_2 + s_3 e_3 + \dots + s_N e_N,$$

$$s_i = \frac{\frac{\partial F}{\partial x_i}}{\sum_{j=1}^N \left[\left(\frac{\partial F}{\partial x_j} \right)^2 \right]^{1/2}}$$

де (14.16)

Іноді характер цільової функції буває достатньо добре відомий, щоб можна було обчислити компоненти вектора градієнта шляхом безпосереднього диференціювання. Якщо таким способом частинні похідні отримати неможливо, то можна знайти їх наближені значення в безпосередньому околі точки, яка розглядається:

$$\frac{\partial F}{\partial x_i} = \frac{F(x_1, x_2, \dots, x_i + \Delta, \dots, x_N) - F(x_1, x_2, \dots, x_i, \dots, x_N)}{\Delta}$$

Тут Δ - невелике зміщення в напрямку x_i . Цю формулу часто називають "наближенням січної". Отриману інформацію про напрямком градієнта можна використовувати різним чином для побудови алгоритму пошуку.

Постановка задачі оптимізації градієнтними методами: мінімізація функції $F(x_1, x_2, x_3, \dots, x_N)$ з N проектними параметрами за допомогою ЕОМ розв'язується ітераційними методами. Розв'язок задачі починається з вибору початкових значень $x_i^{[0]}$ ($i = 1, 2, \dots, N$), які за звичай визначаються із умов розв'язуваної задачі, і потім будують послідовні наближення, використовуючи ітераційну формулу, яку отримуємо з формули (14.12):

$$x_i^{[j+1]} = x_i^{[j]} + \lambda^{[j]} s_i^{[j]}, \quad (i = 1, 2, \dots, N; j = 0, 1, 2, \dots), \quad (14.17)$$

де $\lambda^{[j]}$ - величина кроку ітерації по кожному з параметрів x_i ;

$s_i^{[j]}$ - параметр вибору "напрямок", який звичайно визначається за формулою (14.12).

Дана формула забезпечує збіжність досліджуваної функції до деякого рішення \bar{x}_k при $j \rightarrow \infty$. Величина кроку $\lambda^{[j]}$ на кожній j -ій ітерації визначається одним

з методів оптимізації однопараметричної оптимізації, наприклад методом ділення відрізка навпіл або методом “золотого перетину” або Фібоначчі.

14.3.1 Найшвидший підйом з використанням одномірного пошуку

В деяких методах пошуку інформація про градієнт використовується для ведення одномірного пошуку в напрямку найшвидшого підйому або спуску, причому використовується співвідношення

$$x_{i,нов} = x_{i,ст} + \lambda s_i, \quad (14.18)$$

де λ - величина кроку, значення якого визначаються в напрямку градієнта.

Отримавши одномірний оптимум в напрямку даного градієнта, знаходять новий градієнт і повторюють процес до тих пір, поки наступні обчислення дозволяють покращувати отриманий результат. Головна перевага цього метода полягає в тому, що параметр λ можна використовувати в якості незалежної змінної для пошуку за методом Фібоначчі, і це забезпечує високу ефективність метода. Інша важлива перевага методів, які розглядаються, полягає в тому, що вони дозволяють відходити від сідловин точок поверхні, яка описується цільовою функцією (рис. 14.4).

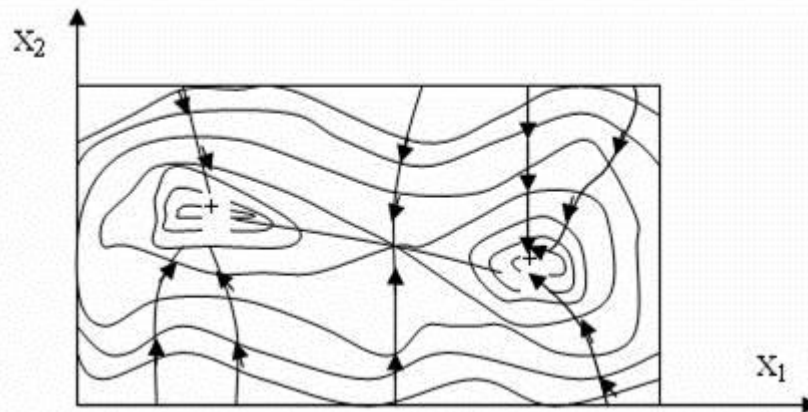


Рисунок 14.4 - Бімодальна цільова функція

Відмітимо, однак, що, як видно з рисунку, для мультимодальних функцій градієнтні методи дозволяють знайти лише локальний оптимум. Тому, якщо характер поверхні недостатньо добре відомий, то слід випробувати декілька початкових точок і переконатися, що в усіх випадках отримується одне й те ж оптимальне рішення. Іншою причиною, яка знижує ефективність градієнтних методів, являються злами ліній рівня цільової функції. Так як такі точки відповідають розриву в нахилі лінії контура, то тут можливі помилки в визначенні напрямку подальшого пошуку. Тому пошук може уповільнитися і йти зигзагами поперек лінії зламу, а час необхідний для отримання рішення, буде настільки великим, що рахунок прийдеться припинити. В дійсності більшість досліджуваних поверхонь має одну або більше ліній зламу, які нерідко проходять через точку оптимуму. Тому, наткнувшись на лінію зламу, слід в подальшому рухатися вздовж неї.

14.3.2 Метод найшвидшого спуску

Даний метод заснований на використанні ітераційної формули

$$x_i^{[j+1]} = x_i^{[j]} + \lambda^{[j]} s_i^{[j]},$$

$$s_i^{[j]} = -\frac{\partial F}{\partial x_i}$$

де $\lambda_i = x_i^{[1]}$;

$\lambda^{[j]}$ - величина кроку, значення якого змінюється (зменшується або обчислюється) методом половинного ділення.

Алгоритм методу найшвидшого спуску:

$$\bar{x}^0 = \left(x_1^{(0)}, x_2^{(0)}, \dots, x_N^{(0)} \right)$$

1. Обираємо початкові значення координат вектора \bar{x}^0 і початкові значення кроку ітераційного процесу λ , які звичайно обираються з умов розв'язуємої конкретної задачі. Хоча загальних правил вибору \bar{x}^0 немає, однак якщо є додаткова інформація про область розташування мінімуму цільової функції, то \bar{x}^0 обираємо в цій області.

2. Задаємо номер ітерації $k = 1$.

3. Обчислюємо значення цільової функції в точці з координатами \bar{x}^0 .

4. Обчислюємо значення градієнта S_i .

5. Обчислюємо норму вектора градієнта NG .

6. Якщо $|NG| < \varepsilon$, то ітераційний процес закінчується і оптимум знайдений.

7. Якщо умова $|NG| < \varepsilon$ не виконується, то визначаються нові координати вектора \bar{x}^1 , які отримуються при русі до мінімуму цільової функції з кроком λ (рис. 14.5).

8. Порівнюємо два значення цільової функції в двох точках з координатами векторів \bar{x}^0 і \bar{x}^1 за формулою

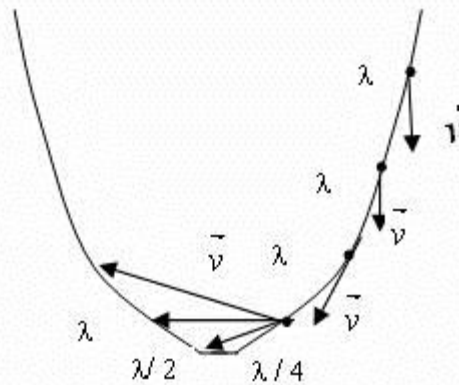


Рисунок 14.5– Послідовність руху до мінімуму з заданим кроком λ

$$f(\bar{x}^1) < f(\bar{x}^0), \quad (14.31)$$

9. Якщо умова не виконується, то крок був вибраний невірний, тобто з цим кроком перескочили через оптимум і крок потрібно зменшити, наприклад в два рази $\lambda = \frac{\lambda}{2}$ і переходимо до пункту 7 (рис. 14.5).

10. Якщо умова (14.31) виконується, то запам'ятовуємо координати вектора \bar{x}^1 і переходимо до пункту 4.

Схема алгоритму описаного методу представлена на рис. 14.6.

14.3.3 Метод Флетчера – Рівса

Цей метод дозволяє знайти мінімум нелінійної цільової функції багатьох змінних вигляду

$$M = F(x_1, x_2, \dots, x_N)$$

при відсутності обмежень. Метод заснований на застосуванні часткових похідних цільової функції по незалежним змінним і перевизначений для дослідження унімодальних функцій. За його допомогою можна досліджувати і мультимодальні функції, однак в цьому випадку слід брати декілька вхідних точок і перевіряти,

однаково чи в усіх випадках рішення. Схема алгоритму метода Флетчера – Рівса представлена на рис.14.6.

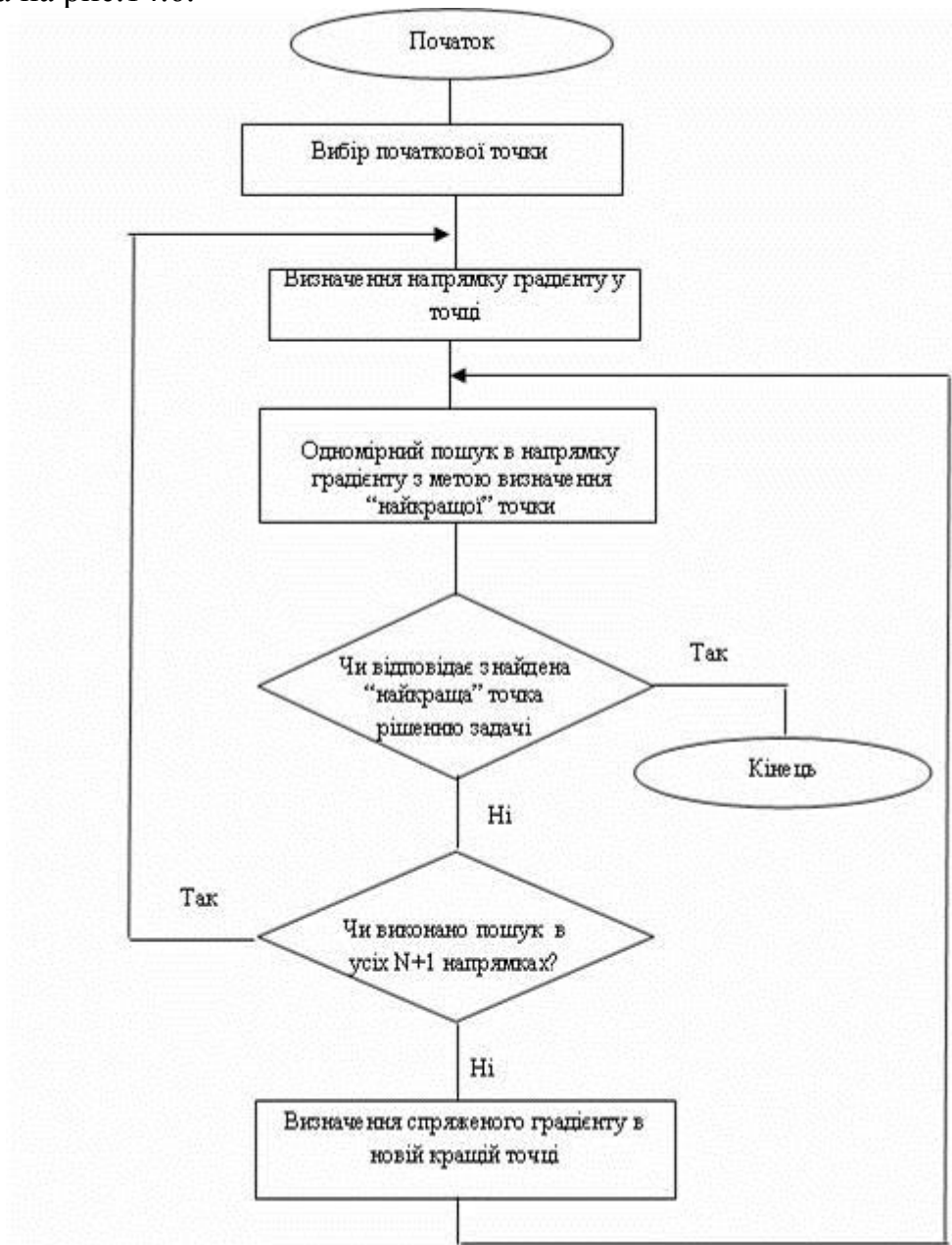


Рисунок 14.6 - Схема алгоритму метода Флетчера - Рівса

Виконується він наступним чином. Спочатку вибирається підходяща початкова точка простору проектування й шляхом обчислення компонент вектора градієнта визначається напрямок найшвидшого спуску. Індекс $k=1$ відповідає вхідній точці. Після цього в напрямку найшвидшого спуску ведеться одномірний пошук по формулі

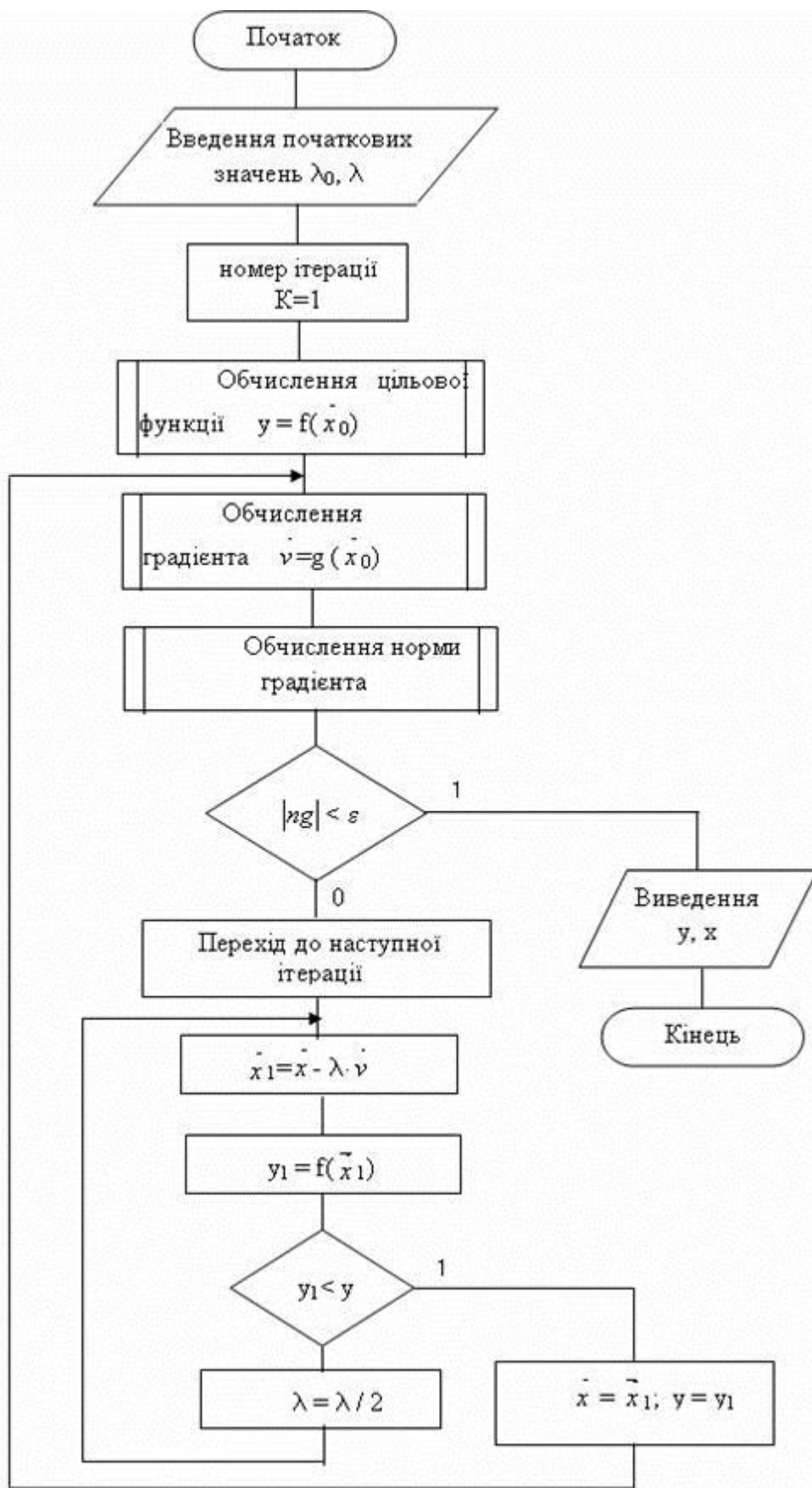


Рисунок 14.7 - Схема алгоритму методу найшвидшого спуску

$$s_j^{(k)} = \frac{-\left(\frac{\partial F}{\partial x_j}\right)^{(k)}}{\left[\sum_{j=1}^N \left(\frac{\partial F}{\partial x_j}\right)^2\right]^{\frac{1}{2}}}, \quad i = 1, 2, \dots, N,$$

$$x_{i,нов} = x_{i,ст} + \lambda s_i, \quad i = 1, 2, \dots, N,$$

де λ – зміщення в напрямку вектора градієнту. Знайшовши мінімум в цьому напрямку, визначають напрямки нових одиничних векторів, які дещо відрізняються від напрямку нового вектора градієнту і являють собою лінійні комбінації вектору градієнту на даному кроці і вектора градієнту, отриманого на попередньому кроці. Нові компоненти одиничних векторів записуються у вигляді

$$x_j^{(k+1)} = \frac{-\frac{\partial F^{(k+1)}}{\partial x_j} + \beta^{(k)} s_j^{(k)}}{\left\{ \sum_{j=1}^N \left[-\left(\frac{\partial F}{\partial x_j}\right)^{(k+1)} + \beta^{(k)} s_j^{(k)} \right]^2 \right\}^{\frac{1}{2}}}, \quad i = 1, 2, \dots, N, \quad (14.19)$$

де

$$\beta^{(k)} = \frac{\sum_{j=1}^N \left[\left(\frac{\partial F}{\partial x_j}\right)^{(k+1)} \right]^2}{\sum_{j=1}^N \left[\left(\frac{\partial F}{\partial x_j}\right)^{(k)} \right]^2}. \quad (14.20)$$

Індекс k вказує на послідовність обчислень в процесі ітерацій. Нові напрямки називаються «спряженими» і відповідають поточній локальній квадратичній апроксимації функції, а фактично представляють собою рух по дну яру (рис. 14.7).

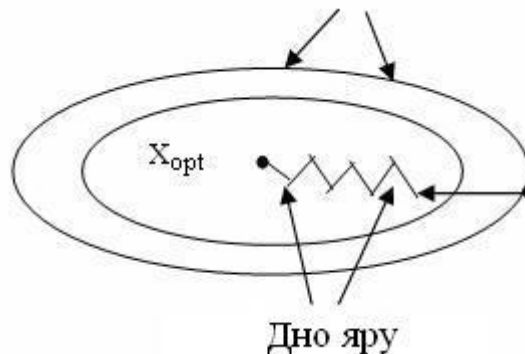


Рисунок 14.7 – Зміна напрямків руху \bar{s}_i по дну яру

Після цього по новому напрямку (іншому склону яру) проводять одномірний пошук і, знайшовши мінімум, перевіряють, чи досягнутий необхідний ступінь збіжності. Якщо перевірка показує, що це так, то рахунок припиняється. В протилежному випадку визначають нові спряжені напрямки, k збільшують на одиницю і продовжують процес до тих пір, доки не буде забезпечена збіжність або доки пошук не буде проведений по всім $N+1$ напрямкам. Закінчивши цикл пошуку по $N+1$ напрямкам, починають новий цикл, в якому знову використовується напрямок найшвидшого спуску. Особливість цього алгоритму полягає в тому, що він дозволяє використати переваги градієнтних методів, які проявляються при дослідженні цільової функції з перервними похідними. Так як $N+1$ напрямків пошуку другої сукупності відрізняються від напрямків одиничних векторів градієнту, то пошук не «зависає на

перегині», а йде вздовж лінії, яка з'єднує точки перегинів лінії рівня, яка, як правило, проходить через точку оптимуму. Взагалі можна стверджувати, що методи, основані на визначенні нових напрямків пошуку на основі накопичених даних про локальну поведінку функції, по своїй природі більш ефективні, ніж методи, в яких напрямок пошуку задається заздалегідь. Саме тому метод Флетчера – Рівса володіє більшими перевагами у порівнянні з методами найшвидшого спуску або підйому. Його недолік полягає в тому, що, він є складнішим ніж вказані методи, він вимагає розробки більш складних програм.

14.3.4 Метод Девідона – Флетчера – Пауела

Метод Девідона – Флетчера – Пауела являє собою алгоритм оптимізації, пристосований для відшукування безумовного мінімуму цільової функції, яка залежить від декількох змінних і яка має вигляд

$$M = F(x_1, x_2, \dots, x_N) \quad (14.21)$$

Необхідні часткові похідні цільової функції по незалежним змінним. Оскільки в основі методу лежить допущення про унімодальність цільової функції, в тих випадках, коли є підстави припускати, що вона не є такою, потрібно брати декілька вхідних точок. На рис. 14.9 представлена схема алгоритму методу Девідона – Флетчера – Пауела. Спочатку в просторі проектування вибирають придатну початкову точку. Після цього, обчислюючи складові вектора градієнту визначають напрямок пошуку.

$$s_j^{(k)} = \frac{\sum_{j=1}^N H_{i,j} \left(\frac{\partial F}{\partial x_j} \right)^{(k)}}{\left\{ \sum_{j=1}^N \left[\sum_{j=1}^N H_{i,j} \frac{\partial F}{\partial x_j} \right]^2 \right\}^{\frac{1}{2}}}, \quad i = 1, 2, \dots, N,$$

тут k – номер ітерації, а $H_{i,j}$ – елементи симетричної позитивно визначеної матриці розмірності $N \times N$. В процесі ітерацій ця матриця перетворюється в матрицю, зворотну матриці Гессе, елементами якої є другі часткові похідні цільової функції. Оскільки звичайно матриця заздалегідь невідома, то в якості початкової можна скористуватися будь-якою симетричною позитивно визначеною матрицею. Як правило, беруть найпростішу з них – одиничну матрицю. В цьому випадку пошук починається вздовж лінії найшвидшого спуску.

Одномірний пошук ведеться вздовж вхідного напрямку у відповідності з співвідношенням

$$x_{i,нов} = x_{i,ст} + \lambda S_i, \quad i = 1, 2, \dots, N, \quad (14.22)$$

де λ – величина кроку в напрямку пошуку. Знайшовши одновимірний оптимум, перевіряють результат на збіжність і, якщо вона досягнута, пошук припиняють. В протилежному випадку для подальшого пошуку вибирають новий напрямок, причому використовують колишнє співвідношення і нову матрицю \underline{H} , що визначається формулою

$$\underline{H}^{(k+1)} = \underline{H}^{(k)} + \underline{A}^{(k)} - \underline{B}^{(k)}. \quad (14.23)$$

Елементи матриць $\underline{A}^{(k)}$ і $\underline{B}^{(k)}$, які мають розмірність $N \times N$ обчислюються по формулам

$$\underline{A}^{(k)} = \frac{\underline{\Delta x}^{(k)} (\underline{\Delta x}^{(k)})^T}{(\underline{\Delta x}^{(k)})^T \underline{\Delta G}^{(k)}}, \quad (14.24)$$

$$\underline{B}^{(k)} = \frac{\underline{H}^{(k)} \underline{\Delta G}^{(k)} (\underline{\Delta G}^{(k)})^t \underline{H}^{(k)}}{(\underline{\Delta G}^{(k)})^t \underline{H}^{(k)} \underline{\Delta G}^{(k)}} \quad (14.25)$$

де верхнім індексом t позначені транспоновані матриці, а $\underline{\Delta x}^{(k)}$ і $\underline{\Delta G}^{(k)}$ – відповідно вектори - стовпці різниць значень x , і градієнтів в двох точках. Вектори - стовпці визначаються виразами

У відповідності з правилами матричного обчислення числівники виразів для $\underline{A}^{(k)}$ і $\underline{B}^{(k)}$ являють собою матриці розмірності $N \times N$, а знаменники є скалярами. Визначивши новий напрямок пошуку, проводять одномірний пошук і продовжують ітераційний процес. При виконанні алгоритму, який описується, пошук після першої спроби ведеться в тих напрямках, в яких цільова функція в найближчому околі має значення, які наближаються до оптимального.

$$\underline{\Delta x}^{(k)} = \underline{x}^{(k+1)} - \underline{x}^{(k)},$$

$$\underline{\Delta G}^{(k)} = \frac{\partial F^{(k+1)}}{\partial x} - \frac{\partial F^{(k)}}{\partial x}.$$

Лише в рідких випадках ці напрямки співпадають з напрямком градієнту. Тому даний алгоритм часто називають методом «відхиленого» градієнту. Вказана властивість методу Девідона – Флетчера – Пауела дозволяє обходити труднощі, які зв'язані з розривами похідних в просторі проектування. Широко розповсюджена думка, що цей метод є найбільш ефективним з усіх градієнтних методів. На відміну від методу Флетчера – Рівса він дає повну інформацію про кривизну поверхні цільової функції в точці мінімуму, однак при цьому вимагається більший об'єм пам'яті і більший час лічби для обробки матриці \underline{H} .

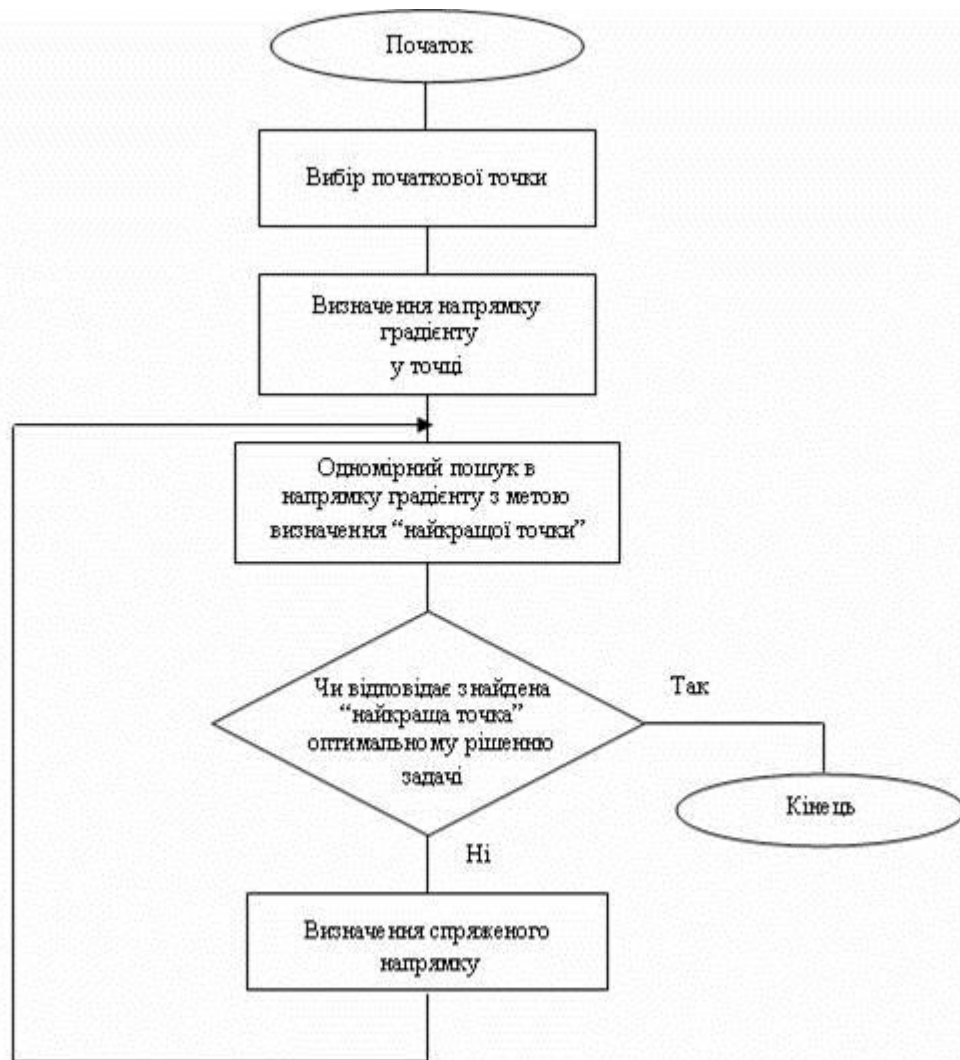


Рисунок 14.9 - Схема алгоритму методу Девідона – Флетчера – Пауела

14.3.5 Метод конфігурацій Хука – Дживса

Цей метод полегшує пошук і не вимагає обчислення похідних. Пошук ведеться вздовж ліній розриву похідних у припущенні, що зміщення в просторі проектування, які опинилися вдалими на ранній стадії пошуку, можуть призвести до успіху і на його більш пізніх стадіях. Метод Хука – Дживса перевизначений для пошуку мінімуму унімодальної функції багатьох змінних

$$M = F(x_1, x_2, \dots, x_n) \quad (14.26)$$

при відсутності обмежень. На рис. 14.10 представлена схема алгоритму цього методу.

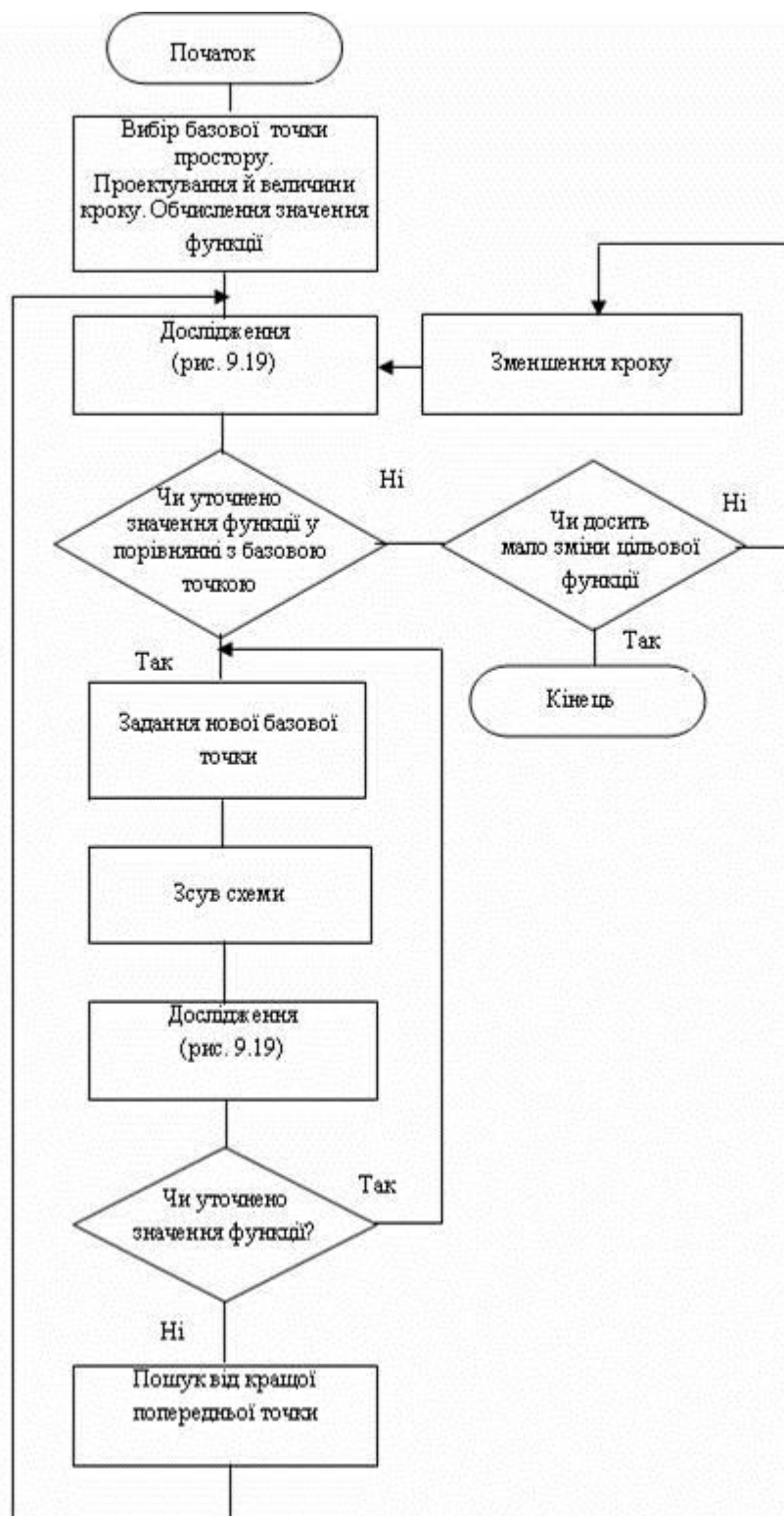


Рисунок 14.10 - Схема алгоритму методу конфігурацій Хука – Дживса

Виконується він наступним чином. Спочатку вибирається вхідна базова точка простору проектування і величини кроків, що будуть використані при дослідженні функції. Після цього у відповідності зі схемою рис. 14.11 проводиться дослідження із заданими приростом у напрямках, відповідних всім незалежним змінним.

Там, де отримано уточнене значення функції, розміщують нову тимчасову базову точку. Закінчивши етап дослідження, вибирають нову базову точку і виконують «зрушення схеми». Ця операція полягає в екстраполяції вздовж лінії, що з'єднує нову і колишню базові точки. Відстань зсуву за нову базову точку декілька перевищує

відстань між двома колишніми базовими точками. Математично екстраполяція визначається формулою (14.27)

$$x_{i,0}^{(k+1)} = x_i^{(k+1)} + \alpha(x_i^{(k+1)} - x_i^{(k)}) \quad (14.27)$$

де $x_{i,0}^{(k+1)}$ – нова тимчасова базова точка, або «точка зростання», i – змінний індекс, k – порядковий номер стадії пошуку, а α – коефіцієнт підсилення, що більше або рівний одиниці. Після цього досліджують окіл нової тимчасової базової точки, щоб з'ясувати, чи не містить вона точку, прийнявши яку за наступну базову можна наблизитися до оптимального рішення. Цей пошук також ведеться по схемі, яка показана на рис. 14.11.

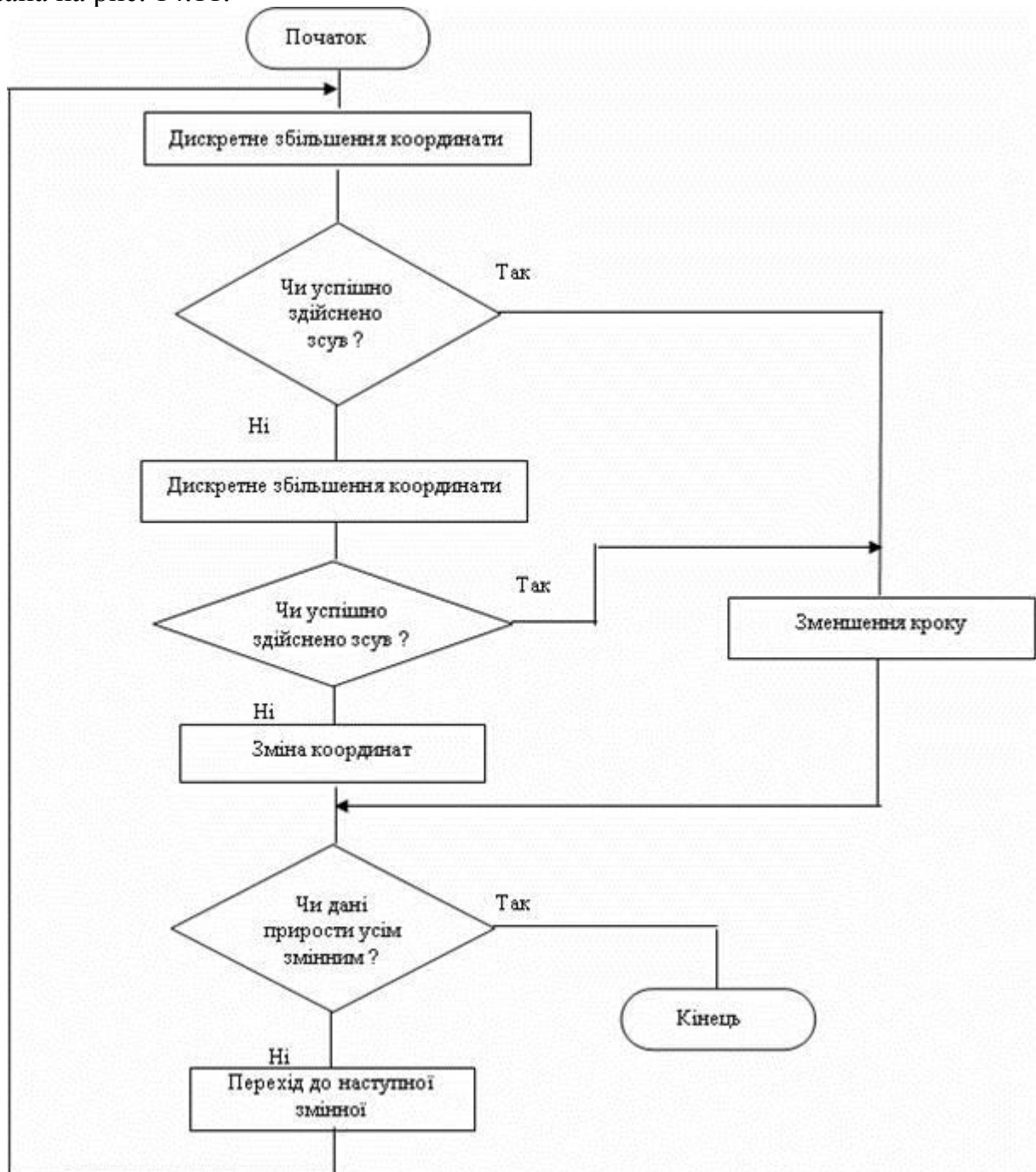


Рисунок 14.11 - Алгоритм дослідження цільової функції на основі методу Хука Джівса

Якщо знайдена тимчасова точка зростання або одна з сусідніх з нею точок має перевагу перед іншими, то вся процедура повторюється з використанням її в якості базової. Завдяки введенню коефіцієнта підсилення, кожне наступне дослідження околу точки здійснюється на все більшому і більшому віддаленні від вхідної точки до

тих пір, доки в процесі пошуку не виявиться пройденим пік або лінія розриву похідної. В цьому випадку вертаються до попередньої «кращої базової точки», звужують область дослідження і повторюють весь процес знову. Якщо крок, який зменшується, послідовно виявляється меншим за деяку заздалегідь задану величину і при цьому відсутня помітна зміна значення цільової функції, пошук припиняється. Після декількох змін напрямку пошуку метод Хука – Дживса забезпечує збіг розподілу розрахункових точок з лінією розриву похідних. Звичайно після завершення вибору схеми пошуку зсуву на кожному наступному кроці збільшується, доки не перевищить величину вхідного кроку в 10 або навіть в 100 раз. Тому у випадку, коли зрушення виявляється невдалим, єдиний засіб продовжити пошук — повернутися до найбільш вдалої з базових точок і почати усе спочатку. Той факт, що даний алгоритм володіє властивістю «прискорюватися», сприяє підвищенню його загальної ефективності. Друга перевага методу Хука – Дживса – можливість отримання за його допомогою наближеного рішення, якість якого безупинно підвищується на всіх стадіях чисельного рішення. Особливо явно переваги подібних засобів виявляються при відшуканні екстремумів на гіперповерхнях, які містять глибокі вузькі западини, тобто тоді, коли градієнтні методи неефективні.

14.3.6 Метод конфігурацій Розенброка

Метод конфігурацій Розенброка заснований на пошуку мінімуму вздовж ліній розриву похідних і часто виявляється ефективним, коли інші методи не дозволяють отримати рішення. Його нерідко називають «методом обертання вісей координат», оскільки дослідження в околі вибраної точки ведеться саме таким методом. На відміну від попередніх методів, в яких вхідним змінним надають незалежні прирости, в методі Розенброка система координат повертається так, щоб одна з вісей була направлена вздовж лінії розриву похідних, положення якої визначається в результаті попереднього дослідження. Решта осей утворюють з нею ортогональну систему координат. Метод Розенброка заснований на припущенні про унімодальність цільової функції і перевизначений для відшукання мінімуму функції багатьох змінних вигляду

$$M = F(x_1, x_2, \dots, x_n) \quad (14.28)$$

при відсутності обмежень. На рис. 14.12 показана схема алгоритму, який використовується в цьому методі.

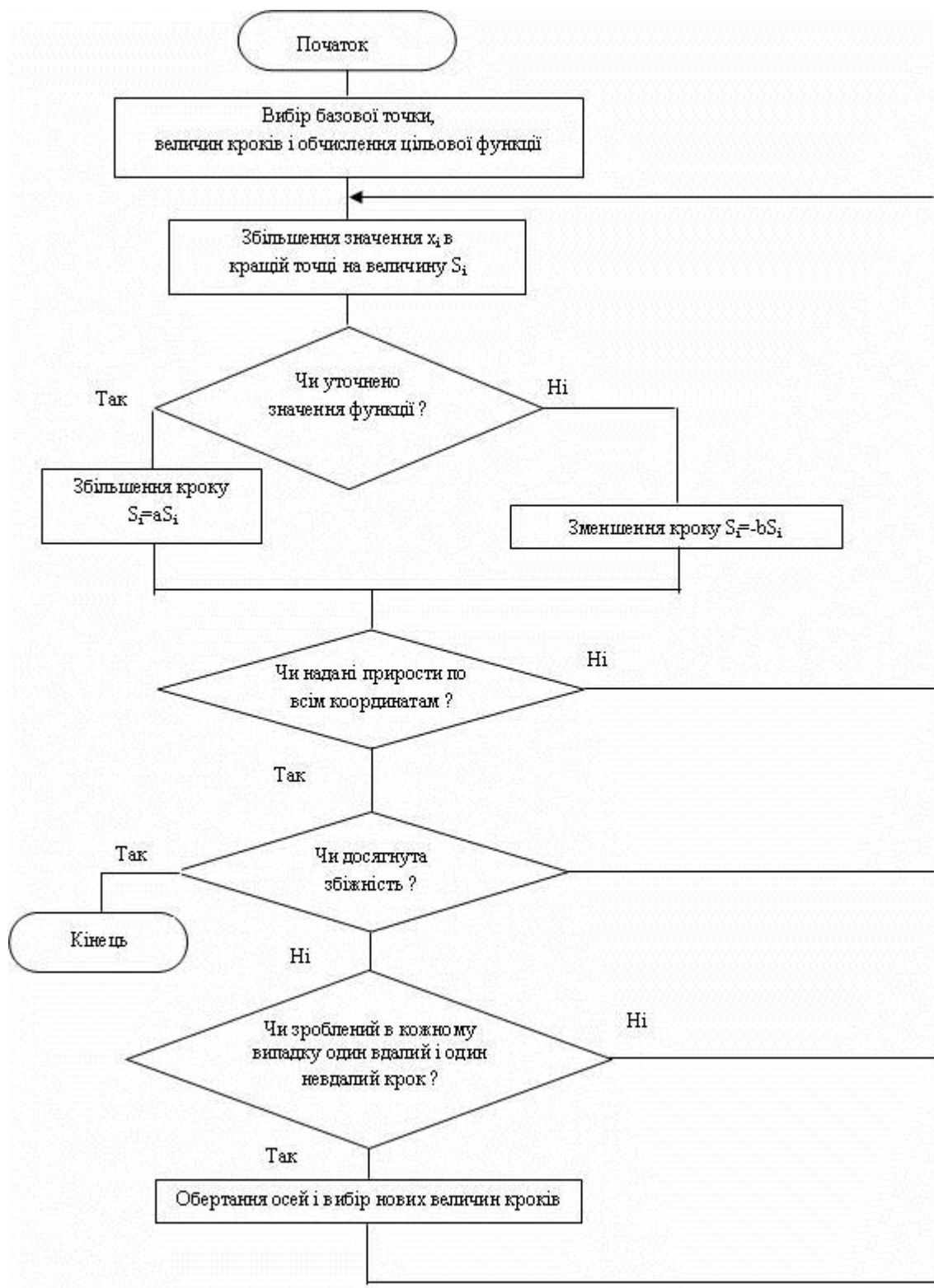


Рисунок 14.12 - Блок-схема алгоритму метода конфігурації Розенброка

Виконується він наступним чином. Спочатку вибирають початкову точку, задають початкові величини кроків $S_i (i = 1, 2, \dots, N)$ і обчислюють цільову функцію. Після цього кожній змінній x_i надають приріст S_i в напрямку, паралельному до відповідної осі координат в просторі проектування, і знову обчислюють цільову функцію F . Якщо її нове значення виявляється меншим за попереднє, то зсув вважається вдалим і наступний крок збільшується у відповідності з формулою

$$S_i = \alpha S_i, \quad (14.29)$$

де $\alpha > 1$. Якщо ж нове значення F виявляється більше за попереднє, то зсув вважається невдалим і наступний крок визначається по формулі

$$S_j = -\beta S_j, \quad (14.30)$$

де $\beta < 1$.

Здійснивши зсув по всім змінним, перевіряють збіжність і, якщо вона досягнута, пошук припиняють. В протилежному випадку вводять додаткову перевірку, щоб з'ясувати, чи були зроблені хоча б один успішний і один безуспішний зсув в кожному напрямку. Якщо такого поєднання зсув не було, вся процедура повторюється, починаючи з першої змінної. Якщо ж подібне поєднання виявлене для кожного напрямку, осі повертають так, щоб вхідний напрямок пошуку співпав з найбільш перспективним з раніше розглянутих напрямків. Після цього вибирають нові значення кроків і продовжують пошук по всім змінним, користуючись новою системою координат.

На відміну від інших, даний алгоритм націлений на пошук оптимальної точки в кожному напрямку, а не просто на фіксований зсув по всім напрямкам. Величина кроку в процесі пошуку безупинно змінюється в залежності від конфігурації рельєфу поверхні. Поєднання обертання вектору, направленою вздовж лінії розриву похідних, з регулюванням масштабу робить метод Розенброка виключно ефективним при розв'язанні складних задач оптимізації.

Література

1. Щуп Т. Решение инженерных задач на ЭВМ. – М.: Мир, 1982. – 235с.
2. Бахвалов Н. С. Численные методы . Т. И. Анализ, алгебра, обычные дифференциальные уравнения. – М.: Наука, 1975. – 631 с.
3. Ляшенко М.Я., Головань М.С. Чисельні методи: Підручник. Либідь. 1996. – 288 с.
4. Крылов В. И. и др. Численные методы . – М.: Наука, 1978. – 1979. – Т. 2. – 400 с.
5. Форсайт Дж., Малькольм., Моулер Р. Машинные методы математических вычислений. – М.: Мир, 1980. – 279с.
6. Д. Мэтьюз, Г. Цинк, Д. Куртис. Численне методи. Использование Matlab, –М. Издательский дом “Вильямс”, 2001. – 720 с. 720 с.
7. Гил Ф., Мюрей У. Численные методы условной оптимизации. – М.: Мир, 1977. – 290 с.
8. Василев Ф. П. Лекции по методам решения экспериментальных задач. – М.: Изд – во Моск. Ун – то, 1984. – 374 с.
9. Полак Е. Численные методы оптимизации. – М. : Мир, 1974.
10. Пшеничний Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. – М. : Наука, 1975.
11. Моделирование и оптимизация на ЭВМ радиоэлектронных устройств / Под ред. З. М. Бенесона. – М. : Радио и связь, 1981. – 272 с.
12. *И.В. Кузьмин, М.М. Биков, С.М. Москвина, А.И. Кузьмин. Методы оптимизации сложных систем. – Винница: ВДГУ, 2003. – 165с.*