МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

В. П. Северин, О. М. Нікуліна

# МЕТОДИ ТА АЛГОРИТМИ ОДНОВИМІРНОЇ ОПТИМІЗАЦІЇ

Навчальний посібник з навчальної дисципліни
«Дослідження операцій» для студентів
напряму «Інформаційні технології»

Заверджено
редакційно-видавничою
радою НТУ «ХПІ»,
протокол № 2 від 27.06.2024

Харків
НТУ «ХПІ»
2025

УДК 519.85(075)
С 28

Рецензенти:
*С. Г. Удовенко*, д-р техн. наук, проф., Харківський національний університет імені С. Кузнеця;
*І. І. Смирнова*, канд. техн. наук, доц., ТОВ «ТЕХНІЧНИЙ УНІВЕРСИТЕТ» «МЕТІНВЕСТ ПОЛІТЕХНІКА»

**Северин В. П.**

**С 28** Методи та алгоритми одновимірної оптимізації : навч. посіб. з курсу «Дослідження операцій» / В. П. Северин, О. М. Нікуліна. – Харків : НТУ «ХПІ», 2025. – 115 с.

**ISBN 978-617-05-0524-8**

Розглянуто теорію і методи одновимірної оптимізації. Викладено теорію необхідних і достатніх умов екстремуму функції однієї змінної. Наведено обґрунтування побудови, теоретичний опис і алгоритми більш ніж десяти основних методів одновимірної оптимізації. Надані алгоритми дозволяють полегшити вивчення методів і прискорити виконання лабораторних робіт. Дано задачі для лабораторних робіт.

Призначено для студентів напряму «Інформаційні технології».

Іл. 26. Табл. 4. Бібліогр. 32 назв.

**УДК 519.85(075)**

ISBN 978-617-05-0524-8

## ЗМІСТ

# ВСТУП

Математичні методи оптимізації, їх алгоритми та комп'ютерні програми є ефективним елементом наукомістких технологій, розробка яких у теперішній час особливо актуальна. Опубліковані книги за методами оптимізації мають великий попит як на прилавках книгарень, так і в бібліотеках.

Курс лекцій з дисципліни «Методи оптимізації» протягом кількох років читався автором в НТУ «ХПІ» для студентів спеціальностей «Інформатика», «Соціальна інформатика», «Системний аналіз». Цей курс включає розділи «Методи одновимірної оптимізації», «Методи безумовної мінімізації», «Методи умовної оптимізації» та «Спеціальні методи оптимізації». У запропонованому навчальному посібнику «Методи та алгоритми одновимірної оптимізації», що складається з чотирьох розділів, дано вступ у методи оптимізації, розглянуто теорію, методи та алгоритми одновимірної оптимізації.

Перший розділ присвячений визначенню та розгляду тих основних понять, які притаманні всьому курсу «Методи оптимізації». Представлено предмет методів оптимізації, введено основні поняття теорії, наведено приклади задач оптимізації та дано класифікацію методів. Розрізняють оптимізацію у широкому та вузькому значенні, обґрунтовується необхідність вивчення багатьох методів оптимізації. Вводяться поняття змінних параметрів, допустимої області, цільової функції, задачі оптимізації та її вирішення, методу оптимізації. Розглядаються приклади задач оптимізації, показується корисність їх розв'язання і важливість застосування сучасних ефективних методів. Надаються різні принципи класифікації методів оптимізації.

У більшості методів багатовимірної оптимізації використовуються методи одновимірного пошуку. У другому розділі розглядаються питання, пов'язані з теорією та практикою одновимірної оптимізації. Викладається теорія необхідних та достатніх умов екстремуму функції однієї змінної. Розглядаються властивості унімодальної функції, вводиться поняття інтервалу невизначеності точки мінімуму унімодальної функції, наводяться основні принципи пошуку інтервалу невизначеності та його зменшення. Обгрунтовуються метод Свенна для пошуку інтервалу невизначеності та метод адаптації кроку для мінімізації функції однієї змінної, який не вимагає знаходження інтервалу невизначеності.

Третій розділ присвячений методам виключення інтервалів для зменшення інтервалу невизначеності. Обґрунтовуються методи рівномірного пошуку, дихотомії, бісекції, поділу інтервалу навпіл. Розглядаються властивості чисел Фібоначчі та конструюється метод Фібоначчі, вивчаються властивості золотого перетину, та обґрунтовується метод золотого перетину. Виконується порівняння методів виключення інтервалів щодо їх ефективності.

У четвертому розділі розглядаються методи поліноміальної інтерполіції нульового, першого та другого порядків, що ґрунтуються на апроксимації унімодальної цільової функції поліномом з подальшою інтерполяцією точки мінімуму функції. Обґрунтовуються методи квадратичної та кубічної інтерполяції, метод січних, одновимірний метод Ньютона. Дається порівняння швидкості асимптотичної збіжності методів поліноміальної інтерполяції.

Наведено реалізовані алгоритми всіх розглянутих методів одновимірної оптимізації. Подані алгоритми дозволяють полегшити розуміння методів та прискорити виконання лабораторних робіт. Надано завдання для лабораторних робіт.

Особливістю посібника є повніше викладення методів одновимірної оптимізації в порівнянні з більшістю навчально-методичних видань. Включено оригінальний матеріал, наприклад, метод адаптації кроку, який не викладено в іншій навчально-методичній літературі.

Для засвоєння матеріалу достатньо володіння стандартними курсами математичного аналізу та лінійної алгебри.

# РОЗДІЛ 1. ВСТУП У МЕТОДИ ОПТИМІЗАЦІЇ

Цей розділ присвячений визначенню та розгляду тих основних понять, які вивчатимуться у всьому курсі «Методи оптимізації». Розрізняється оптимізація у широкому та вузькому значенні, показується походження терміна «оптимізація», визначається предмет методів оптимізації, обґрунтовується необхідність вивчення багатьох методів оптимізації. Вводяться поняття змінних параметрів оптимізації, простору параметрів та допустимої області, критерію якості та цільової функції, задачі оптимізації та вирішення задачі оптимізації, методу оптимізації. Розглядаються приклади задач оптимізації, показується корисність їх вирішення та важливість застосування сучасних ефективних методів. Розглядаються різні принципи класифікації методів оптимізації і дається класифікація методів за застосовуваними засобами, за математичним формулюванням задач оптимізації, за порядком похідних. Наводиться опис лабораторної роботи для початку практичного вивчення методів оптимізації.

## 1.1. Предмет методів оптимізації

Слово «*оптимальний*» у перекладі з грецької optimus означає «найкращий», тому *оптимізація у широкому значенні* означає пошук найкращого варіанта з безлічі різних альтернатив.

Оптимізація різних об'єктів та їх систем є основою інженерної діяльності, оскільки мета праці будь-якого інженера – створення нових, ефективніших систем за меншими можливими ресурсами, зокрема фінансовими. У теперешній час знання методів оптимізації студентами вищої школи так само потрібне, як і знання основ математичного аналізу, фізики, інформатики та інших дисциплін.

Для пошуку найкращого варіанта деякої системи необхідно задати для неї критерій якості та визначити, за допомогою яких параметрів на нього можна впливати. Вибраний критерій якості являє собою функцію змінних параметрів, яку необхідно мінімізувати або максимізувати. *Оптимізація у вузькому значенні* полягає в мінімізації або максимізації функції за наявності або відсутності обмежень

Оптимізація може здійснюватися різними засобами – як за допомогою простої арифметики, так і за допомогою складних аналітичних і числових процедур. Існує безліч методів оптимізації, найпотужнішими з яких є числові методи, які найбільш повно використовують можливості сучасної обчислювальної техніки.

Термін «оптимізація» вперше запровадив у математиці Лейбніц. Методи оптимізації почали розвиватися після Другої світової війни, а більшість класичних методів оптимізації створено у 50–60-ті роки минулого століття. Поряд з поняттям методів оптимізації застосовують його синонім – *математичне програмування*. З іншого боку, математичне програмування є частиною *дослідження операцій* – підрозділу обчислювальної математики, що вивчає властивості різних операцій.

*Операцією* в математиці називається відображення декартового добутку деяких множин в іншу множину:

$$M_1 \times M_2 \times \ldots M_n \to M \ .$$

Поширеним прикладом операції є функція кількох змінних $f(\mathbf{x})$, тому дослідження функцій – це частина дослідження операцій.

Запити людської діяльності ставлять безліч задач оптимізації різних типів, але навіть для різних задач однакового типу немає єдиного універсального методу вирішення. Це приводить до нагальної необхідності вивчення багатьох методів оптимізації та аналізу їх переваг і недоліків з метою найбільш ефективного застосування методів.

Людський геній винайшов багато ефективних числових методів оптимізації, які мають велику пізнавально-методичну і практичну цінність. Числові методи для розв’язання різних типових задач оптимізації і становлять предмет навчальної дисципліни «Методи оптимізації».

## 1.2. Поняття та визначення теорії оптимізації

Для формального подання задач оптимізації та методів їх вирішення необхідно запровадити основні поняття та визначення теорії оптимізації.

1. *Параметри оптимізації* $x_1$, $x_2$, …, $x_n$ – це ті незалежні змінні, які можна змінювати при вирішенні задачі оптимізації, $n$ – кількість змінних. Ці змінні також називаються *варійованими параметрами, змінними параметрами* або просто *змінними*. Впорядкований набір змінних утворює *вектор змінних параметрів* $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Вважатимемо, що кожен параметр належить множині дійсних чисел $R = (-\infty, +\infty)$: $x_i \in R$, $i = \overline{1, n}$.

2. *Простір параметрів* – це множина векторів змінних параметрів $\mathbf{x}$, що утворюють $n$-вимірний простір $R^n$. При цьому вважають, що простір параметрів являє собою декартовий добуток множин дійсних чисел $R^n = R \times R \times \ldots \times R$. Для фіксованого вектора $\mathbf{x}$ його можна уявити точкою простору параметрів. Розмірність простору параметрів $n$ визначається розмірністю його вектора $\mathbf{x}$, для чого використовується позначення $n = \dim \mathbf{x}$. Тут $\dim$ – це скорочення від англійського слова dimension – розмірність.

3. *Допустима область* – це множина векторів параметрів, що задовольняють обмеженням задачі оптимізації, вона позначається символом $D$. Обмеження можуть задаватися у вигляді рівностей та нерівностей:

$$D = \{\, \mathbf{x} \mid g_i(\mathbf{x}) \le 0,\, i = \overline{1, p};\;\; h_j(\mathbf{x}) = 0,\; j = \overline{1, r}\,\}.$$

За наявності обмежень говорять про задачу умовної оптимізації, а за їх відсутності – про задачу безумовної оптимізації.

4. *Критерій якості* – це функціонал, що відображає якість системи, що оптимізується. Функціонал – це будь-який оператор, результатом якого є число. Критерій якості зазвичай подається функцією вектора змінних параметрів $f(\mathbf{x})$, що приймає дійсні значення. Така функція

виконує математичну операцію відображення простору $R^n$ в $R$, тобто $f: R^n \to R$.

5. *Цільова функція* – це критерій якості із зазначенням типу його оптимальності – мінімальності чи максимальності: $f(\mathbf{x}) \to \min$ або $f(\mathbf{x}) \to \max$. При розробці сучасного програмного забезпечення методів оптимізації для визначеності вважають, що цільову функцію необхідно мінімізувати. Якщо необхідна максимізація цільової функції – $g(\mathbf{x}) \to \max$, то переходять до допоміжної функції $f(\mathbf{x}) = -g(\mathbf{x})$, мінімізація якої приведе до максимізації вихідної функції:

$$f(\mathbf{x}) =\to \min \quad \Rightarrow \quad g(\mathbf{x}) \to \max .$$

Якщо $g(\mathbf{x}) \neq 0$, то можна також покласти $f(\mathbf{x}) = 1/g(\mathbf{x})$.
Таким чином, будь-яку задачу максимізації цільової функції можна звести до відповідної задачі мінімізації.

6. *Задача оптимізації* – це задача, яка полягає в мінімізації цільової функції $f(\mathbf{x})$ за відсутності або наявності додаткових умов. Якщо додаткові умови відсутні, має місце задача безумовної оптимізації, що полягає у мінімізації цільової функції $f(\mathbf{x})$ і коротко записується як $\min f(\mathbf{x})$. Якщо є додаткові умови обов'язкової приналежності вектора змінних $\mathbf{x}$ до допустимої області $D$, то ставиться задача умовної оптимізації, яка записується як $\min\limits_{\mathbf{x} \in D} f(\mathbf{x})$. За відомої цільової функції та допустимої області говорять, що задана модель задачі оптимізації.

7. *Розв'язання задачі оптимізації* – це значення вектора змінних параметрів $\mathbf{x}^*$, яке задовольняє умову задачі оптимізації. Розв'язання задачі безумовної оптимізації визначають рівностями:

$$f(\mathbf{x}^*) = \min f(\mathbf{x}) \quad \text{або} \quad \mathbf{x}^* = \arg\min f(\mathbf{x}) .$$

Оптимальне значення цільової функції позначається через $f^* = f(\mathbf{x}^*)$.

8. *Метод оптимізації* – це метод для виконання ітераційного процесу вирішення задачі оптимізації, який дозволяє виключити повний перебір варіантів і отримати розв'язок задачі $\mathbf{x}^*$ більш раціональним способом. Будь-який метод оптимізації формує послідовність точок $\{\mathbf{x}_k\}$ у просторі параметрів $R^n$, яка повинна збігатися до розв'язку задачі оптимізації: $\mathbf{x}_k \underset{k\to\infty}{\to} \mathbf{x}^*$. Ефективність методу оптимізації визначається як отриманою точністю розв'язання задачі оптимізації, так і величиною витрачених на її вирішення обчислювальних ресурсів.

### 1.3. Приклади задач оптимізації

Існує багато задач оптимізації, які розрізняються за ступенем популярності та корисності. Розглянемо найбільш відомі та показові задачі оптимізації.

1. Задача Дідони.

За давнім переказом легендарна цариця Дідона заснувала місто Карфаген у 825 році до нашої ери на морському узбережжі північної Африки за іронічної умови вождів місцевих племен виділити їй стільки землі, скільки вона зможе охопити бичачою шкірою. Дідона розрізала шкуру на тонкі мотузки, зв'язала їх та з урахуванням морського берега охопила прямокутну ділянку, що має максимальну площу і достатня для заснування міста. Так була вирішена одна з найдавніших задач оптимізації.

Позначимо через $x_1$ і $x_2$ довжини сторін прямокутної ділянки, площу якої $S(x) = x_1 x_2$ необхідно максимізувати при обмеженій довжині трьох сторін $l(\mathbf{x}) = 2x_1 + x_2$, де $\mathbf{x} = (x_1, x_2)$, $l(\mathbf{x}) = C$, тому що четверта сторона прямокутника була обмежена морем (рис. 1.1). Таким чином, необхідно максимізувати цільову функцію $f(\mathbf{x}) = S(\mathbf{x})$ при обмеженні-рівності $2x_1 + x_2 = C$. Зауважимо, що вирішення цієї задачі дозволило мінімізувати кількість будівельного матеріалу для побудови оборонної стіни навколо міста при найбільшій можливій його площі.

Задача Дідони є задачею нелінійного програмування, оскільки в ній цільова функція нелінійна щодо змінних.

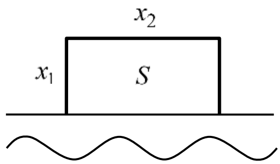


Рис. 1.1. Задача Дідони

2. Задача оптимізації плану виробництва.

Необхідно скласти план виробництва $n$ видів продукції з використанням $m$ типів ресурсів таким чином, щоб отримати максимальний прибуток підприємства. Введемо позначення: $a_{ij}$ – норма витрати $i$-го ресурсу виробництва одиниці $j$-тої продукції, $i=\overline{1,m}$, $j=\overline{1,n}$; $b_i$ – запас $i$-го ресурсу, $i=\overline{1,m}$; $c_j$ – норма прибутку від продажу одиниці $j$-тої продукції, $j=\overline{1,n}$; $x_j$ – запланована кількість $j$-го виду продукції, $j=\overline{1,n}$. Загальна вартість виробленої продукції визначається цільовою функцією

$$f(\mathbf{x})=\sum_{j=1}^{n}c_jx_j\rightarrow\max,$$

яка має бути максимізована вибором $n$ значень $x_j\ge 0$, які є змінними параметрами. Але при цьому мають бути враховані обмеження наявності ресурсів:

$$\sum_{j=1}^{n}a_{ij}x_j\le b_i,\quad i=\overline{1,m}.$$

Тут цільова функція та функції обмежень є лінійними, тому така задача є задачею лінійного програмування.

3. Транспортна задача.

Необхідно скласти план перевезень деякого товару з $m$ складів до $n$ магазинів таким чином, щоб витрати на ці перевезення були мінімальними. Введемо позначення: $a_i$ – кількість одиниць товару на $i$-тому складі, $i=\overline{1,m}$; $b_j$ – кількість одиниць товару, замовлена $j$-тим магазином, $j=\overline{1,n}$; $c_{ij}$ – вартість перевезення одиниці товару з $i$-го складу в $j$-тий магазин; $x_{ij}$ – запланована кількість товару, що перевозиться з $i$-го складу в $j$-тий магазин. Загальна вартість перевезень визначається цільовою функцією

$$f(\mathbf{x})=\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ij}x_{ij}\rightarrow\min,$$

мінімум якої має бути досягнутий вибором $mn$ значень змінних параметрів $x_{ij}\geq 0$. При цьому повинні бути враховані обмеження наявності товару на складах та потреб магазинів:

$$\sum_{j=1}^{n}x_{ij}\leq a_i,\quad i=\overline{1,m};\qquad \sum_{i=1}^{m}x_{ij}=b_j,\quad j=\overline{1,n}.$$

Цільова функція та функції обмежень є лінійними, тому транспортна задача також є задачею лінійного програмування.

4. Задача мінімізації вартості ракети.

При створенні триступеневої ракети для реалізації космічних програм США в 60-х роках минулого століття розроблена нелінійна цільова функція її вартості в мільйонах доларів $f(\mathbf{x})\rightarrow\min$, що залежить від 25 змінних параметрів, що включають значення мас елементів ракети і палива, тягові параметри, час роботи ступенів, а також інші конструктивні та технологічні параметри. На змінні параметри накладено 26 обмежень: 9 рівностей та 17 нерівностей, 10 обмежень представлені нелінійними функціями. Таким чином, ця задача умовної оптимізації є задачею нелінійного програмування. Американськими вченими Б. С. Рашем, Дж. Бракеном і Г. П. МакКорміком отримано розв’язок задачі методом послідовної безумовної мінімізації зі значенням цільової функції

$f^* = 2330$. Проте, подальші дослідження показали, що цільова функція багатоекстремальна, і насправді було знайдено її локальний мінімум. Глобальний мінімум з $f^* = 1359$ був обчислений методом $\Psi$-перетворення В. К. Чичинадзе, спеціально призначеним для вирішення багатоекстремальних задач оптимізації. Таким чином, застосування ефективнішого методу оптимізації дозволило б заощадити 971 мільйон доларів, тобто 42 % від спочатку розрахованої вартості ракети.

5. Задача комівояжера.

Є граф із виваженими ребрами. Вершини відповідають містам, а ваги ребер – вартості переміщення між містами. Потрібно знайти оптимальний шлях комівояжера, який представляє замкнутий цикл, що включає всі вершини і має мінімальну вагу. Ця задача є задачею дискретної оптимізації, оскільки змінними є вершини графа – дискретні значення.

6. Задача оптимізації системи управління.

Модель системи автоматичного управління є системою диференціальних рівнянь:

$$\frac{d\mathbf{X}}{dt} = F(\mathbf{x}, \mathbf{X}, u), \quad y = \mathbf{CX},$$

де $\mathbf{x}$ – вектор змінних параметрів системи; $\mathbf{X}$ – вектор її стану; $u$ – зовнішній вплив; $y$ – керована величина. При ступінчастій зміні $u = 1(t)$, де

$$1(t) = \begin{cases} 0, \, t < 0, \\ 1, \, t \geq 0, \end{cases}$$

система переходить до нового стану. Характер цього переходу визначається функцією керованої величини $y = y(\mathbf{x}, t)$, за якою визначають прямі показники якості системи управління: перерегулювання $\sigma(\mathbf{x})$, показник коливань $\zeta(\mathbf{x})$ і час регулювання $t_c(\mathbf{x})$. На перші два показники накладаються обмеження $\sigma(\mathbf{x}) \leq \sigma_m$, $\zeta(\mathbf{x}) \leq \zeta_m$, а третій показник мінімізується $t_c(\mathbf{x}) \to \min$. Ця задача є задачею нелінійного програмування.

### 1.4. Класифікація методів оптимізації

Для успішного застосування методів оптимізації необхідно чітко уявляти їх призначення та можливості. Існує багато методів оптимізації, які можна класифікувати за різними ознаками. Розглянемо основні ознаки класифікації.

I. За застосовуваними засобами.

1. Аналітичні методи. Використовують класичні методи диференціального обчислення. Для розв'язання задач безумовної оптимізації застосовується необхідна умова екстремуму – рівність нулю похідних цільової функції:

$$\frac{\partial f}{\partial x_i} = 0, \quad i = \overline{1, n}.$$

Для вирішення задач умовної оптимізації використовується метод множників Лагранжа. Аналітичні методи дають можливість отримати точний розв'язок задачі оптимізації. Загальний їх недолік у тому, що вони можуть застосовуватися тільки для простих задач з цільовими функціями, що диференціюються, заданими аналітично, при невеликій кількості змінних.

2. Методи перебору варіантів. Аналізуються різні варіанти вирішення задачі оптимізації та шляхом їх порівняння обирається найкращий варіант. Недолік цих методів полягає у їхній великій трудомісткості та малій точності.

3. Графічні методи. Основані на побудові графіків функцій однієї чи кількох змінних. Екстремум функції визначається безпосередньо шляхом аналізу її графіка. Вони прості та наочні, відразу показують, чи існує розв'язок. Застосовні лише для малого числа змінних.

4. Числові методи. Найширший клас методів, що найбільш повно

використовує можливості сучасної обчислювальної техніки. Це ітераційні методи, які полягають у послідовному переході від одного рішення $\mathbf{x}_k$ до кращого рішення $\mathbf{x}_{k+1}$. На відміну від аналітичних методів, за допомогою яких одержують точний розв'язок, числові методи надають наближений розв'язок.

5. Експериментальні методи. У деяких задачах побудову моделі задачі оптимізації важко або неможливо виконати, тому в них оптимальний розв'язок $\mathbf{x}^*$ отримують методом оптимального експерименту – як і в числових методах, перехід до кращого рішення проводиться ітераційно, проте значення самої цільової функції $f(\mathbf{x})$ визначаються експериментально. Це найдорожчі методи.

Для практичних додатків найефективніші числові методи оптимізації.

II. За математичним формулюванням задачі оптимізації.

1. Методи одновимірної оптимізації. Полягають у знаходженні мінімуму функції однієї змінної: $\min f(x)$, $\dim x = 1$. Ці методи також називаються методами одновимірного пошуку.

2. Методи багатовимірної безумовної оптимізації. Полягають у знаходженні мінімуму функції багатьох змінних: $\min f(\mathbf{x})$, $\dim \mathbf{x} \geq 2$.

3. Методи умовної оптимізації. Необхідно знайти $\min\limits_{\mathbf{x}\in D} f(\mathbf{x})$, де

$$D = \{\mathbf{x} \mid g_i(\mathbf{x}) \leq 0,\ i = \overline{1, p};\quad h_j(\mathbf{x}) = 0,\ j = \overline{1, r}\}.$$

Якщо хоча б одна з функцій задачі оптимізації нелінійна, то методи називаються методами нелінійного програмування.

4. Методи опуклої оптимізації. Формулювання задачі аналогічне пункту 3, однак $f(\mathbf{x})$ і $D$ є опуклими. У цьому випадку існує єдиний розв'язок $\mathbf{x}^*$.

5. Методи лінійного програмування. Формулювання задачі аналогічно пункту 3, тільки всі функції лінійні.

6. Методи мінімізації суми квадратів. Шукається мінімум функції

$$S(\mathbf{x}) = \sum_{k=1}^{m} [\, f_k(\mathbf{x})]^2 \,,$$

що становить суму квадратів функцій $f_k(\mathbf{x})$, $k = \overline{1, m}$. Якщо $f_k(\mathbf{x})$ – лінійні функції, то застосовується метод найменших квадратів, який зводить розв'язання задачі оптимізації до розв'язання системи лінійних алгебраїчних рівнянь.

7. Методи дискретного програмування. Задачі оптимізації, що розв'язуються цими методами, характеризуються тим, що змінні параметри $x_i$, $i = \overline{1, n}$ можуть набувати тільки дискретних значень. Якщо $x_i$ набувають лише цілі значення, то відповідна задача є задачею цілочислового програмування. Вона вирішується відповідними методами цілочислового програмування. Цей клас методів широко використовує дискретну математику та різні комбінаторні схеми.

III. За порядком похідних.

Числові методи оптимізації у свою чергу класифікують за максимальним порядком застосовуваних у цих методах похідних цільової функції.

1. Методи нульового порядку (методи прямого пошуку). У цих методах оптимізації цільової функції використовуються лише значення самої функції, а значення її похідних не використовуються. Перевага методів нульового порядку полягає в порівняно простій модифікації для вирішення широкого кола різних задач оптимізації.

2. Методи першого порядку. Ці методи застосовуються для оптимізації диференційованих цільових функцій і основані на використанні перших похідних функції. Перші похідні задаються градієнтом цільової функції, що являє собою вектор-стовпець її частинних похідних за змінними параметрами

$$\operatorname{grad} f(\mathbf{x}) = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)^{\mathrm{T}},$$

де Т – знак транспонування. Тому ці методи часто називаються градієнтними методами. Для градієнта також широко використовується позначення $\nabla f(\mathbf{x}) = \mathrm{grad}\, f(\mathbf{x})$ .

3. Методи другого порядку. Ці методи застосовуються для мінімізації двічі диференційованих цільових функцій і основані на використанні матриці других частинних похідних цільової функції – матриці Гессе

$$H(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Матриця Гессе або гесіан також позначається $\nabla^2 f(\mathbf{x}) = H(\mathbf{x})$ .

**Лабораторна робота**

Тема лабораторної роботи.

Розробка загальних підпрограм для методів одновимірної оптимізації.

Мета лабораторної роботи.

Розробити загальні підпрограми, які дозволяють відображати на екрані комп’ютера процес мінімізації функції однієї змінної у вигляді таблиці та графіка.

Порядок виконання лабораторної роботи.

1. Отримати у викладача варіант цільової функції $f(x)$ однієї змінної $x \in R$ з початковою точкою пошуку мінімуму $x_0 \in R$ .

2. Аналітично знайти точку $x^* \in R$ мінімуму заданої функції $f(x)$

та обчислити мінімальне значення функції $f^* = f(x^*)$.

3. Написати підпрограму обчислення значень функції $f(x)$ з вхідним параметром $x$ та вихідним параметром $f$.

4. Протестувати підпрограму обчислення функції $f(x)$ при обчисленні значень $f_0 = f(x_0)$ і $f^* = f(x^*)$ та порівняти $f^*$ з $f_0$.

5. Написати підпрограму ітераційного процесу наближення до точки мінімуму $x^*$ функції $f(x)$ з початкової точки $x_0$ відповідно до вимог:

1) ітерації повинні виконуватися за рекурентними формулами:

$$h_k = x^* - x_k\,, \qquad x_{k+1} = x^* + \alpha h_k\,, \qquad f_k = f(x_k)\,, \qquad k = 0, 1, 2, \ldots,$$

де $\alpha \in (0; 1)$ – задане число;

2) ітерації продовжуються доти, доки виконуються нерівності:

$$|h_k| > \varepsilon\,, \qquad k \le N\,,$$

де $\varepsilon$ – припустима похибка, $N$ – гранична кількість точок;

3) вхідні параметри підпрограми: $f(x)$, $x_0$, $x^*$, $\alpha$, $\varepsilon$, $N$;

4) вихідні параметри підпрограми: $X$ – масив усіх точок пошуку $x_k$, $F$ – масив відповідних значень функції $f_k = f(x_k)$;

5) на кожній ітерації виводити на екран комп'ютера рядок, що містить номер точки пошуку $k$, крок пошуку $h_k$, значення функції $f_k$ та значення змінної $x_k$;

6) після закінчення процесу оптимізації на екрані повинна відображатися таблиця, що має відповідний заголовок і являє процес мінімізації функції;

7) під таблицею необхідно відобразити кількість обчислень цільової функції, кінцевий крок, кінцеве значення функції та відповідне йому значення незалежної змінної.

6. Написати підпрограму відображення ітераційного процесу на

двовимірному графіку функції $f(x)$ відповідно до вимог:

1) вхідні параметри підпрограми: $f(x)$, $X$, $F$;

2) визначити межі інтервалу осі абсцис:

$$a = \min\{X_k\}, \qquad b = \max\{X_k\};$$

3) відобразити на екрані графік функції $f(x)$ на інтервалі $[a, b]$;

4) на побудованому графіку функції відобразити всі послідовні точки пошуку $(x_k, f_k)$, $k = 0, 1, 2, \ldots$, сполучені відрізками прямої лінії;

5) на тому ж графіку відобразити початкову та кінцеву точки пошуку маркерами, різними за формою та кольором.

7. Написати загальну програму для виконання ітераційного процесу та його відображення на екрані відповідно до вимог:

1) задати у програмі значення $x_0$, $x^*$, $\alpha$, $\varepsilon$, $N$;

2) шляхом виклику підпрограми ітераційного процесу з використанням підпрограми обчислення функції $f(x)$ та заданих значень $x_0$, $x^*$, $\alpha$, $\varepsilon$, $N$ реалізувати ітераційний процес з побудовою таблиці процесу на екрані та отриманням масивів значень змінної $X$ і функції $F$;

3) за допомогою підпрограми графічного відображення ітераційного процесу для мінімізації функції $f(x)$ за масивами $X$ і $F$ відобразити графік процесу на екрані комп'ютера.

8. Запустити загальну програму ітераційного процесу та виконати аналіз отриманих результатів:

1) на підставі табличного та графічного уявлення ітераційного процесу зробити висновок про його збіжність до точки мінімуму функції;

2) встановити, чи є послідовність значень функції монотонно спадною послідовністю;

3) зробити висновок про правильність роботи загальної програми та всіх підпрограм, які вона використовує.

9. Відповісти на запитання викладача за темою лабораторної роботи.

10. Оформити та здати звіт про проведену лабораторну роботу.

З м і с т  з в і т у:

1) титульний аркуш встановленого зразка із зазначенням міністерства та навчального закладу, назви навчальної дисципліни, теми роботи, номера варіанта, прізвища та ініціалів виконавця та приймаючого, міста, року;

2) вступ з коротким описом предмета методів оптимізації та метою лабораторної роботи;

3) опис цільової функції, значення початкової точки пошуку, аналітичне рішення задачі мінімізації цільової функції, точка мінімуму та мінімальне значення цільової функції;

4) теоретичний опис ітераційного процесу мінімізації цільової функції;

5) роздруківки використовуваних комп'ютерних програм;

6) табличне надання ітераційного процесу оптимізації та його результати – кількість обчислень цільової функції, кінцевий крок, кінцеве значення функції та відповідне йому значення незалежної змінної;

7) двовимірний графік ітераційного процесу з усіма точками пошуку;

8) аналіз одержаних результатів;

9) висновки про виконану лабораторну роботу.

**Контрольні запитання**

1. Що означає слово «оптимальний»?

2. Чому неприпустимі вирази «найоптимальніший», «оптимальніше», «оптимальніший»?

3. Що розуміють під оптимізацією у широкому сенсі?

4. Що розуміють під оптимізацією у вузькому сенсі?

5. Які методи оптимізації найбільш повно використовують можливості сучасної обчислювальної техніки?

6. Хто вперше ввів у математиці термін «оптимізація»?

7. Наведіть синоніми поняття методів оптимізації.

8. Чому необхідно вивчати багато методів оптимізації, а не один

метод?

9. Що становить предмет навчальної дисципліни «Методи оптимізації»?

10. Дайте визначення параметрів оптимізації та вектора змінних параметрів.

11. Що таке простір параметрів і яка його розмірність?

12. Дайте визначення допустимої області.

13. Що становить критерій якості?

14. Дайте визначення цільової функції.

15. У чому полягає задача оптимізації?

16. Дайте визначення розв'язання задач оптимізації.

17. Що таке метод оптимізації?

18. Наведіть приклади задач оптимізації.

19. У чому полягає корисність розв'язання задач оптимізації?

20. Які основні ознаки класифікації методів оптимізації?

21. Наведіть класифікацію методів оптимізації за застосовуваними засобами.

22. Наведіть класифікацію методів оптимізації за математичним формулюванням задач оптимізації.

23. Наведіть класифікацію методів оптимізації за порядком використовуваних похідних цільової функції.

# РОЗДІЛ 2. ОСНОВИ ОДНОВИМІРНОЇ ОПТИМІЗАЦІЇ

Одновимірна оптимізація використовується в більшості методів багатовимірної оптимізації та багато в чому визначає їх практичну ефективність. У даному розділі розглядаються питання, пов'язані з основними положеннями теорії одновимірної оптимізації. Даються визначення мінімуму, максимуму та екстремуму функції однієї змінної, наводяться властивості екстремуму. Формулюються і доводяться необхідні та достатні умови екстремуму функції однієї змінної. Вводиться поняття унімодальної функції, розглядаються її властивості, даються поняття опуклої та увігнутої функцій. Вводиться поняття інтервалу невизначеності точки мінімуму унімодальної функції, наводяться основні принципи пошуку інтервалу невизначеності та його зменшення, правила виключення інтервалів. Обґрунтовується метод Свенна для пошуку інтервалу невизначеності та надається алгоритм методу. Надається метод адаптації кроку, який не використовує інтервал невизначенсті та дозволяє обчислити мінімум функції однієї змінної. Наводиться опис лабораторної роботи з дослідження методів Свенна і адаптації кроку.

## 2.1. Екстремум функції однієї змінної

Методи одновимірної оптимізації призначені для обчислення екстремуму цільової функції $f(x)$ однієї змінної $x \in R$.

Якщо у деякій точці $x^*$ функція $f(x)$ набуває значення $f^* = f(x^*)$ , найменше в порівнянні з її значеннями в околиці цієї точки $C_\varepsilon(x^*) = (x^* - \varepsilon, x^* + \varepsilon)$, де $\varepsilon > 0$, то точка $x^*$ називається *точкою мінімуму* функції $f(x)$, а число $f^*$ називають просто *мінімумом* функції (рис. 2.1).

Якщо ж для всіх $x \in C_\varepsilon(x^*)$ і $x \neq x^*$ буде $f(x) > f(x^*)$, то в точці

$x^*$ функція $f(x)$ має *максимум* $f^*$. Мінімум та максимум мають загальну назву *екстремум*.

Для дослідження функції $f(x)$ на екстремум використовують поняття похідної цієї функції

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

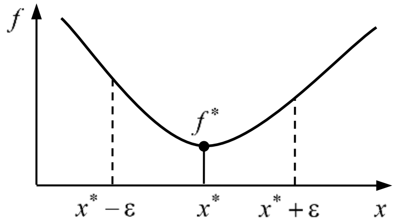


Рис. 2.1. Мінімум функції

Перша похідна на підставі цього визначення обчислюється за формулою кінцевої різниці

$$f'_x = \frac{f(x+\delta) - f_x}{\delta}, \qquad (2.1)$$

де $x$ – точка, у якій обчислюється похідна; $\delta$ – малий приріст аргументу; $f_x$ – значення функції в точці $x$. Формула (2.1) потребує одного додаткового обчислення значення функції. Похибка цієї формули становить $o(\delta)$. Диференційована функція безперервна, її графік гладкий без зломів і розривів. Тому диференційована функція називається *гладкою функцією*. Основана на визначенні похідної корисна властивість екстремуму дається теоремою Ферма.

Теорема Ферма. *Якщо функція $f(x)$ в деякій внутрішній точці $x^*$ інтервалу $X$ набуває в цьому інтервалі найменшого або найбільшого значення і диференційована в точці $x^*$, то $f'(x^*) = 0$.*

Доведення. Нехай для певності функція $f(x)$ в точці $x^*$ набуває найменшого в інтервалі $X$ значення. Візьмемо такий приріст $\Delta x$, що $x^* + \Delta x \in X$. Складемо відношення

$$d(\Delta x) = \frac{f(x^* + \Delta x) - f(x^*)}{\Delta x}.$$

Його чисельник завжди додатний. Тому якщо $\Delta x > 0$, то $d(\Delta x) > 0$, а значить, і $\lim\limits_{\Delta x \to +0} d(\Delta x) \geq 0$, тобто $f'(x^* + 0) \geq 0$. Якщо ж $\Delta x < 0$, то $d(\Delta x) < 0$, отже, $\lim\limits_{\Delta x \to -0} d(\Delta x) \leq 0$, тобто $f'(x^* - 0) \leq 0$. Оскільки функція $f(x)$ диференційована в точці $x^*$, з нерівностей $f'(x^* + 0) \geq 0$ і $f'(x^* - 0) \leq 0$ випливає, що $f'(x^* + 0) = f'(x^* - 0) = 0$, тобто $f'(x^*) = 0$, що і потрібно було довести.

Геометричний зміст теореми Ферма полягає в тому, що дотична до графіка функції, проведена в точці екстремуму, паралельна осі абсцис, що показано на рис. 2.2.

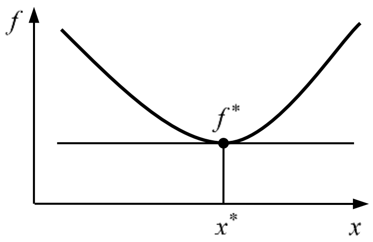



Рис. 2.2. Зміст теореми Ферма

Точки $x$, у яких виконується рівність $f'(x) = 0$, називаються *стаціонарними точками*.

Функція, що має кілька мінімумів чи максимумів, називається *багатоекстремальною*. Найменший з усіх мінімумів багатоекстремальної функції називається *глобальним мінімумом*, інші ж мінімуми називаються *локальними* (рис. 2.3). При цьому глобальний мінімум можна визначити шляхом знаходження всіх локальних мінімумів та вибору найменшого з них. Аналогічно визначаються *глобальний* та *локальні*

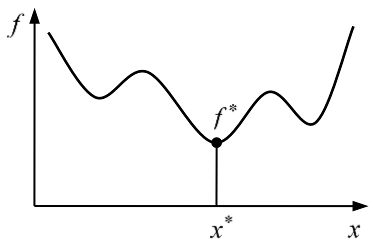



Рис. 2.3. Глобальний мінімум

*максимуми*. Методи одновимірної оптимізації, що розглядаються в даному посібнику, призначені для обчислення локального мінімуму цільової функції.

### 2.2. Умови екстремуму функції однієї змінної

Необхідна ознака існування екстремуму функції однієї змінної $f(x)$ в точці $x^*$ надається наступною теоремою.

Теорема 1. *Якщо функція $f(x)$ має у точці $x^*$ екстремум, то її похідна у цій точці або дорівнює нулю, або не існує.*

Доведення. Припустимо, що функція $f(x)$ має у точці $x^*$ екстремум. Тоді існує такий інтервал, що містить у собі точку $x^*$ і $f^* = f(x^*)$ є найбільшим або найменшим значенням у цьому інтервалі. При цьому можливі два випадки.

1. Функція $f(x)$ диференційована у точці $x^*$. Тоді за теоремою Ферма $f'(x^*) = 0$ (рис. 2.2).

2. Функція $f(x)$ не диференційована у точці $x^*$. У цьому випадку $f'(x^*)$ не існує (рис. 2.4).

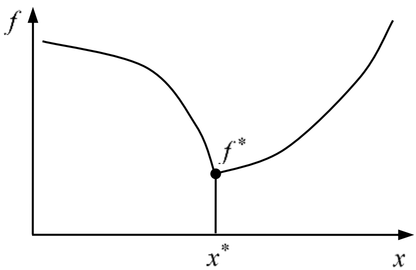


Рис. 2.4. Функція не диференційована

Достатньою умовою ця ознака не є, тому що рівність нулю похідної або її відсутність у цій точці ще не означає, що ця точка є екстремальною.

Необхідною і достатньою ознакою існування екстремуму функції $f(x)$ в точці $x^*$ є зміна знака похідної $f'(x)$ при переході через точку $x^*$ (рис. 2.5). Якщо при зростанні змінної $x$ похідна $f'(x)$ в точці $x^*$ змінює знак з мінусу на плюс, то це точка мінімуму функції (рис. 2.5, *а*). Якщо ж знак $f'(x)$ у точці $x^*$ змінюється з плюсу на мінус, це точка максимуму (рис. 2.5, *б*). Тому має місце така теорема.

Т е о р е м а 2 . *Для того щоб диференційована функція $f(x)$, мала в точці $x^*$ екстремум, необхідно і достатньо, щоб її похідна $f'(x)$ при переході через цю точку змінювала знак.*

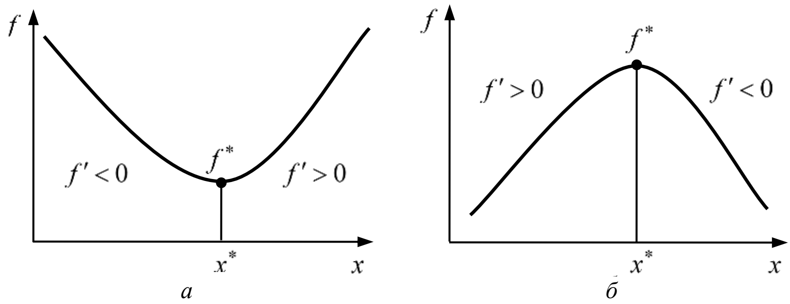



Рис. 2.5. Необхідна та достатня ознака екстремуму

Якщо ж $f'(x^*) = 0$ або $f'(x^*)$ не існує, але при переході через точку $x^*$ знак $f'(x)$ не змінюється, то екстремум у точці $x^*$ не існує. Прикладом є точка перегину графіка функції з горизонтальною дотичною (рис. 2.6).

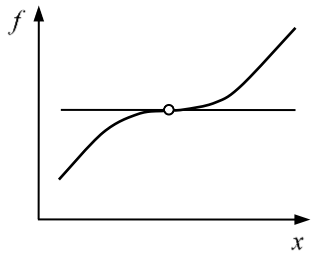



Рис. 2.6. Точка перегину

Дослідити функцію на екстремум можна і за допомогою другої похідної.

Т е о р е м а  3 . *Нехай у точці* $x^*$ *функція* $f(x)$ *двічі диференційована та* $f'(x^*) = 0$, $f''(x^*) \neq 0$. *Тоді функція* $f(x)$ *в точці* $x^*$ *має екстремум: мінімум, якщо* $f''(x^*) > 0$*, і максимум, якщо* $f''(x^*) < 0$.

Д о в е д е н н я . Нехай $f''(x^*) > 0$. Це означає, що

$$\lim_{x \to x^*} \frac{f'(x) - f'(x^*)}{x - x^*} > 0 .$$

А оскільки $f'(x^*) = 0$, то

$$\lim_{x \to x^*} \frac{f'(x)}{x - x^*} > 0 .$$

Звідси випливає, що при $x$ досить близьких до $x^*$ також буде

$$\frac{f'(x)}{x - x^*} > 0 .$$

Нехай $x < x^*$, тоді $x - x^* < 0$, а значить, і $f'(x) < 0$. Якщо ж $x > x^*$, то $x - x^* > 0$, тоді й $f'(x) > 0$. Отже, при переході через точку $x^*$ похідна $f'(x)$ змінює знак з мінусу на плюс, тому функція $f(x)$ має у цій точці мінімум. Аналогічно доводиться існування максимуму.

Доведена теорема 3 є більш «вузькою», ніж попередня теорема 2, оскільки у точці екстремуму навіть $f'(x)$ може не існувати, а тоді тим більше не існує $f''(x)$. У подібних випадках можна застосовувати тільки теорему 2, яка може бути застосована і в тому випадку, коли $f'(x^*) = 0$ і $f''(x^*) = 0$.

Якщо одночасно потрібно визначити першу і другу похідні функції, то вони обчислюються за формулами кінцевих різниць:

$$f'_x = \frac{f(x+\delta) - f(x-\delta)}{2\delta}, \qquad f''_x = \frac{f(x+\delta) - 2f_x + f(\delta - h)}{\delta^2}, \qquad (2.2)$$

де $x$ – точка, у якій обчислюються похідні; $\delta$ – малий приріст аргументу; $f_x$ – значення функції в точці $x$. Формули (2.2) потребують двох додаткових обчислень значення функції. Похибки цих формул становлять $o(\delta^2)$.

Всі наведені умови існування екстремумів прадиві як для локальних екстремумів, так і для глобальних. Тому вони не дозволяють відрізнити глобальний екстремум від локального екстремуму.

## 2.3. Унімодальні функції та їх властивості

Існують функції, для яких умови оптимальності дають можливість визначити глобальний екстремум. Розглянемо низку визначень.

Функція $f(x)$ називається *строго спадноою* в інтервалі $X \subseteq R$, якщо для будь-яких $u, v \in E$, $u < v$ виконується нерівність $f(u) > f(v)$ (рис. 2.7, *а*). Для диференційованої строго спадної функції $f'(x) < 0$.

Функція $f(x)$ називається *строго зростаючою* в інтервалі $X \subseteq R$, якщо для будь-яких $u, v \in E$, $u < v$ виконується нерівність $f(u) < f(v)$ (рис. 2.7, *б*). Для диференційованої строго зростаючої функції $f'(x) > 0$.

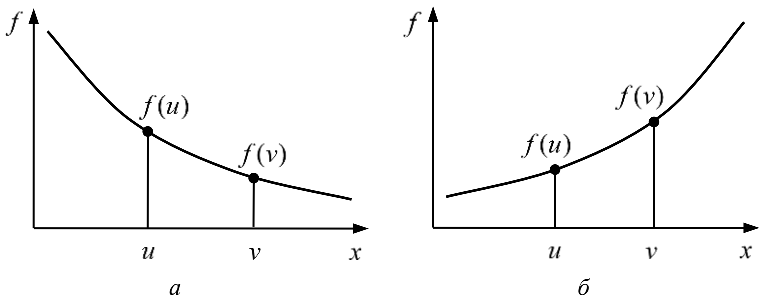



Рис. 2.7. Строго спадна і строго зростаюча функції

Функція $f(x)$ називається *унімодальною*, якщо існує єдина точка її мінімуму $x^*$ і зліва від $x^*$ ця функція строго зменшується, а праворуч від $x^*$ строго зростає (рис. 2.8, *а, б*). Це означає, що для будь-яких $u, v \in R$, $u < v < x^*$ виконується нерівність $f(u) > f(v)$, а будь-яких $u, v \in R$, таких що $x^* < u < v$, виконується нерівність $f(u) < f(v)$. Для диференційованої унімодальної функції $f'(x) < 0$ для будь-якого $x < x^*$ та $f'(x) > 0$ за будь-якого $x > x^*$ (рис. 2.5, *а*).

Унімодальна функція не обов'язково безперервна (рис. 2.8, *б*).

Функція називається *опуклою* в інтервалі $X \subseteq R$, якщо для всіх $u, v \in X$ і $\lambda \in [0;1]$ виконується нерівність

$$f(\lambda u + (1-\lambda)v) \le \lambda f(u) + (1-\lambda) f(v) . \quad (2.3)$$

Якщо ж ця нерівність є строгою для всіх $u, v \in R$, $u \neq v$ і $\lambda \in (0;1)$, то функція $f(x)$ називається *строго опуклою*.

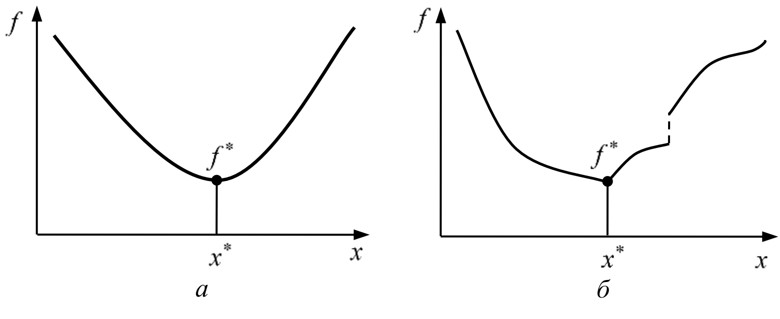
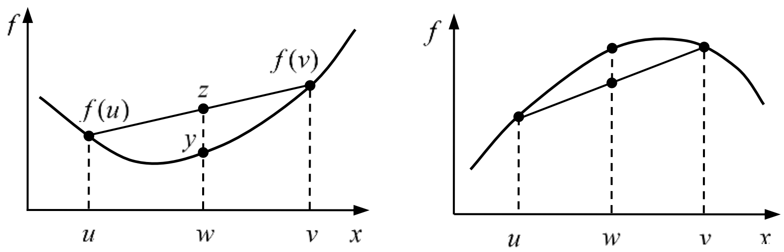



Рис. 2.8. Унімодальні функції

Візьмемо фіксовані $u$, $v$, $\lambda \in (0;1)$ і визначимо $w = \lambda u + (1-\lambda)v$, $y = f(w)$, $z = \lambda f(u) + (1-\lambda) f(v)$. Точка $(w, y)$ лежить на графіку функції, а точка $(w, z)$ – на графіку січної (рис. 2.9, *а*). Нерівність (2.3) набуде вигляду $y \le z$. Таким чином, геометрично опуклість функції $f(x)$ означає, що будь-яка точка довільної хорди графіка $f(x)$ розташовується не нижче за відповідну точку самого графіка. Це також означає, що будь-яка дотична до графіка функції $f(x)$ розташована не вище графіка самої функції. Якщо функція $f(x)$ двічі диференційована і строго опукла, то $f''(x) > 0$.

Функція $f(x)$ називається (строго) *увігнутою* або *угнутою*, якщо $-f(x)$ (строго) опукла (рис. 2.9, *б*). Для диференційованої строго увігнутої функції $f''(x) < 0$.

Якщо строго опукла функція має мінімум, цей мінімум єдиний. Така функція є унімодальною. Однак визначенню унімодальної функції можуть задовольняти і функції, які не є безперервними і опуклими.




*а* *б*

Рис. 2.9. Строго опукла і строго увігнута функції

Якщо функція має властивість унімодальності, то локальний мінімум є і глобальним мінімумом. Якщо ж функція не є унімодальною, то можлива наявність кількох локальних мінімумів.

Майже всі методи одновимірної оптимізації, що далі розглядаються, призначені для числової мінімізації унімодальних функцій.

## 2.4. Інтервал невизначеності та виключення інтервалів

Замкнутий інтервал $[a, b] \subset R$, що містить невідому точку мінімуму $x^*$ функції $f(x)$, називається *інтервалом невизначеності*.

Отже, якщо відомо, що $x^* \in [a, b]$, але саме значення $x^*$ невідоме заздалегідь, то $[a, b]$ – інтервал невизначеності.

Одновимірний пошук мінімуму цільової функції зазвичай складається з двох основних етапів:

1) пошук інтервалу невизначеності;

2) зменшення інтервалу невизначеності.

Пошук інтервалу невизначеності полягає в тому, що з деякої початкової точки послідовно здійснюються кроки до тих пір, поки значення функції в пробних точках не перестануть зменшуватися. Якщо для унімодальної функції $f(x)$ знайдено такі три різні точки $a$, $u$ і $b$ зі значеннями функції $f_a = f(a)$, $f_u = f(u)$ і $f_b = f(b)$, що значення функції у внутрішній точці не перевищує її значень у крайніх точках, то крайні точки є межами інтервалу невизначеності. Нехай, наприклад, при $a < u < b$ виконуються нерівності $f_a > f_u$, $f_u \le f_b$. Тоді інтервали $(-\infty, a)$ і $(b, +\infty]$, в яких не може бути мінімуму унімодальної функції, виключають. Після цього залишиться інтервал невизначеності $[a, b]$, що показано на рис. 2.10. Таке правило знаходження інтервалу невизначення *називається першим правилом виключення інтервалів* або *правилом*

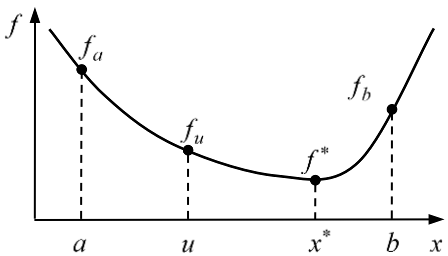


Рис. 2.10. Правило трьох точок

*трьох точок*. Зазначимо, що точка $u$ є внутрішньою точкою, яка лежить між крайніми точками $a$ і $b$, і надамо формулювання правилу трьох точок.

*Якщо відомі три точки графіка функції та значення функції у внутрішній точці не перевершує значень функції у крайніх точках, то інтервал невизначеності обмежується крайніми точками.*

Методи зменшення інтервалу невизначеності теж ґрунтуються на правилах виключення інтервалів.

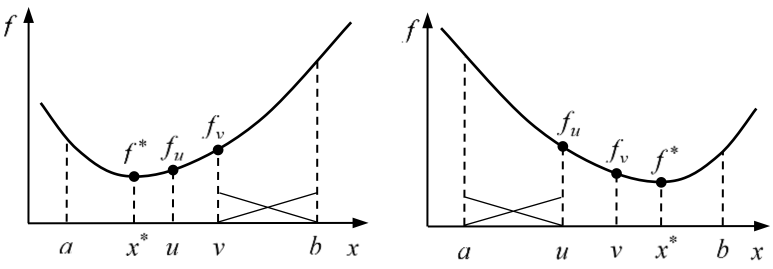
Т е о р е м а . *Нехай* $[a, b]$ *– інтервал невизначеності унімодальної функції* $f(x)$, *мінімум якої* $f^*$ *досягається у точці* $x^*$. *Нехай* $u$, $v$ *– такі внутрішні точки інтервалу невизначеності, що* $a < u < v < b$, $f_u = f(u)$ *і* $f_v = f(v)$ *– значення функції у цих точках. Тоді, якщо* $f_u < f_v$, *то* $x^* \in [a, v]$, *якщо ж* $f_u \geq f_v$, *то* $x^* \in [u, b]$.

Д о в е д е н н я . Розглянемо випадок, коли $f_u < f_v$. Припустимо, що твердження теореми неправильне, тобто $x^* \in (v, b]$. Оскільки $x^*$ – точка мінімуму унімодальної функції $f(x)$, то за визначенням лівіше точки $x^*$ функція монотонно зменшується. Це означає, що для $u < v < x^*$ повинна виконуватись нерівність $f_u > f_v$, що суперечить умові. Аналогічно доводиться твердження для випадку $f_u \geq f_v$.

Цю теорему, яка дозволяє зменшити інтервал невизначеності з використанням двох значень цільової функції, називають *другим правилом виключення інтервалів* або *правилом двох точок*. З неї випливає, що з виконанням нерівності $f_u < f_v$ напіввідкритий інтервал $(v, b]$ не містить точку мінімуму $x^*$, і його слід виключити, прийнявши за новий інтервал невизначеності інтервал $[a, v]$ меншої довжини, ніж вихідний інтервал $[a, b]$ (рис. 2.11, *а*). Якщо ж виконується протилежна нерівність $f_u \geq f_v$, то не містить точку мінімуму $x^*$ напіввідкритий інтервал $[a, u)$, і його слід виключити, прийнявши за новий інтервал невизначеності менший інтервал $[u, b]$ (рис. 2.11, *б*).

Для інтервалу невизначенності $[a, b]$ $u$ і $v$ є внутрішніми точками, які поділяють інтервал на три нових інтервала: лівий інтервал $[a, u)$, середній інтервал $[u, v]$ і правий інтервал $(v, b]$. Тому надамо формулювання правилу двох точок.

*Якщо відомий інтервал невизначеності та відомі дві внутрішні точки графіка функції, то для зменшення інтервалу при значенні функції лівої точки, меншому ніж значення функції правої точки, виключається правий інтервал, інакше виключається лівий інтервал.*



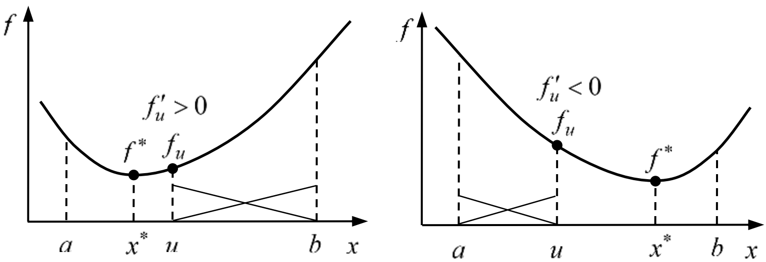

*а* *б*

Рис. 2.11. Правило двох точок

Таким чином, виключається зовнішній інтервал з боку більшого значення функції.

Зменшити інтервал невизначеності можна і з обчисленням похідної функції. Нехай $[a, b]$ – інтервал невизначеності унімодальної функції $f(x)$, мінімум якої $f^*$ досягається у точці $x^*$. І нехай $u$ – внутрішня точка інтервалу невизначеності, $u \in (a,b)$, $f_u = f(u)$ і $f_u' = f'(u)$ – значення функції та її похідної у цій точці. Тоді, якщо $f_u' > 0$, то $x^* \in [a, u]$, а якщо $f_u' < 0$, то $x^* \in [u, b]$. Якщо ж $f_u' = 0$, то $x^* = u$.

Таке правило виключення інтервалів з використанням тільки одного значення похідної цільової функції називають *третім правилом виключення інтервалів* або *правилом однієї точки*. З нього випливає, що з виконанням нерівності $f_u' > 0$ функція зростає у точці $u$, напіввідкритий інтервал $(u, b]$ не містить точку мінімуму $x^*$, і його слід виключити, прийнявши за новий інтервал невизначеності інтервал $[a, u]$ меншої довжини, ніж вихідний інтервал $[a, b]$ (рис. 2.12, *а*). Якщо ж виконується протилежна нерівність $f_u' < 0$, то не містить точку мінімуму $x^*$ напіввідкритий інтервал $[a, u)$, і його слід виключити, прийнявши за новий інтервал невизначеності менший інтервал $[u, b]$ (рис. 2.11, *б*).




*а* *б*

Рис. 2.12. Правило однієї точки

Для інтервалу невизначенності $[a, b]$ точка $u$ є внутрішньою точкою, яка поділяє інтервал на два інтервали: лівий інтервал $[a, u)$ і правий інтервал $(v, b]$. Тому надамо формулювання правилу однієї точки.

*Якщо відомий інтервал невизначеності та відоме значення похідної функції в одній внутрішній точці цього інтервалу, то при додат-*

*ному значенні похідної виключається правий інтервал, інакше виключається лівий інтервал.*

Таким чином, за правилами виключення інтервалів можна побудувати процедури пошуку точки мінімуму цільової функції шляхом обчислення значень функції або її похідних у внутрішніх точках поточного інтервалу невизначеності та подальшого виключення частин цього інтервалу. Пошук завершується, коли інтервал невизначеності зменшиться до достатньо малого розміру. На правилах виключення інтервалів ґрунтуються методи одновимірної оптимізації, які називаються *методами виключення інтервалів*.

### 2.5. Пошук інтервалу невизначеності методом Свенна

Для пошуку інтервалу невизначеності $[a, b]$ унімодальної функції $f(x)$ Свенном було запропоновано метод подвоєння кроку. Для початку роботи методу задається початкова точка $x_0$ та початковий крок $h > 0$. У точці $x_0$ обчислюється значення функції $f_0 = f(x_0)$ і виконується перехід до наступної точки $x_1 = x_0 + h$ з обчисленням значення функції $f_1 = f(x_1)$. Якщо $f_1 < f_0$, то виконуються ітерації в позитивному напрямку з точки $x_0$ за рекурентною формулою подвоєння кроку

$$x_{k+1} = x_k + 2^k h, \quad k = \overline{1, m} \tag{2.4}$$

з обчисленням значень функції $f_k = f(x_k)$ до тих пір, поки не буде отримана точка з не меншим значенням функції, ніж значення функції в попередній точці: $f_{m-1} > f_m$, $f_m \le f_{m+1}$. Тоді буде виконане правило трьох точок. Відповідний процес обчислень для значення $m = 2$ поданий на рис. 2.13.

Очевидно, що в цьому випадку будуть визначені кінці інтервалу невизначеності $a = x_{m-1}$, $b = x_{m+1}$ зі значеннями функції $f_a = f_{m-1}$, $f_b = f_{m+1}$ і внутрішня точка $u = x_m$ інтервалу зі значенням функції $f_u = f_m$. Отже, аналогічно рис. 2.10 за правилом трьох точок маємо інтервал невизначеності $[a, b]$.

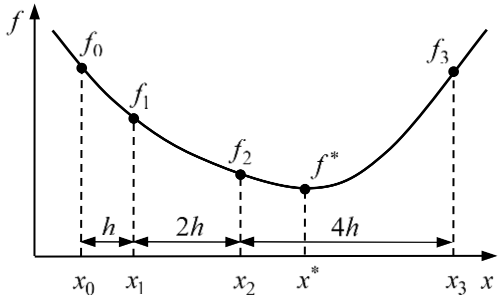



Рис. 2.13. Пошук у позитивному напрямку

Якщо після початкового кроку $f_1 \geq f_0$, то напрямок пошуку змінюється на протилежний і визначаються $x_2 = x_0 - h$, $f_2 = f(x_2)$. Якщо $f_2 \geq f_0$, то виконується правило трьох точок, інтервал невизначеності знайдено і $a = x_2$, $b = x_1$, $f_a = f_2$, $f_b = f_1$, $u = x_0$, $f_u = f_0$ (рис. 2.14).

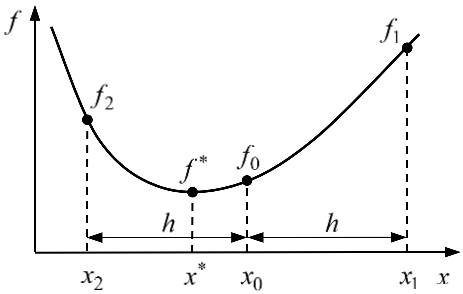



Рис. 2.14. Зміна напряму пошуку

Якщо ж $f_2 < f_0$, то ітерації виконуються у зворотному напрямку за формулою подвоєння кроку

$$x_{k+1} = x_k - 2^{k-1}h, \quad k = \overline{2, n} \tag{2.5}$$

з обчисленням значень функції $f_k = f(x_k)$ доти, доки значення функції не перестануть зменшуватися, виконується правило

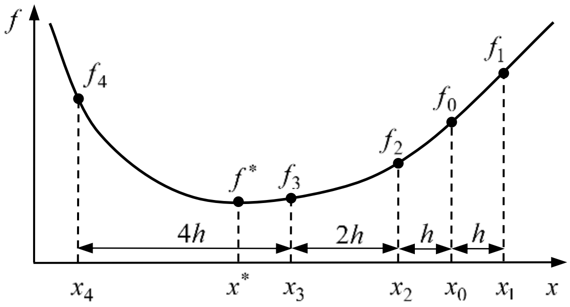



Рис. 2.15. Пошук у зворотному напрямку

трьох точок: $f_{n-1} > f_n$, а $f_n \le f_{n+1}$, що подано на рис. 2.15. Таким чином, буде визначено інтервал невизначеності зі значеннями кінців $a = x_{n+1}$, $b = x_{n-1}$, $f_a = f_{n+1}$, $f_b = f_{n-1}$ і внутрішня точка інтервалу $u = x_n$ зі значенням функції $f_u = f_n$.

У результаті роботи методу Свенна встановлюються значення меж інтервалу невизначеності. У будь-якому випадку інтервал невизначеності $[a, b]$ визначається трьома останніми точками пошуку методу Свенна. При цьому стануть відомими значення функції на кінцях інтервалу невизначеності $f_a = f(a)$ і $f_b = f(b)$, а також внутрішня точка інтервалу $u$ та значення функції у ній $f_u = f(u)$ (рис. 2.10).

Метод Свенна з використанням формул (2.4) і (2.5) можна реалізувати за наступним алгоритмом.

Алгоритм 1 методу Свенна.

Вхідні параметри: $u$ і $h$ – початкові точка і крок пошуку.

Вихідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях інтервалу; $u$ і $f_u$ – внутрішня точка інтервалу та значення у ній функції.

1. Обчислити $f_u = f(u)$, $b = u + h$, $f_b = f(b)$.

2. Якщо $f_b \ge f_u$, то перейти до кроку 6.

3. Покласти $a = u$, $f_a = f_u$, $u = b$, $f_u = f_b$, $h = 2h$.

4. Обчислити $b = u + h$, $f_b = f(b)$.

5. Якщо $f_b < f_u$, то перейти до кроку 3, інакше перейти до кроку 11.

6. Обчислити $a = u - h$, $f_a = f(a)$.

7. Якщо $f_a \geq f_u$, то перейти до кроку 11.

8. Покласти $b = u$, $f_b = f_u$, $u = a$, $f_u = f_a$, $h = 2h$.

9. Обчислити $a = u - h$, $f_a = f(a)$.

10. Якщо $f_a < f_u$, то перейти до кроку 8.

11. Вийти.

На кроках 1–5 виконується пошук інтервалу невизначеності у позитивному напрямку, а на кроках 6–10 – у зворотному напрямку, тобто в алгоритмі 1 використовуються два цикли.

Метод Свенна також можна реалізувати з одним циклом за наступним алгоритмом.

Алгоритм 2 методу Свенна.

Вхідні параметри: $u$ і $h$ – початкові точка і крок пошуку.

Вихідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях інтервалу; $u$ і $f_u$ – внутрішня точка інтервалу та значення у ній функції.

1. Обчислити $f_u = f(u)$, $b = u + h$, $f_b = f(b)$.

2. Якщо $f_b < f_u$, то покласти $a = u$, $f_a = f_u$, $u = b$, $f_u = f_b$, $h = 2h$, інакше покласти $a = b$, $f_a = f_b$, $h = -h$.

3. Обчислити $b = u + h$, $f_b = f(b)$.

4. Якщо $f_b < f_u$, то покласти $a = u$, $f_a = f_u$, $u = b$, $f_u = f_b$, $h = 2h$ і перейти до кроку 3.

5. Якщо $a > b$, то покласти $v = a$, $f_v = f_a$, $a = b$, $f_a = f_b$, $b = v$, $f_b = f_v$.

6. Вийти.

На кроках 1–2 обчислюються дві перші точки пошуку та ініціюються параметри ітераційного циклу, на кроках 3–4 виконується ітераційний пошук інтервалу невизначеності, а на кроці 5 коректуються кінці інтервалу невизначеності. У алгоритмі 2 використовується один цикл, цей алгоритм простіше алгоритму 1.

Метод Свенна – це евристичний алгоритм, оснований на простій ідеї подвоєння кроеу пошуку. У цьому методі для пошуку інтервалу невизначеності використовуються лише значення цільової функції, тому це метод нульового порядку.

## 2.6. Метод адаптації кроку

Метод адаптації кроку, розроблений харківським вченим В. Ф. Коропом, також є евристичним і використовує ідею подвоєння кроку для знаходження точки мінімуму унімодальної функції, але цей метод не використовує інтервал невизначеності. Метод полягає у множенні поточного кроку пошуку на деякий коефіцієнт, який обчислюється на підставі інформації попереднього пошуку.

Для початку роботи методу адаптації кроку задаються початкова точка $x_0$, початковий крок $h_0 > 0$ та початкове значення множника кроку $r_0 = 0$. У точці $x_0$ обчислюється значення цільової функції $f_0 = f(x_0)$.

Ітерація з номером $k+1$ методу адаптації кроку виконується за результатами ітерації з номером $k = 0, 1, 2, \ldots$ . На початку ітерації обчислюється така пробна точка пошуку:

$$y_{k+1} = x_k + h_k, \qquad g_{k+1} = f(y_{k+1}). \tag{2.6}$$

Потім визначається найкраща точка пошуку шляхом порівняння нового значення цільової функції $g_{k+1}$ з найкращим значенням попередніх ітерацій $f_k$ :

$$x_{k+1} = \begin{cases} y_{k+1}, & g_{k+1} < f_k; \\ x_k, & g_{k+1} \geq f_k; \end{cases} \qquad f_{k+1} = \begin{cases} g_{k+1}, & g_{k+1} < f_k; \\ f_k, & g_{k+1} \geq f_k. \end{cases} \tag{2.7}$$

На підставі порівняння $g_{k+1}$ з $f_k$ і попереднього значення множника кроку $r_k$ обчислюється нове значення множника кроку

$$r_{k+1} = \begin{cases} 2, & g_{k+1} < f_k \wedge r_k \geq 0{,}5; \\ 0{,}5, & g_{k+1} < f_k \wedge r_k < 0{,}5; \\ 0{,}25, & g_{k+1} \geq f_k \wedge r_k = 2; \\ -0{,}5, & g_{k+1} \geq f_k \wedge r_k \neq 2. \end{cases} \tag{2.8}$$

Помноженням цього значення на величину попереднього кроку обчислюється значення наступного кроку

$$h_{k+1} = r_{k+1} h_k \,. \tag{2.9}$$

Ітерації продовжуються доти, доки виконується умова $\mid h_{k+1} \mid > \varepsilon$, де $\varepsilon$ – допустима похибка.

Розглянемо особливості процесу одновимірного пошуку для методу адаптації кроку за формулами (2.6)–(2.9). Нехай у першій ітерації з точки $x_0$ зі значенням функції $f_0$ виконується початковий крок з переходом до пробної точки $y_1 = x_0 + h_0$ і обчисленням значення функції $g_1 = f(y_1)$. Якщо $g_1 < f_0$, то покладають $x_1 = y_1$, $f_1 = g_1$, множнику адаптації кроку задають значення $r_1 = 0{,}5$, обчислюють наступний крок $h_1 = r_1 h_0 = 0{,}5 h_0$ і переходять до наступної пробної точки $y_2 = x_1 + h_1$ з обчисленням значення функції $g_2 = f(x_2)$.

Якщо значення функції продовжують зменшуватися, то виконуються ітерації в позитивному напрямку за рекурентною формулою подвоєння кроку

$$r_k = 2\,, \quad h_k = r_k h_{k-1}, \quad y_{k+1} = x_k + h_k\,, \quad k = \overline{3, m} \tag{2.10}$$

з обчисленням значень функції $g_k = f(y_k)$ та виконанням рівностей $x_k = y_k$, $f_k = g_k$ доти, поки не буде отримана точка $y_{k+1}$ з не меншим значенням функції $g_{k+1} = f(y_{k+1})$, ніж значення функції в попередній точці: $f_{m-1} > f_m$, $f_m \le g_{m+1}$. Цей процес обчислень для $m = 3$ поданий на рис. 2.16, де $h_2 = r_2 h_1 = h_0$, $h_3 = r_3 h_2 = 2h_0$.

В цьому випадку дії методу адаптації кроку схожі на дії методу Свенна за винятком більш обережного кроку $h_1 = h_0/2$, що підтверджується порівнянням формул (2.4) та (2.10), а також рис. 2.13 та рис. 2.16:

$$x_{k+1} = x_k + 2^{k-2} h_0, \quad k = \overline{1, m-1}.$$

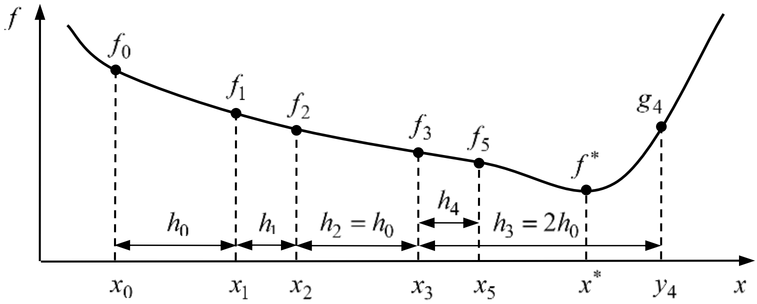



Рис. 2.16. Пошук методом адаптації кроку у позитивному напрямку

Далі відповідно до формул (2.7) і (2.8) покладають $x_{m+1} = x_m$, $f_{m+1} = f_m$, $r_{m+1} = 0{,}25$, обчислюють значення наступного кроку $h_{m+1} = r_{m+1} h_m$ і виконують наступний крок: $y_{m+2} = x_{m+1} + h_{m+1}$, $g_{m+2} = f(x_{m+2})$. Якщо $g_{m+2} < f_{m+1}$, то вважають $x_{m+2} = y_{m+2}$, $f_{m+2} = g_{m+2}$, $r_{m+2} = 0{,}5$ і обчислюють $h_{m+2} = r_{m+2} h_{m+1}$ (рис. 2.16).

Якщо ж після початкового кроку $g_1 \geq f_0$, то відповідно до формули (2.7) в ітерації з номером $k=1$ покладають $x_1 = x_0$, $f_1 = f_0$, множнику адаптації кроку за формулою (2.8) задають значення $r_1 = -0{,}5$ і обчислюють $h_1 = r_1 h_0 = -0{,}5 h_0$, $y_2 = x_1 + h_1 = x_0 - 0{,}5 h_0$, $g_2 = f(y_2)$, що подано на рис. 2.17.

Якщо і в наступних ітераціях з номерами $k = 2, 3, \ldots$ значення цільової функції не зменшуються в порівнянні зі значенням $f_0$, тобто виконується умова $g_k = f(y_k) \geq f_{k-1} = f_0$, то покладають

$$r_k = -0{,}5, \qquad h_k = r_k h_{k-1}, \qquad y_{k+1} = x_k + h_k = x_0 + h_k .$$

Таким чином, у цьому випадку на кожній ітерації напрямок пошуку змінюється на протилежний (рис. 2.17).

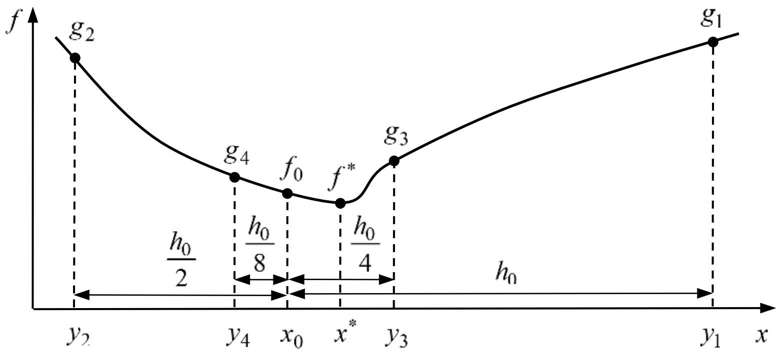



Рис. 2.17. Зміна напряму пошуку методом адаптації кроку

Якщо на першій ітерації $g_1 \geq f_0$, то за формулами (2.7) та (2.8) покладають $x_1 = x_0$, $f_1 = f_0$, $r_1 = -0{,}5$, $h_1 = r_1 h_0 = -0{,}5 h_0$. На другій ітерації обчислюються $y_2 = x_1 + h_1 = x_0 - 0{,}5 h_0$, $g_2 = f(y_2)$ і, якщо $g_2 < f_0$, то покладають $x_2 = y_2$, $f_2 = g_2$, множнику адаптації кроку за формулою

(2.8) задають значення $r_2 = 0{,}5$ і обчислюють $h_2 = r_2 h_1 = -0{,}25 h_0$.

Якщо значення функції продовжують зменшуватися, то виконуються ітерації в протилежному напрямку за рекурентною формулою подвоєння кроку (3.32) з обчисленням значень функції $g_k = f(y_k)$ і виконанням рівностей $x_k = y_k$, $f_k = g_k$, доки не буде отримана точка $y_{k+1}$ з не меншим значенням функції $g_{k+1} = f(y_{k+1})$, ніж значення функції у попередній точці: $f_{m-1} > f_m$, $f_m \le g_{m+1}$. Цей процес обчислень для $m = 4$ представлений на рис. 2.18, де $h_3 = r_3 h_2 = -0{,}5 h_0$, $h_4 = r_4 h_3 = -h_0$.

У цьому випадку дії методу адаптації кроку також схожі на дії методу Свенна за винятком більш обережних трьох перших кроків у негативному напрямку $h_1 = -h_0/2$, $h_2 = -h_0/4$ і $h_3 = -h_0/2$, що підтверджується порівнянням формул (2.5) і (2.9), а також рис. 2.15 та рис. 2.18:

$$x_{k+1} = x_k - 2^{k-4} h_0, \quad k = \overline{2, m-1}.$$

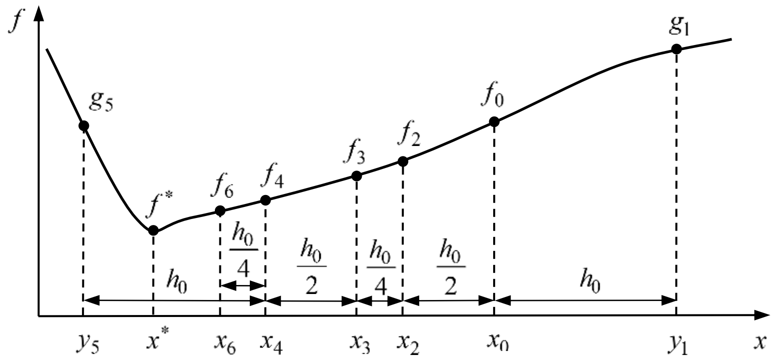


Рис. 2.18. Пошук методом адаптації кроку в негативному напрямку

Далі відповідно до формул (3.29) і (3.30) покладають $x_{m+1} = x_m$,

$f_{m+1} = f_m$, $r_{m+1} = 0{,}25$, обчислюють значення наступного кроку $h_{m+1} = r_{m+1}h_m$ і виконують наступний крок: $y_{m+2} = x_{m+1} + h_{m+1}$, $g_{m+2} = f(x_{m+2})$. Якщо $g_{m+2} < f_{m+1}$, то покладають $x_{m+2} = y_{m+2}$, $f_{m+2} = g_{m+2}$, $r_{m+2} = 0{,}5$ і обчислюють $h_{m+2} = r_{m+2}h_{m+1}$ (рис. 2.18).

У будь-якому випадку найкращі точки пошуку $x_k$ будуть наближатися до оптимальної точки $x^*$ доти, доки виконується умова $|\, h_{k+1} \,| > \varepsilon$. Таким чином, метод адаптації кроку оснований як на ідеї подвоєння кроку в методі Свенна при вдалому напрямку пошуку, так і на ідеї поділу кроку на два або чотири, що дозволяє швидко наближатися до оптимальної точки. Якщо ж задати початкове значення $r_0 \geq 0{,}5$, то подібність методу адаптації кроку до методу Свенна збільшиться.

А л г о р и т м   м е т о д у   а д а п т а ц і ї   к р о к у.

Вхідні параметри: $x$ – початкова точка пошуку; $h$ – початковий крок; $\varepsilon$ – припустима похибка.

Вихідні параметри: $x$ і $f_x$ – найкраща точка і значення функції в ній.

1. Покласти $r = 0$ і обчислити $f_x = f(x)$.

2. Обчислити $y = x + h$, $f_y = f(y)$.

3. Якщо $f_y < f_x$, то покласти $x = y$, $f_x = f_y$ і якщо $r \geq 0{,}5$, то покласти $r = 2$, інакше, покласти $r = 0{,}5$, інакше, якщо $r = 2$, то покласти $r = 0{,}25$, інакше покласти $r = -0{,}5$.

4. Покласти $h = h \cdot r$.

5. Якщо $|h| > \varepsilon$, то перейти до кроку 2.

6. Вийти.

Порівнюючи алгоритм методу адаптації кроку з алгоритмом методу Свенна, який тільки визначає інтервал невизначеності, можна зробити висновок, що алгоритм методу адаптації кроку не складніший за алгоритм методу Свенна. У методі адаптації кроку застосовуються лише прості операції додавання, подвоєння кроку, поділу кроку на два і на чотири. Тому даний метод є надійним і ефективним на практиці.

**Лабораторна робота**

Тема лабораторної роботи.

Розробка підпрограм методів Свенна та адаптації кроку.

Мета лабораторної роботи.

Розробити підпрограми методу Свенна для пошуку інтервалу невизначеності унімодальної цільової функції та методу адаптації кроку для мінімізації функції.

Порядок виконання лабораторної роботи.

1. Отримати у викладача варіант унімодальної цільової функції $f(x)$ однієї змінної $x \in R$ з початковою точкою пошуку мінімуму $x_0 \in R$.

2. Аналітично знайти точку $x^* \in R$ мінімуму заданої функції $f(x)$ та обчислити мінімальне значення функції $f^* = f(x^*)$.

3. Написати підпрограму обчислення значень функції $f(x)$ з вхідним параметром $x$ та вихідним параметром $f$.

4. Протестувати підпрограму обчислення функції $f(x)$ при обчисленні значень $f_0 = f(x_0)$ і $f^* = f(x^*)$, порівняти $f^*$ з $f_0$.

5. На підставі алгоритму методу Свенна написати підпрограму для пошуку інтервалу невизначеності точки мінімуму $x^*$ функції $f(x)$ з початкової точки $x_0$ відповідно до вимог:

1) за основу підпрограми для пошуку інтервалу невизначеності взяти підпрограму ітераційного процесу з попередньої лабораторної роботи та модифікувати цю підпрограму згідно з алгоритмом методу Свенна;

2) вхідні параметри підпрограми: $f(x)$ – цільова функція, $x_0$ – початкова точка пошуку, $h$ – початковий крок, $N$ – граничне число точок;

3) вихідні параметри підпрограми: $X$ – масив усіх точок пошуку $x_k$; $F$ – масив відповідних значень функції $f_k = f(x_k)$; $a$ і $b$ – межі інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях інтервалу; $u$ і $f_u$ – внутрішня точка інтервалу та значення в ній функції;

4) на кожній ітерації виводити на екран комп'ютера рядок, що містить номер точки пошуку $k$, крок пошуку $h_k$, значення функції $f_k$ та значення змінної $x_k$;

5) ітерації продовжуються доти, доки виконується нерівність $k \le N$;

6) після закінчення процесу оптимізації на екрані повинна відображатися таблиця, що має відповідний заголовок і представляє процес мінімізації функції;

7) під таблицею необхідно надати кількість обчислень цільової функції, кінцевий крок, кінцеве значення функції і відповідне значення незалежної змінної, а також значення вихідних параметрів $a$, $b$, $f_a$, $f_b$, $u$, $f_u$.

6. Написати загальну програму для виконання обчислювального процесу пошуку інтервалу невизначеності та відображення на екрані цього процесу відповідно до вимог:

1) за основу загальної програми для пошуку інтервалу невизначеності взяти загальну програму ітераційного процесу з попередньої лабораторної роботи та модифікувати цю програму;

2) задати у програмі значення $x_0$, $x^*$, $h$, $\varepsilon$, $N$;

3) шляхом виклику підпрограми методу Свенна з використанням підпрограми обчислення функції $f(x)$ та значень $x_0$, $h$, $N$ реалізувати ітераційний процес пошуку інтервалу невизначеності з побудовою таблиці процесу на екрані та отриманням масивів значень змінної $X$ та функції $F$, меж інтервалу невизначеності $a$ і $b$, значень функції на кінця $f_a$ і $f_b$, внутрішньої точки інтервалу $u$ та значення у ній функції $f_u$;

4) за допомогою підпрограми графічного відображення ітераційного процесу, розробленої на попередній лабораторній роботі, за масивами $X$ і $F$ відобразити графік процесу пошуку інтервалу невизначеності на екрані комп'ютера.

7. Запустити на рахунок загальну програму пошуку інтервалу невизначеності з декількох початкових точок

$$x_{01} < x^*, \quad x_{02} = x^*, \quad x_{03} > x^*$$

при кількох значеннях кроку

$$h_{01} < \left| x_0 - x^* \right|, \quad h_{02} = \left| x_0 - x^* \right|, \quad h_{03} > \left| x_0 - x^* \right|.$$

8. Виконати аналіз отриманих результатів:

1) на підставі табличної та графічної інформації про роботу методу пошуку інтервалу невизначеності визначити кінцеві точки методу Свенна $(a,\ f_a)$, $(b,\ f_b)$, $(u,\ f_u)$ і знайдений цим методом інтервал невизначеності $[a, b]$;

2) на підставі табличного та графічного надання процесу пошуку інтервалу невизначеності зробити висновок про правильне знаходження меж інтервалу;

3) зробити висновок про правильність роботи загальної програми та всіх використовуваних нею підпрограм.

9. На підставі алгоритму методу адаптації кроку написати підпрограму для мінімізації функції $f(x)$ з початкової точки $x_0$ і виконати тестування підпрограми аналогічно тестуванню підпрограми методу Свенна.

10. Відповісти на запитання викладача на тему лабораторної роботи.

11. Оформити та здати звіт про проведену лабораторну роботу.

З м і с т  з в і т у:

1) титульний аркуш встановленого зразка із зазначенням організації, назви навчальної дисципліни, теми роботи, номера варіанта, виконавця та приймаючого, міста, року;

2) вступ з коротким загальним описом призначення методів одновимірного пошуку та метою лабораторної роботи;

3) цільова функція, початкова точка пошуку, аналітичне рішення задачі мінімізації функції, точка мінімуму та мінімальне значення функції;

4) теоретичний опис методів Свенна та адаптації кроку;

5) алгоритми методів Свенна та адаптації кроку;

6) роздруківки комп’ютерних програм, що використовуються;

7) табличне надання процесів обчислень та їх результати;

8) двовимірні графіки процесів обчислень з усіма точками;

9) аналіз одержаних результатів;

10) висновки про виконану лабораторну роботу.

**Контрольні запитання**

1. Для розв’язання якої задачі призначені методи одновимірної оптимізації?

2. Дайте визначення мінімуму і максимуму функції однієї змінної.

3. Сформулюйте та доведіть теорему Ферма.

4. Дайте визначення стаціонарної точки.

5. Дайте визначення глобального та локального мінімумів.

6. Сформулюйте необхідну ознаку існування екстремуму функції однієї змінної.

7. Сформулюйте необхідну і достатню ознаку існування екстремуму функції однієї змінної.

8. Як дослідити функцію однієї змінної на екстремум за допомогою другої похідної цієї функції?

9. Які функції називаються строго спадними і строго зростаючими?

10. Дайте визначення унімодальної функції.

11. Які функції називаються опуклими та строго опуклими?

12. Які функції називаються увігнутими та строго увігнутими?

13. У чому полягає особливість опуклої функції, що має мінімум?

14. Дайте визначення інтервалу невизначеності.

15. З яких етапів складається більшість методів одновимірного пошуку?

16. Яку найменшу кількість точок на графіку функції треба знати, щоб визначити інтервал невизначеності?

17. Сформулюйте правило трьох точок.

18. Сформулюйте правило двох точок.

19. Сформулюйте правило однієї точки.

20. Яку найменшу кількість точок всередині інтервалу невизначеності необхідно визначити, щоб зменшити цей інтервал?

21. Сформулюйте правила виключення інтервалів.

22. У чому полягають методи виключення інтервалів?

23. Для чого призначений метод Свенна?

24. За яким принципом змінюється величина кроку у методі Свенна?

25. Як закінчуються обчислення у методі Свенна?

26. Якого порядку метод Свенна?

27. Яке призначення методу адаптації кроку?

28. Опишіть метод адаптації кроку.

29. За яким правилом змінюється крок пошуку у методі адаптації кроку?

30. Якого порядку метод адаптації кроку?

31. Які переваги та недоліки методу адаптації кроку?

# 3. МЕТОДИ ВИКЛЮЧЕННЯ ІНТЕРВАЛІВ

Методи одновимірної оптимізації, що дозволяють обчислити екстремум унімодальної функції однієї змінної з використанням правил виключення інтервалів, називаються *методами виключення інтервалів*. Цей розділ і присвячений методам виключення інтервалів унімодальної функції. Обґрунтовуються методи рівномірного пошуку, дихотомії, бісекції, поділу інтервалу навпіл. Розглядаються властивості чисел Фібоначчі, виходячи з яких конструюється метод Фібоначчі. Вивчаються властивості золотого перетину, якими обгрунтовується метод золотого перетину. Для розглянутих методів дано оцінку їх ефективності. Виконується порівняння методів зменшення інтервалу невизначеності за коефіцієнтом зменшення довжини інтервалу при заданій кількості обчислень функції. Наводяться алгоритми всіх розглянутих методів. Дається опис лабораторної роботи з вивчення методів зменшення інтервалу невизначеності точки мінімуму функції однієї змінної.

## 3.1. Метод рівномірного пошуку

Для обчислення мінімуму функції однієї змінної $f(x)$ існують пасивні та послідовні стратегії пошуку. При пасивній стратегії всі $n$ точок $x_1$, $x_2$, …, $x_n$, у яких будуть обчислені значення функції $f(x)$, визначено заздалегідь.

*Метод рівномірного пошуку* відповідає пасивній стратегії з рівномірним заповненням $n$ точками пошуку інтервалу невизначеності $[a,b]$, які ділять цей інтервал на $n+1$ частину однакової довжини. Для цього знаходиться довжина інтервалу невизначеності $L=b-a$, визначається відстань між точками $\delta = L/(n+1)$, обчислюються координати точок $x_k = a + k\delta$ для $k=\overline{1,n}$ і відповідні їм значення функції $f_k = f(x_k)$ (рис. 3.1).

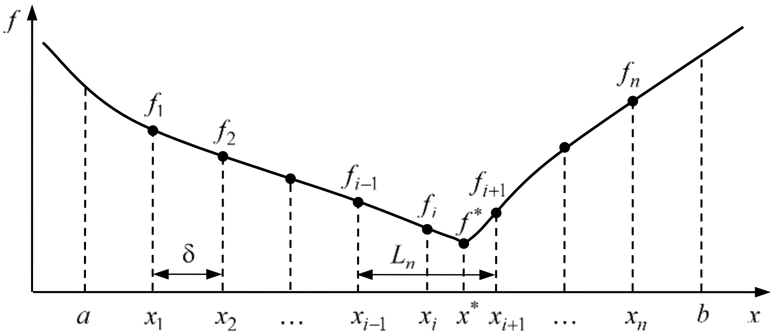



Рис. 3.1. Метод рівномірного пошуку

Далі визначається точка $x_i$ з найменшим значенням функції $f_i = \min_k f_k$. Оскільки у двох сусідніх точках відносно точки $x_i$ значення функції не менше, ніж у цій точці, то за правилом трьох точок у результаті виконання рівномірного пошуку буде отримано кінцевий інтервал невизначеності $[x_{i-1}, x_{i+1}]$ довжини $L_n = 2\delta$ (рис. 3.1). Точка мінімуму $x^*$ буде знаходитися в цьому інтервалі, похибка обчислення точки $x^*$ відносно кращої точки $x_i$ не перевищить $\delta$.

Ефективність методу рівномірного пошуку можна оцінити за допомогою відношення довжин кінцевого та початкового інтервалів невизначеності:

$$r_n = \frac{L_n}{L} = \frac{2\delta}{L} = \frac{2L}{L(n+1)}, \qquad r_n = \frac{2}{n+1}. \tag{3.1}$$

Ця величина визначає *коефіцієнт зменшення* довжини вихідного інтервалу при $n$ обчисленнях цільової функції, оскільки $L_n = r_n L$. Чим менший коефіцієнт зменшення, тим ефективніший метод оптимізації.

З'ясуємо, скільки разів необхідно обчислити значення функції, щоб зменшити інтервал невизначеності в сто разів:

$$r_n = \frac{1}{100}, \quad \frac{2}{n+1} = \frac{1}{100}, \quad n+1 = 200, \quad n = 199.$$

Якщо ж інтервал необхідно зменшити у 1000 разів, то $r_n = 0{,}001$ і $n = 1999$.

Метод рівномірного пошуку можна реалізувати за таким алгоритмом.

А л г о р и т м   м е т о д у   р і в н о м і р н о г о   п о ш у к у.

Вхідні параметри: $a$ і $b$ – межі інтервалу невизначеності; $n$ – кількість точок поділу інтервалу.

Вихідні параметри: $x$ і $f_x$ – найкраща точка і значення функції в ній.

1. Обчислити $L = b - a$, $\delta = L/(n+1)$, $x = \delta + h$, $f_x = f(x)$.
2. Покласти $k = 2$, $u = x$.
3. Покласти $u = u + \delta$ і обчислити $f_u = f(u)$.
4. Якщо $f_u < f_x$, то покласти $x = u$, $f_x = f_u$.
5. Якщо $k < n$, то покласти $k = k + 1$ і перейти до кроку 3.
6. Вийти.

Метод рівномірного пошуку є методом нульового порядку і дозволяє обчислити мінімум навіть багатоекстремальної цільової функції в заданому інтервалі з гарантованою точністю. Однак для унімодальних функцій існують набагато більш ефективні послідовні методи пошуку.

### 3.2. Метод дихотомії

На відміну від пасивного методу мінімізації функції однієї змінної $f(x)$ в послідовних методах точки пошуку вибирають послідовно з урахуванням результатів попередніх обчислень. Усі далі розглянуті методи одновимірної оптимізації є послідовними.

Розглянемо поділ інтервалу невизначеності $[a, b]$ довжини

$L = b - a$ двома внутрішніми точками $u$ і $v$, такими що $u < v$, зі значеннями функції $f_u = f(u)$, $f_v = f(x_v)$. За правилом двох точок отримаємо новий інтервал невизначеності: $[a, v]$, якщо $f_u < f_v$ (рис. 2.11, *а*); $[u, b]$, якщо $f_u \geq f_v$ (рис. 2.11, *б*). Позначимо довжини нових інтервалів $l_1 = v - a$ і $l_2 = b - u$.

Оскільки заздалегідь невідомо, $f_u < f_v$ або $f_u \geq f_v$, то за двох обчисленнях функції $n = 2$ довжина $L_2$ нового інтервалу невизначеності задовольнятиме умові $L_2 \leq \max\{l_1, l_2\}$. Величина $L_2$ не залежатиме від результату порівняння значень функції у внутрішніх точках $u$ і $v$, якщо $l_1 = l_2$. Таким чином, точки $u$ і $v$ доцільно розташовувати симетрично відносно центру $c = (a + b)/2$ вихідного інтервалу $[a,b]$.

Нехай $\delta$ – відстань точок $u$ і $v$ від центру інтервалу $c$. Тоді ці точки $u = c - \delta$, $v = c + \delta$, довжина нового інтервалу $L_2 = L/2 + \delta$ (рис. 3.2). При цьому чим менше значення $\delta$, тим менша довжина нового інтервалу невизначеності. Тому $\delta$ беруть досить малим, так що можна вважати $L_2 = L/2$, тобто інтервал невизначеності зменшується вдвічі. Залежно від результату порівняння $f_u$ і $f_v$ за наступний інтервал невизначеності беруть $[a,v]$ або $[u,b]$, і до нового інтервалу застосовують аналогічні дії. У цьому і полягає *метод дихотомії*. Термін *дихотомія* походить від грецького слова dichotomia – поділ надвоє. Ітерації продовжуються доти, доки інтервал невизначеності не зменшиться до заданого значення. Похибка

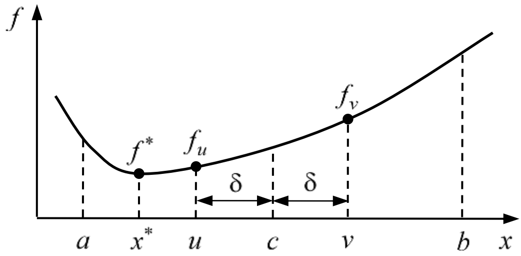


Рис. 3.2. Метод дихотомії

обчислення точки мінімуму $x^*$ не перевищить половини довжини кінцевого інтервалу невизначеності.

На кожній ітерації методу дихотомії з номером $k$ обчислюються два значення функції, тому кількість обчислень функції $n = 2k$ – парне число. Після першої ітерації довжина нового інтервалу невизначеності $L_2 = L/2 + \delta$. Після другої ітерації отримаємо

$$L_4 = \frac{L_2}{2} + \delta = \frac{L}{4} + \frac{3\delta}{2}.$$

Після третьої ітерації

$$L_6 = \frac{L_4}{2} + \delta = \frac{L_2}{4} + \frac{3\delta}{2} = \frac{L}{8} + \frac{7\delta}{4}.$$

Після ітерації з номером $k$

$$L_{2k} = \frac{L}{2^k} + \frac{2^k - 1}{2^{k-1}}\delta.$$

У цій рівності другий доданок правої частини не перевищує $2\delta$ і прямує до цього значення при $k \to \infty$. Оскільки $k = n/2$, останню рівність перепишемо у вигляді

$$L_n = \frac{L}{2^{n/2}} + \frac{2^{n/2} - 1}{2^{n/2-1}}\delta.$$

При знехтуваному малому значенні $\delta$ в результаті $n$ обчислень функції отримаємо інтервал довжини

$$L_n = \frac{L}{2^{n/2}}.$$

Ефективність методу дихотомії визначається аналогічно методу рівномірного пошуку шляхом обчислення коефіцієнта зменшення інтервалу як відношення довжин кінцевого і початкового інтервалів невизначеності:

$$r_n = \frac{L_n}{L}, \qquad r_n = \frac{1}{2^{n/2}}. \tag{3.2}$$

Нехай, наприклад, інтервал невизначеності необхідно зменшити не менш ніж у сто разів:

$$r_n \le \frac{1}{100}, \quad \frac{1}{2^{n/2}} \le \frac{1}{100}, \quad 2^{n/2} \ge 100, \quad n \ge 2\log_2 100 = 4\log_2 10 = 13{,}3.$$

Це означає, що можна покласти $n = 14$, тобто $k = n/2 = 7$ – в семи ітераціях функцію достатньо обчислити 14 разів.

Якщо ж інтервал невизначеності необхідно зменшити щонайменше ніж у 1000 разів, то $n \ge 2\log_2 1000 = 19{,}9$ і можна покласти $n = 20$.

Отже, послідовний метод дихотомії набагато ефективніший за пасивний метод рівномірного пошуку.

Недоліком методу дихотомії є необхідність обчислення двох значень функції на кожній ітерації та залежність надійності методу від значення параметра. При надто малому значенні $\delta$ та помилках обчислення значень функції інтервал невизначеності може бути втрачений. При надто великому значенні $\delta$ ефективність методу дихотомії суттєво знижується.

А л г о р и т м  м е т о д у  д и х о т о м і ї.

Вхідні параметри: $a$ и $b$ – кінці інтервалу невизначеності; $\delta$ – відстань внутрішніх точок від центру інтервалу; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – найкраща точка і значення функції в ній.

1. Обчислити $c = (a+b)/2$.

2. Обчислити $u = c - \delta$, $f_u = f(u)$, $v = c + \delta$, $f_v = f(v)$.

3. Якщо $f_u < f_v$, то покласти $b = v$, інакше, покласти $a = u$.

4. Якщо $v - a > \varepsilon$, то перейти до кроку 1.

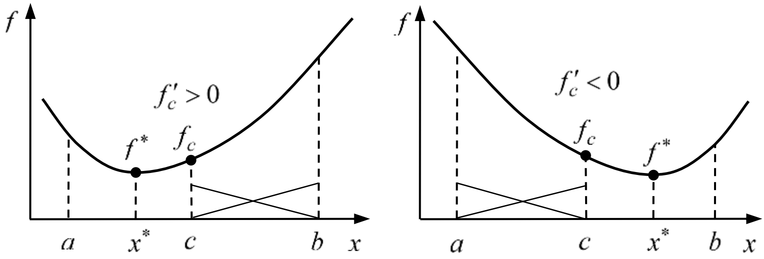
5. Вийти.

У цьому алгоритмі між його вхідними параметрами мають виконуватися співвідношення: $a < b$, $\delta < \varepsilon/4$.

Метод дихотомії – це метод нульового порядку.

### 3.3. Метод бісекції

*Метод бісекції* для зменшення інтервалу невизначеності $[a, b]$ функції $f(x)$ шляхом виключення інтервалів використовує правило однієї точки з обчисленням похідної цільової функції в середній точці інтервалу. У цьому методі обчислюються середня точка інтервалу, а також значення функції та її похідної в середній точці: $c = (a+b)/2$, $f_c = f(c)$, $f_c' = f'(c)$. Якщо $f_c' = 0$, то знайдено точку мінімуму $x^* = c$, і обчислення припиняються. Якщо $f_c' > 0$, то виключимо інтервал $[a, c)$, покладаючи $b = c$ (рис 3.3, *а*). Якщо ж $f_c' < 0$, то виключимо інтервал $(c, b]$, покладаючи $a = c$ (рис 3.3, *б*).




*а* *б*

Рис. 3.3. Метод бісекції

Для нового інтервалу невизначеності повторюються ті ж дії. Обчислення тривають, поки довжина інтервалу невизначеності $L = b - a$ перевищує допустиму похибку $\varepsilon$: $L > \varepsilon$.

Значення похідної цільової функції обчислюється за формулою

$$f'_x = \frac{f(x+\delta) - f_x}{\delta},$$

де $x$ – точка, у якій обчислюється похідна; $\delta$ – малий приріст аргументу; $f_x$ – значення функції в точці $x$. Оскільки при числовому обчисленні похідної використовується одне додаткове обчислення значення функції, то на кожній ітерації значення функції обчислюються двічі, коефіцієнт зменшення інтервалу невизначеності визначається формулою $r_n = 1/2^{n/2}$, що відповідає ефективності методу дихотомії.

За наведеним описом методу бісекції складемо алгоритм.

Алгоритм методу бісекції.

Вхідні параметри: $f$ – цільова функція; $a$ і $b$ – кінці інтервалу невизначеності; $\varepsilon$ – допустима похибка.

Вихідні параметри: $c$ і $f_c$ – кінцева точка та значення функції у ній.

1. Обчислити $c = 0{,}5 \cdot (a+b)$, $f_c = f(c)$, $f'_c = f'(c)$.
2. Якщо $f'_c = 0$, то перейти до кроку 5.
3. Якщо $f'_c < 0$, то покласти $a = c$, інакше покласти $b = c$.
4. Якщо $b - a > \varepsilon$, то перейти до кроку 1.
5. Вийти.

Метод бісекції – це метод першого порядку. Він використовується для підвищення надійності методів одновимірної оптимізації з використанням похідних цільової функції.

### 3.4. Метод поділу інтервалу навпіл

Метод поділу інтервалу навпіл застосовується для зменшення інтервалу невизначеності точки мінімуму $x^*$ цільової функції $f(x)$ і є удосконаленням методу дихотомії з метою виключення недоліків цього методу. На першій ітерації поділимо інтервал невизначеності $[a,b]$ довжини $L = b - a$ навпіл точкою $c = (a+b)/2$ і обчислимо в ній значення цільової функції $f_c = f(c)$. В результаті отримаємо два менші інтервали: $[a,c]$ і $[c,b]$ (рис. 3.4).

Розділимо лівий інтервал $[a,c]$ навпіл точкою $u = (a+c)/2$ і обчислимо в ній значення цільової функції $f_u = f(u)$. Якщо $f_u < f_c$, то за правилом двох точок виключимо правий інтервал $(c,b]$ і отримаємо новий інтервал невизначеності $[a,c]$ з відомою середньою точкою $(u, f_u)$ (рис. 3.4).

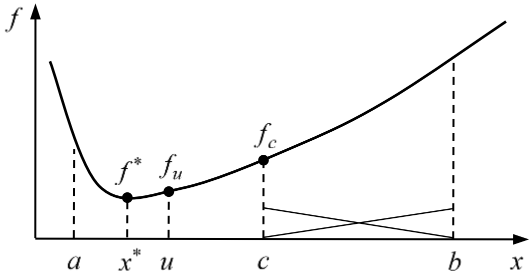


Рис. 3.4. Виключення правого интервалу

Якщо ж $f_u \geq f_c$, то розділимо правий інтервал $[c,b]$навпіл точкою $v = (c+b)/2$ і обчислимо в ній значення цільової функції $f_v = f(v)$.

Якщо $f_v < f_c$, то за правилом двох точок виключимо лівий інтервал $[a,c)$ і отримаємо новий інтервал невизначеності $[c,b]$ з відомою середньою точкою $(v, f_v)$ (рис. 3.5).

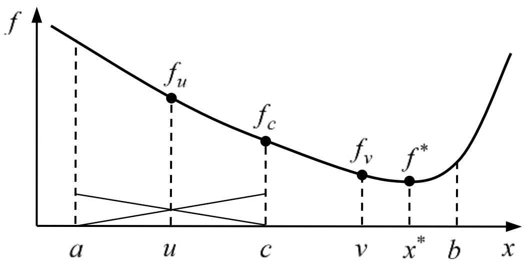


Рис. 3.5. Виключення лівого интервалу

Якщо ж $f_v \geq f_c$, то виключаються крайні інтервали $[a,u)$ і $(v,b]$, а інтервал $[u,v]$ стає новим інтервалом невизначеності з відомою середньою точкою $(c, f_c)$ (рис. 3.6).

У будь-якому випадку після двох або трьох обчислень цільової функції буде знайдено новий інтервал невизначеності, удвічі менший за вихідний інтервал, з відомою середньою точкою. У наступних ітераціях, які виконуються аналогічно першій, але без обчислення середньої точки, інтервал невизначеності буде ділитися навпіл вже після одного або двох обчислень цільової функції. Ітерації продовжуються доти, доки довжина інтервалу невизначеності більше заданого значення $\varepsilon$. Похибка обчислення точки мінімуму $x^*$ не перевищить половини довжини кінцевого інтервалу невизначеності.

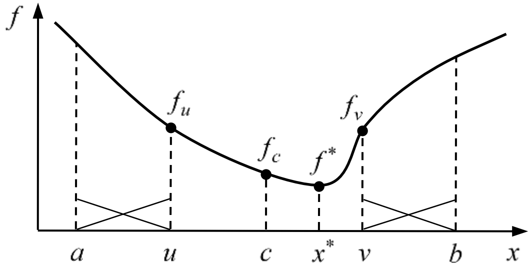


Рис. 3.6. Виключення крайніх интервалів

Якщо на кожній ітерації інтервалом невизначеності стає ліва половина попереднього інтервалу невизначеності, то в результаті $n$ обчислень функції отримаємо кінцевий інтервал невизначеності довжини

$$L_n = \frac{L}{2^{n-1}}.$$

Якщо ж на кожній ітерації інтервалом невизначеності стає не ліва половина попереднього інтервалу невизначеності, то в результаті $n$ обчислень функції отримаємо кінцевий інтервал невизначеності довжини

$$L_n = \frac{L}{2^{(n-1)/2}}.$$

У загальному випадку

$$\frac{L}{2^{n-1}} \le L_n \le \frac{L}{2^{(n-1)/2}}.$$

Ефективність методу поділу інтервалу навпіл визначається аналогічно ефективності методу дихотомії за коефіцієнтом зменшення:

$$r_n = \frac{L_n}{L}, \qquad \frac{1}{2^{n-1}} \le r_n \le \frac{1}{2^{(n-1)/2}}. \tag{3.3}$$

Нехай, наприклад, інтервал невизначеності необхідно зменшити не менш як у сто разів і $r_n \le 0{,}01$:

$$\frac{1}{2^{n-1}} \le \frac{1}{100}, \qquad 2^{n-1} \ge 100, \qquad n-1=7, \qquad n=8;$$

$$\frac{1}{2^{(n-1)/2}} \le \frac{1}{100}, \qquad 2^{n/2} \ge 100, \qquad \frac{n-1}{2}=7, \qquad n=15.$$

Тут використано нерівність $2^7 = 128 > 100$. Таким чином, $8 \le n \le 15$.

Якщо ж інтервал невизначеності необхідно зменшити не менше ніж у 1000 разів, то з використанням нерівності $2^{10} = 1024 > 1000$ отримаємо $11 \le n \le 21$.

Отже, метод поділу інтервалу навпіл у найгіршому випадку може витратити лише на одне обчислення функції більше, ніж метод дихотомії, а у найкращому випадку він використовує майже вдвічі менше обчислень. До того ж метод поділу інтервалу навпіл не використовує малий параметр, а в ньому застосовуються лише прості арифметичні операції додавання чисел і поділу на два. Тому даний метод надійніший і в багатьох випадках ефективніший за метод дихотомії.

Алгоритм методу поділу інтервалу навпіл.

Вхідні параметри: $a$, $b$ – кінці інтервалу невизначеності; $\varepsilon$ – допустима похибка.

Вихідні параметри: $c$ і $f_c$ – найкраща точка і значення функції в ній.

1. Обчислити $c=(a+b)/2$, $f_c = f(c)$.

2. Обчислити $u=(a+c)/2$, $f_u = f(u)$.

3. Якщо $f_u < f_c$, то покласти $b=c$, $c=u$, $f_c = f_u$, інакше, обчислити $v=(c+b)/2$, $f_v = f(v)$ і, якщо $f_v < f_c$, то покласти $a=u$, $c=v$, $f_c = f_v$, інакше покласти $a=u$, $b=v$.

4. Якщо $c-a>\varepsilon$, то перейти до кроку 2.

5. Вийти.

У цьому алгоритмі на перших двох кроках обчислюються середні точки інтервалу невизначеності та лівої половини інтервалу, а на кроці 3 – середня точка правої половини інтервалу.

Метод поділу інтервалу навпіл також називається *методом Больцано*. Це метод нульового порядку.

### 3.5. Числа Фібоначчі та їх властивості

Леонардо Пізано Фібоначчі – відомий математик епохи Середньовіччя, який у 1202 році написав книгу «Liber abacci», що містить майже всі арифметичні та алгебраїчні відомості того часу. Зокрема, за цією книгою європейці познайомилися з індуськими чи арабськими цифрами. Там же містилася знаменита «задача про розмноження кроликів».

«Нехай у огородженому місці є пара кроликів (самка та самець) у перший день січня. Ця пара кроликів народжує нову пару кроліків в перший день лютого і потім в перший день кожного наступного місяця. Кожна новонароджена пара кроликів стає зрілою вже через місяць і потім через місяць дає життя новій парі кроликів. Скільки пар кроликів буде в огородженому місці за рік, тобто через 12 місяців від початку розмноження?»

Розв’язання цієї задачі приводить до *чисел Фібоначчі*:

$$F_0=1,\ F_1=1,\ F_2=2,\ F_3=3,\ F_4=5,\ F_5=8,\ F_6=13,\ F_7=21,\ \ldots,$$

де індекс числа Фібоначчі відповідає номеру місяця, а значення числа –

кількості пар кроликів у цьому місяці. Такій послідовності чисел, починаючи з $F_2$, відповідає закономірність: *кожен член послідовності дорівнює сумі двох попередніх*. Це правило може бути записано у вигляді першої в історії математики *рекурентної формули*

$$F_n = F_{n-1} + F_{n-2}\,. \tag{3.4}$$

Числа Фібоначчі $F_n$ для номера $n$ від 0 до 51 наведені в табл. 3.1.

Таблиця 3.1 – Числа Фібоначчі

| $n$ | $F_n$ | $n$ | $F_n$ | $n$ | $F_n$ | $n$ | $F_n$ |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 13 | 377 | 26 | 196418 | 39 | 102334155 |
| 1 | 1 | 14 | 610 | 27 | 317811 | 40 | 165580141 |
| 2 | 2 | 15 | 987 | 28 | 514229 | 41 | 267914296 |
| 3 | 3 | 16 | 1597 | 29 | 832040 | 42 | 433494437 |
| 4 | 5 | 17 | 2584 | 30 | 1346269 | 43 | 701408733 |
| 5 | 8 | 18 | 4181 | 31 | 2178309 | 44 | 1134903170 |
| 6 | 13 | 19 | 6765 | 32 | 3524578 | 45 | 1836311903 |
| 7 | 21 | 20 | 10946 | 33 | 5702887 | 46 | 2971215073 |
| 8 | 34 | 21 | 17711 | 34 | 9227465 | 47 | 4807526976 |
| 9 | 55 | 22 | 28657 | 35 | 14930352 | 48 | 7778742049 |
| 10 | 89 | 23 | 46368 | 36 | 24157817 | 49 | 12586269025 |
| 11 | 144 | 24 | 75025 | 37 | 39088169 | 50 | 20365011074 |
| 12 | 233 | 25 | 121393 | 38 | 63245986 | 51 | 32951280099 |

Таким чином, за цією таблицею отримаємо розв'язання наведеної «задачі про розмноження кроликів»: через рік у огородженому місці буде $F_{13} = 377$ пар кроликів.

Числа Фібоначчі, крім опису кількісних характеристик процесу розмноження кроликів, є зручним математичним інструментом для опису різних явищ природи: процесів розвитку рослин і тварин, форм живих організмів і неживої матерії.

Числа Фібоначчі мають багато цікавих математичних властивостей. Зокрема,

$$F_1^2 - F_0F_2 = 1^2 - 1\cdot 2 = -1, \quad F_2^2 - F_1F_3 = 2^2 - 1\cdot 3 = 1,$$

$$F_3^2 - F_2F_4 = 3^2 - 2\cdot 5 = -1, \quad F_4^2 - F_3F_5 = 5^2 - 3\cdot 8 = 1, \quad \dots .$$

Методом математичної індукції доведемо правдивість загальної математичної формули

$$F_n^2 - F_{n-1}F_{n+1} = (-1)^n . \tag{3.5}$$

Нехай ця формула правильна при деякому $n = k$, тобто

$$F_k^2 - F_{k-1}F_{k+1} = (-1)^k . \tag{3.6}$$

Доведемо справедливість формули (3.5) і при $n = k+1$, тобто що

$$F_{k+1}^2 - F_kF_{k+2} = (-1)^{k+1} . \tag{3.7}$$

На підставі рівності (3.4) маємо

$$F_{k+1}^2 - F_kF_{k+2} = F_{k+1}(F_k + F_{k-1}) - F_k(F_{k+1} + F_k) .$$

Розкриваючи дужки у правій частині, отримаємо

$$F_{k+1}^2 - F_kF_{k+2} = F_{k+1}F_k + F_{k+1}F_{k-1} - F_kF_{k+1} - F_k^2 ,$$

та з урахуванням рівності (3.6)

$$F_{k+1}^2 - F_kF_{k+2} = -(F_k^2 - F_{k+1}F_{k-1}) = -(-1)^k$$

приходимо до рівності (3.7).

Доведемо тепер правдивість формули

$$F_{n-1}F_{n-2} - F_nF_{n-3} = (-1)^n . \tag{3.8}$$

На підставі рівності (3.4) маємо

$$F_{n-1}F_{n-2}-F_nF_{n-3}=F_{n-1}(F_n-F_{n-1})-F_n(F_{n-1}-F_{n-2})\ .$$

Розкриваючи дужки, отримаємо

$$F_{n-1}F_{n-2}-F_nF_{n-3}=F_{n-1}F_n-F_{n-1}^2-F_nF_{n-1}+F_nF_{n-2}\,.$$

З урахуванням рівності (3.5)

$$F_{n-1}F_{n-2}-F_nF_{n-3}=-(F_{n-1}^2-F_nF_{n-2})=-(-1)^{n-1}\,,$$

звідки і випливає формула (3.8).

Отримаємо вирази для чисел Фібоначчі у явному вигляді. Для цього шукатимемо розв'язання рекурентного рівняння (3.4) серед геометричних прогресій з $n$-м членом $x^n$

$$x^n=x^{n-1}+x^{n-2}\,.$$

Скорочуючи на $x^{n-2}$, отримаємо рівняння

$$x^2-x-1=0\,, \tag{3.9}$$

корені якого

$$x_1=\frac{1+\sqrt{5}}{2}\,,\qquad x_2=\frac{1-\sqrt{5}}{2}\,. \tag{3.10}$$

Отже, послідовності $\{x_1^n\}$ і $\{x_2^n\}$ задовольняють рекурентному рівнянню (3.4). Йому ж задовольняє і будь-яка лінійна комбінація цих послідовностей $\{c_1x_1^n+c_2x_2^n\}$. Виберемо коефіцієнти $c_1$ і $c_2$ так, щоб

$$F_n=c_1x_1^n+c_2x_2^n\,. \tag{3.11}$$

Для цього достатньо вимагати виконання двох умов – $F_0=1$ і $F_1=1$:

$$\begin{cases} c_1+c_2=1, \\ c_1x_1+c_2x_2=1. \end{cases}$$

Розв'яжемо цю систему рівнянь за правилом Крамера:

$$\Delta = \begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix} = x_2 - x_1 = \frac{1-\sqrt{5}}{2} - \frac{1+\sqrt{5}}{2} = \frac{-2\sqrt{5}}{2} = -\sqrt{5}\,,$$

$$\Delta_1 = \begin{vmatrix} 1 & 1 \\ 1 & x_2 \end{vmatrix} = x_2 - 1 = \frac{1-\sqrt{5}}{2} - 1 = \frac{-1-\sqrt{5}}{2} = -\frac{1+\sqrt{5}}{2} = -x_1\,,$$

$$\Delta_2 = \begin{vmatrix} 1 & 1 \\ x_1 & 1 \end{vmatrix} = 1 - x_1 = 1 - \frac{1+\sqrt{5}}{2} = \frac{1-\sqrt{5}}{2} = x_2\,,$$

$$c_1 = \frac{\Delta_1}{\Delta} = \frac{-x_1}{-\sqrt{5}} = \frac{x_1}{\sqrt{5}}\,, \qquad c_2 = \frac{\Delta_2}{\Delta} = \frac{x_2}{-\sqrt{5}} = -\frac{x_2}{\sqrt{5}}\,.$$

Підставляючи знайдені значення $c_1$ і $c_2$ у формулу (3.11), отримаємо

$$F_n = \frac{x_1}{\sqrt{5}} x_1^n - \frac{x_2}{\sqrt{5}} x_2^n = \frac{x_1^{n+1} - x_2^{n+1}}{\sqrt{5}}\,.$$

З урахуванням значень коренів (3.10) маємо остаточно

$$F_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}}{\sqrt{5}}\,. \tag{3.12}$$

Ця рівність називається формулою Біне. Таким чином, усі числа Фібоначчі, що за визначенням входять до множини цілих чисел, можуть бути виражені у явному вигляді через різницю степенів дробів, що включають ірраціональне число $\sqrt{5}$.

### 3.6. Метод Фібоначчі

*Метод Фібоначчі* оснований на числах Фібоначчі та застосовується для зменшення інтервалу невизначеності точки мінімуму $x^*$

цільової функції $f(x)$. При розгляді методу дихотомії для інтервалу невизначеності $[a, b]$ довжини $L = b - a$ з двома внутрішніми точками показано, що точки поділу доцільно розташовувати симетрично відносно центру інтервалу на малій відстані між ними. Але у методі дихотомії кожної ітерації значення цільової функції необхідно обчислювати у двох точках.

У методі Фібоначчі також дві перші точки поділу інтервалу невизначеності $x_1$ і $x_2$ розташуємо симетрично відносно центру інтервалу $[a, b]$. Обчислимо в них значення функції та застосуємо правило двох точок. На наступній ітерації маємо новий інтервал невизначеності з однією відомою внутрішньою точкою. Знайдемо другу внутрішню точку $x_3$, симетричну відносно центру інтервалу внутрішній точці, що залишилася. Обчислимо в точці $x_3$ значення функції та знову застосуємо правило двох точок. Отже, зберігаючи принцип симетрії внутрішніх точок, на кожній наступній ітерації будемо обчислювати тільки одне значення функції.

Нехай на зменшення інтервалу невизначеності відводиться $n$ обчислень цільової функції. Щоб забезпечити найбільше зменшення інтервалу, останню точку пошуку $x_n$, як і в методі дихотомії, розташуємо симетрично відносно внутрішньої точки $x_{n-1}$, що залишилася, на малій відстані $\delta$ від неї. Позначимо кінцеву довжину інтервалу невизначеності, отриману при $n$ обчисленнях цільової функції, через $L_n$. Закінчення процесу поділу інтервалу невизначеності показано на рис. 3.7.

Знайдемо довжини інтервалів невизначеності:

$$L_{n-1} = L_n + (L_n - \delta) = 2L_n - \delta\,, \qquad L_{n-2} = L_{n-1} + L_n = 3L_n - \delta\,,$$

$$L_{n-3} = L_{n-2} + L_{n-1} = 5L_n - 2\delta\,, \qquad L_{n-4} = L_{n-3} + L_{n-2} = 8L_n - 3\delta\,, \qquad \ldots\,.$$

Перепишемо ці рівності з використанням чисел Фібоначчі:

$$L_{n-1} = F_2 L_n - F_0 \delta\,, \qquad L_{n-2} = F_3 L_n - F_1 \delta\,,$$

$$L_{n-3} = F_4 L_n - F_2\delta\,, \qquad L_{n-4} = F_5 L_n - F_3\delta\,, \qquad \ldots\,.$$

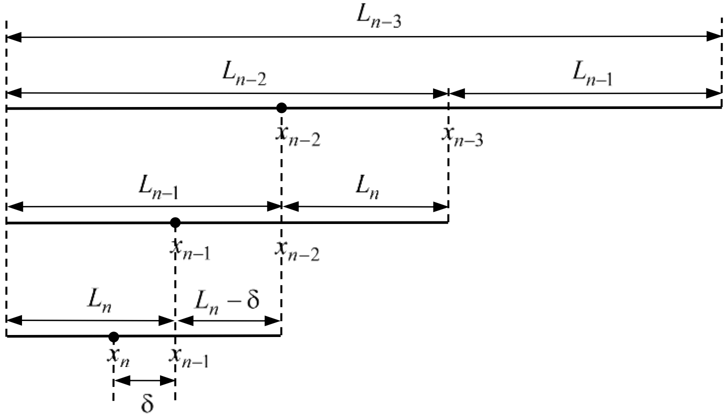



Рис. 3.7. Закінчення поділу інтервалу методом Фібоначчі

Звідси виведемо загальну формулу для довжин інтервалів:

$$L_{n-j} = F_{j+1} L_n - F_{j-1}\delta\,, \qquad j = \overline{1, n-1}\,. \tag{3.13}$$

Покладаючи $j = n-1$, знайдемо вихідну довжину інтервалу $L_1 = L = b - a$

$$L = F_n L_n - F_{n-2}\delta\,.$$

Звідси отримаємо, що при $n$ обчисленнях функції у методі Фібоначчі кінцева довжина інтервалу невизначеності дорівнюватиме

$$L_n = \frac{L}{F_n} + \frac{F_{n-2}}{F_n}\delta\,. \tag{3.14}$$

Для початку пошуку необхідно знайти положення першої точки поділу вихідного інтервалу, яка повинна лежати на відстані $L_2$ від кінця інтервалу. Покладаючи у формулі (3.13) $j = n-2$, отримаємо

$$L_2 = F_{n-1}L_n - F_{n-3}\delta .$$

Підставимо в цю формулу вираз для $L_n$ з формули (3.14):

$$L_2 = F_{n-1}\left(\frac{L}{F_n} + \frac{F_{n-2}}{F_n}\delta\right) - F_{n-3}\delta = \frac{F_{n-1}}{F_n}L + \frac{F_{n-1}F_{n-2} - F_nF_{n-3}}{F_n}\delta .$$

З урахуванням рівності (3.8) остаточно отримаємо

$$L_2 = \frac{1}{F_n}[(b-a)F_{n-1} + (-1)^n\delta] . \qquad (3.15)$$

За властивістю симетрії для двох перших точок поділу вихідного інтервалу невизначеності $[a, b]$ можна покласти:

$$x_1 = b - L_2 , \qquad x_2 = a + L_2 .$$

Звідси випливає, що

$$L_2 = b - x_1 , \qquad x_2 = a + b - x_1 .$$

Якщо на деякій ітерації відома внутрішня точка $x_k$ інтервалу $[a_k, b_k]$, то симетрична їй точка знаходиться за формулою

$$x_{k+1} = a_k + b_k - x_k . \qquad (3.16)$$

Шляхом порівняння $x_k$ і $x_{k+1}$, $f_k = f(x_k)$ і $f_{k+1} = f(x_{k+1})$ за правилом двох точок зменшується інтервал невизначеності.

При знехтувано малому значенні $\delta$ в результаті $n$ обчислень функції за формулою (3.14) отримаємо інтервал довжини

$$L_n = \frac{L}{F_n} .$$

Ефективність методу Фібоначчі визначимо коефіцієнтом зменшення:

$$r_n = \frac{L_n}{L}, \quad r_n = \frac{1}{F_n}. \tag{3.17}$$

Якщо інтервал невизначеності необхідно зменшити не менше ніж у сто разів, то за числами Фібоначчі із табл. 3.1 отримаємо:

$$r_n \le \frac{1}{100}, \quad \frac{1}{F_n} \le \frac{1}{100}, \quad F_n \ge 100, \quad F_{11} = 144 > 100, \quad n = 11.$$

Якщо ж інтервал невизначеності необхідно зменшити щонайменше у 1000 разів, то $F_n \ge 1000$, $F_{16} = 1597 > 1000$, і можна покласти $n = 16$.

Отже, метод Фібоначчі ефективніший за метод дихотомії.

Формули (3.15) і (3.16) разом із правилом вилучення інтервалів є основою алгоритму методу Фібоначчі.

А л г о р и т м  1  м е т о д у  Ф і б о н а ч ч і.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $n$ – число обчислень функції; $\delta$ – відстань між внутрішніми точками на останній ітерації.

Вихідні параметри: $x$ і $f_x$ – найкраща точка і значення функції в ній.

1. Покласти $L = b - a$, $F_0 = 1$, $F_1 = 1$, $F = F_0 + F_1$, $s = 1$, $i = 3$.

2. Покласти $F_0 = F_1$, $F_1 = F$, $F = F_0 + F_1$, $s = -s$.

3. Якщо $i < n$, то покласти $i = i + 1$ і перейти до кроку 2.

4. Обчислити $L = (L \cdot F_1 + \delta \cdot s) / F$, $u = b - L$, $f_u = f(u)$, $v = a + L$, $f_v = f(v)$ і покласти $k = 2$.

5. Обчислити $L = v - u$.

6. Якщо $f_u < f_v$, то покласти $a = u$, $u = v$, $f_u = f_v$ і обчислити $v = b - L$, $f_v = f(v)$, інакше покласти $b = v$, $v = u$, $f_v = f_u$ і обчислити $u = a + L$, $f_u = f(u)$.

7. Якщо $k < n$, то покласти $k = k + 1$ і перейти до кроку 5.

8. Вийти.

Недоліком методу Фібоначчі є необхідність завдання кількості обчислень $n$ для значень цільової функції та залежність надійності методу від значення параметра $\delta$, яке має задовольняти умові $\delta < L/F_{n+1}$. Але при надто малому значенні $\delta$ та помилках обчислення значень функції інтервал невизначеності може бути втрачений. Щоб зменшити кількість вхідних параметрів і привести алгоритм методу Фібоначчі до виду алгоритмів попередніх методів, модифікуємо попередній алгоритм.

Алгоритм 2 методу Фібоначчі.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $\delta$ – відстань між внутрішніми точками на останній ітерації; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – найкраща точка і значення функції в ній.

1. Покласти $L = b - a$, $F_0 = 1$, $F_1 = 1$, $F = F_0 + F_1$, $s = 1$.
2. Покласти $F_0 = F_1$, $F_1 = F$, $F = F_0 + F_1$, $s = -s$.
3. Якщо $L > \varepsilon \cdot F$, то перейти до кроку 2.
4. Обчислити $L = (L \cdot F_1 + \delta \cdot s)/F$, $u = b - L$, $f_u = f(u)$, $v = a + L$, $f_v = f(v)$.
5. Обчислити $L = v - u$.
6. Якщо $f_u < f_v$, то покласти $b = v$, $v = u$, $f_v = f_u$ і обчислити $u = a + L$, $f_u = f(u)$, інакше покласти $a = u$, $u = v$, $f_u = f_v$ і обчислити $v = b - L$, $f_v = f(v)$.
7. Якщо $L > \varepsilon$, то перейти до кроку 5.
8. Вийти.

Метод Фібоначчі – це метод нульового порядку. Він є найефективнішим методом одновимірної оптимізації унімодальної функції однієї змінної довільного виду.

## 3.7. Золотий перетин і його властивості

Найвідоміший математичний твір античної науки «Початки», написаний Евклідом у третьому столітті до нашої ери, містить геометричну задачу «про поділ відрізка в крайньому та середньому відношенні». Зміст цієї задачі пояснює рис. 3.8: потрібно розділити відрізок $AB$ точкою $C$ в такому відношенні, щоб довжина більшої частини відрізка $CB$ так відносилася до довжини меншої частини $AC$, як довжина всього відрізка $AB$ до довжини його більшої частини $CB$

$$\frac{AB}{CB} = \frac{CB}{AC} . \qquad (3.18)$$


Рис. 3.8. Поділ відрізка у крайньому та середньому відношенні

Позначимо відношення (3.18) через $x$. Тоді, враховуючи, що $AB = AC + CB$, відношення (3.18) подамо у вигляді

$$x = \frac{AC + CB}{CB} = 1 + \frac{AC}{CB} = 1 + \frac{1}{\frac{CB}{AC}} = 1 + \frac{1}{x} ,$$

звідки випливає алгебриїчне рівняння

$$x^2 = x + 1 . \qquad (3.19)$$

Це рівняння збігається з рівнянням (3.9), що має корені (3.10). З геометричного змісту відношення (3.18) випливає, що розв’язанням задачі є додатний корінь

$$\tau = \frac{1+\sqrt{5}}{2} = 1{,}6180339887498949\ldots \; . \qquad (3.20)$$

Леонардо да Вінчі назвав це ірраціональне число *золотим перетином*. Рівняння (3.19) називається *рівнянням золотого перетину*. На відрізку $AB$ існує ще одна точка $D$ (рис. 3.8), яка ділить його золотим перетином, оскільки

$$\frac{AB}{AD} = \frac{AD}{DB} = \frac{1+\sqrt{5}}{2}.$$

Золотий перетин широко зустрічається у геометрії. З «Початків» Евкліда відомий спосіб геометричної побудови золотого перетину з використанням лінійки та циркуля, показаний на рис. 3.9.

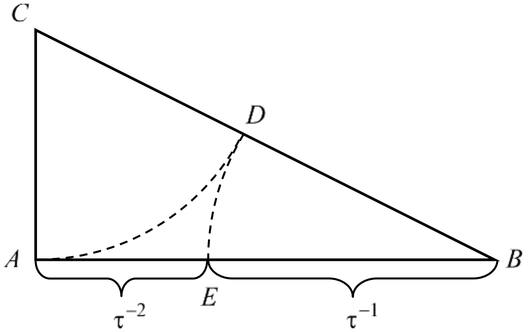



Рис. 3.9. Геометрична побудова золотого перетину

Побудуємо прямокутний трикутник $ABC$ із сторонами $AB = 1$ і $AC = 1/2$. Тоді за теоремою Піфагора

$$CB = \sqrt{1+\left(\frac{1}{2}\right)^2} = \frac{\sqrt{5}}{2}.$$

Провівши дугу кола $AD$ з центром у точці $C$ до її перетину з відрізком $CB$ у точці $D$, отримаємо відрізок

$$BD = CB - CD = \frac{\sqrt{5}-1}{2} = \tau^{-1}.$$

Провівши дугу $DB$ з центром у точці $B$ до її перетину з відрізком $AB$ у точці $E$, отримаємо поділ відрізка $AB$ золотим перетином, оскільки

$$\frac{AB}{EB}=\frac{EB}{AE}=\tau\,,\qquad AB=1=EB+AE=\tau^{-1}+\tau^{-2}\,.$$

Знаменитий астроном Йоганн Кеплер висловив своє захоплення золотим перетином: «У геометрії існує два скарби – теорема Піфагора та поділ відрізка у крайньому та середньому відношенні. Перший можна порівняти з цінністю золота, другий можна назвати дорогоцінним каменем».

Золотий перетин має цікаві алгебраїчні властивості. З рівняння золотого перетину (3.19)

$$\tau^2=\tau+1\,. \tag{3.21}$$

Ця тотожність може бути подана у вигляді

$$\tau=1+\frac{1}{\tau}\,, \tag{3.22}$$

звідки

$$\tau-1=\frac{1}{\tau}\,. \tag{3.23}$$

Щоб отримати зворотне до $\tau$ число, достатньо відняти від $\tau$ одиницю.

Якщо в праву частину рівності (3.22) замість $\tau$ підставити його значення, що задається цією ж рівністю, то отримаємо подання $\tau$ у вигляді багатоповерхового дробу

$$\tau=1+\cfrac{1}{1+\cfrac{1}{\tau}}\,.$$

Якщо продовжити таку підстановку у правій частині нескінченне число разів, то в результаті отримаємо ланцюговий дріб

$$\tau = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \ldots}}}}.$$

Повернімося до тотожності (3.21), яку подамо у формі

$$\tau = \sqrt{1 + \tau}\ . \qquad (3.24)$$

Якщо в правій частині цієї тотожності замість $\tau$ підставити його значення, що задається цією ж тотожністю, то отримаємо подання $\tau$ у вигляді

$$\tau = \sqrt{1 + \sqrt{1 + \tau}}\ .$$

Підставляючи в праву частину цієї тотожності вираз (3.24) замість $\tau$ і повторюючи цю операцію нескінченну кількість разів, отримаємо подання золотого перетину в радикалах

$$\tau = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \ldots}}}}\ .$$

Помножимо всі члени тотожності (3.21) на $\tau$, а потім розділимо їх на $\tau$ і отримаємо дві нові тотожності:

$$\tau^3 = \tau^2 + \tau, \qquad \tau = 1 + \tau^{-1}.$$

У результаті приходимо до тотожності, що зв'язує степені золотого перетину,

$$\tau^n = \tau^{n-1} + \tau^{n-2},$$

де $n = 0, \pm 1, \pm 2, \pm 3, \ldots$ . Будь-який цілий степінь золотого перетину дорівнює сумі двох попередніх.

Розглянемо послідовність степенів золотого перетину

$$\ldots,\ \tau^{-n},\ \tau^{-(n-1)},\ \ldots,\ \tau^{-2},\ \tau^{-1},\ \tau^0 = 1,\ \tau^1,\ \tau^2,\ \ldots,\ \tau^{n-1},\ \tau^n,\ \ldots,$$

названу золотою прогресією. Це єдина послідовність, яка одночасно є геометричною та арифметичною прогресією. Цілі частини додатних степенів золотого перетину $\tau^n$ для $n$ від 1 до 52 подані в табл. 3.2.

Таблиця 3.2 – Цілі частини степенів золотого перетину

| $n$ | $[\tau^n]$ | $n$ | $[\tau^n]$ | $n$ | $[\tau^n]$ | $n$ | $[\tau^n]$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 14 | 842 | 27 | 439204 | 40 | 228826127 |
| 2 | 2 | 15 | 1364 | 28 | 710646 | 41 | 370248451 |
| 3 | 4 | 16 | 2206 | 29 | 1149851 | 42 | 599074578 |
| 4 | 6 | 17 | 3571 | 30 | 1860497 | 43 | 969323029 |
| 5 | 11 | 18 | 5777 | 31 | 3010349 | 44 | 1568397607 |
| 6 | 17 | 19 | 9349 | 32 | 4870846 | 45 | 2537720636 |
| 7 | 29 | 20 | 15126 | 33 | 7881196 | 46 | 4106118243 |
| 8 | 46 | 21 | 24476 | 34 | 12752042 | 47 | 6643838879 |
| 9 | 76 | 22 | 39602 | 35 | 20633239 | 48 | 10749957122 |
| 10 | 122 | 23 | 64079 | 36 | 33385282 | 49 | 17393796001 |
| 11 | 199 | 24 | 103681 | 37 | 54018521 | 50 | 28143753123 |
| 12 | 321 | 25 | 167761 | 38 | 87403803 | 51 | 45537549124 |
| 13 | 521 | 26 | 271442 | 39 | 141422324 | 52 | 73681302247 |

Враховуючи (3.20), маємо:

$$\tau = \frac{1+\sqrt{5}}{2}, \qquad \tau^{-1} = \frac{-1+\sqrt{5}}{2}.$$

Формулу Біне (3.12) подамо у вигляді

$$F_n = \frac{\tau^{n+1} - (-\tau)^{-(n+1)}}{\sqrt{5}}. \tag{3.25}$$

Таким чином, формула Біне встановлює зв’язок між числами Фібоначчі та золотим перетином.

### 3.8. Метод золотого перетину

Недоліками розглянутого методу Фібоначчі є необхідність задання кількості обчислень $n$ для значень цільової функції $f(x)$ та залежність надійності методу від значення малого параметра $\delta$. На кожній ітерації методу Фібоначчі внутрішня точка поточного інтервалу невизначеності ділить його у відношенні двох послідовних чисел Фібоначчі. Так на початковому етапі відповідно до рівності (3.15) при незначному малому значенні $\delta$ інтервал невизначеності $[a, b]$ довжини ділиться $L = b - a$ у відношенні

$$\frac{L_2}{L} = \frac{F_{n-1}}{F_n} .$$

За формулою Біне (3.25) при $n \to \infty$ число Фібоначчі $F_n$ еквівалентне $\tau^{n+1} / \sqrt{5}$, тому

$$\lim_{n \to \infty} \frac{F_{n-1}}{F_n} = \frac{\tau^n \sqrt{5}}{\sqrt{5} \tau^{n+1}} = \frac{1}{\tau} .$$

Отже, у методі Фібоначчі при великих $n$ інтервал невизначеності ділиться по золотому перетину. Позначимо $t = 1/\tau$. Відповідно до рівності (3.22) $t = \tau - 1$ і за формулою (3.20)

$$t = \frac{\sqrt{5} - 1}{2} = 0{,}6180339887498949\ldots \quad . \tag{3.26}$$

Як і метод Фібоначчі, метод золотого перетину використовує принцип симетрії двох внутрішніх точок поділу інтервалу невизначеності. Дві перші точки визначаються за золотим перетином:

$$x_1 = b - tL , \qquad x_2 = a + tL ,$$

де $x_1 < x_2$. У цих точках обчислюються значення функції $f_1 = f(x_1)$, $f_2 = f(x_2)$ і застосовується правило виключення інтервалів, після чого інтервал невизначеності зменшується в $\tau$ разів і має довжину $L_2 = tL$. Внутрішня точка, що залишилася, ділить новий інтервал невизначеності також по золотому перетину. Наступна точка пошуку визначається симетрично точці, що залишилася за формулою (3.16).

На кожній ітерації методу золотого перетину обчислюється значення функції лише в одній внутрішній точці, а друга точка зі своїм значенням функції переходить із попередньої ітерації. Після цього застосовується правило виключення інтервалів і інтервал невизначеності зменшується в $\tau$ разів. Ітерації тривають до того часу, поки довжина інтервалу невизначеності стане менше допустимої похибки $\varepsilon$, тобто доки виконається умова $L_n < \varepsilon$.

Оцінимо ефективність методу золотого перетину. В результаті $n$ обчислень функції отримаємо інтервал довжини

$$L_n = t^{n-1}L = \frac{L}{\tau^{n-1}}.$$

Коеффіцієнт зменшення інтервалу становитиме

$$r_n = \frac{1}{\tau^{n-1}}. \tag{3.27}$$

Якщо інтервал невизначеності необхідно зменшити не менше ніж у сто разів, то на підставі степенів золотого перетину $\tau^n$ з таблиці. 3.2 отримаємо:

$$r_n \le \frac{1}{100}, \quad \frac{1}{\tau^{n-1}} \le \frac{1}{100}, \quad \tau^{n-1} \ge 100, \quad [\tau^{10}] = 122 > 100, \quad n = 11.$$

Якщо ж інтервал невизначеності необхідно зменшити щонайменше у 1000 разів, то $\tau^{n-1} \ge 1000$, $[\tau^{15}] = 1364 > 1000$, і можна покласти $n = 16$. Ці результати показують, що ефективність методу золотого перетину практично така сама, як і методу Фібоначчі. Для більш точного порівняння ефективності цих методів знайдемо межу відношення для них довжин кінцевих інтервалів невизначеності $L_n$:

$$\lim_{n\to\infty}\frac{L/\tau^{n-1}}{L/F_n}=\lim_{n\to\infty}\frac{F_n}{\tau^{n-1}}=\lim_{n\to\infty}\frac{\tau^{n+1}/\sqrt{5}}{\tau^{n-1}}=\frac{\tau^2}{\sqrt{5}}=1{,}1708\ldots\ .$$

При великих $n$ довжина інтервалу невизначеності після $n$ обчислень значень цільової функції у методі золотого перетину лише приблизно на 17 % більша, ніж у методі Фібоначчі. Однак метод золотого перетину має цю істотну перевагу, тому що в ньому обчислення можна припинити у будь-який момент, оскільки точки, в яких проводяться обчислення, не залежать від їх загального числа $n$.

А л г о р и т м   м е т о д у   з о л о т о г о   п е р е т и н у.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – кінцеві точка і значення функції.

1. Обчислити $t=0{,}5\cdot(\sqrt{5}-1)$, $L=t\cdot(b-a)$, $u=b-L$, $f_u=f(u)$, $v=a+L$, $f_v=f(v)$.

2. Обчислити $L=v-u$.

3. Якщо $f_u<f_v$, то покласти $b=v$, $v=u$, $f_v=f_u$ і обчислити $u=a+L$, $f_u=f(u)$, інакше покласти $a=u$, $u=v$, $f_u=f_v$ і обчислити $v=b-L$, $f_v=f(v)$.

4. Якщо $L>\varepsilon$, то перейти до кроку 2.

5. Вийти.

Вочевидь, що алгоритм методу золотого перетину простіший за алгоритм методу Фібоначчі. Параметр $t$ методу золотого перетину слід задавати якомога точніше за формулою (3.26), інакше в результаті похибок обчислень інтервал невизначеності може бути втрачений. Ознакою цього у наведеному алгоритмі є виконання нерівності $u\geq v$. Тоді рекомендується перейти до кроку 1 алгоритму.

Метод золотого перетину – це метод нульового порядку.

### 3.9. Порівняння методів виключення інтервалів

Ми розглянули різні методи виключення інтервалів, призначені для зменшення інтервалу невизначеності точки мінімуму $x^*$ унімодальної цільової функції однієї змінної $f(x)$ за правилами виключення інтервалів. Ці методи, що дозволяють знаходити мінімум унімодальної функції, використовують значення меж інтервалу невизначеності. За обсягом використанням похідних їх можна поділити на дві групи.

1. Методи нульового порядку, що дозволяють знаходити мінімум унімодальної функції довільного вигляду. Це методи рівномірного пошуку, дихотомії, поділу інтервалу навпіл, Фібоначчі та золотого перетину.

2. Метод першого порядку, що дозволяє знаходити мінімум диференційованої унімодальної функції. Це лише метод бісекції.

Для роботи цих методів необхідно знайти інтервал невизначеності методом Свенна. Ефективність усіх прямих методів виключення інтервалів можна оцінити за значеннями коефіцієнта зменшення довжини інтервалу невизначеності при обчисленнях значень цільової функції

$$r_n = \frac{L_n}{L} .$$

Формули цього коефіцієнта для методу рівномірного пошуку (МРП) (3.1), методів дихотомії та бісекції (МДБ) (3.2), методу поділу інтервалу навпіл (МПІП) (3.3), методу Фібоначчі (МФ) (3.17) та методу золотого перетину (МЗП) ) (3.27) для зручності порівняння зведені у табл. 3.3. За цією таблицею, табл. 3.1 і табл. 3.2 для чисел Фібоначчі та степенів золотого перетину складена табл. 3.4. В ній для тих же методів представлено кількість обчислень цільової функції, необхідне для досягнення необхідного зменшення інтервалу невизначеності, яке задано значенням коефіцієнта зменшення інтервалу $r_n$ .

Таблиця 3.3 – Коефіцієнти зменшення інтервалу

| МРП | МДБ | МПІП | МФ | МЗП |
|---|---|---|---|---|
| $r_n = \dfrac{2}{n+1}$ | $r_n = \dfrac{1}{2^{n/2}}$ | $\dfrac{1}{2^{n-1}} \le r_n \le \dfrac{1}{2^{(n-1)/2}}$ | $r_n = \dfrac{1}{F_n}$ | $r_n = \dfrac{1}{\tau^{n-1}}$ |

Таблиця 3.4 – Кількість обчислень функції

| $r_n$ | МРП | МДБ | МПІП | МФ | МЗП |
|---|---|---|---|---|---|
| $10^{-1}$ | 19 | 8 | [5; 9] | 6 | 6 |
| $10^{-2}$ | 199 | 14 | [8; 15] | 11 | 11 |
| $10^{-3}$ | 1999 | 20 | [11; 21] | 16 | 16 |
| $10^{-4}$ | 19999 | 28 | [15; 29] | 20 | 21 |
| $10^{-5}$ | 199999 | 34 | [18; 35] | 25 | 25 |
| $10^{-6}$ | 1999999 | 40 | [21; 41] | 30 | 30 |
| $10^{-7}$ | 19999999 | 48 | [25; 49] | 35 | 35 |
| $10^{-8}$ | 199999999 | 54 | [28; 55] | 39 | 40 |
| $10^{-9}$ | 1999999999 | 60 | [31; 61] | 44 | 45 |
| $10^{-10}$ | 19999999999 | 68 | [35; 69] | 49 | 49 |

З цієї таблиці видно, що теоретично найефективнішим методом зменшення інтервалу невизначеності є метод Фібоначчі. Незначно поступається йому в ефективності метод золотого перетину. Верхня оцінка кількості обчислень функції методом поділу інтервалу навпіл лише на одиницю перевищує ті ж показники для методу дихотомії, проте нижня оцінка майже вдвічі нижча за метод дихотомії і навіть менша, ніж у методу Фібоначчі. Метод рівномірного пошуку задля досягнення тієї ж точності, як і інші методи, витрачає на кілька порядків більше обчислень значень функції. Найнадійнішим з цих методів зменшення інтервалу невизначеності є метод поділу інтервалу навпіл, що не залежить від додаткових параметрів. Його надійність можна порівняти з надійністю методу адаптації кроку.

**Лабораторна робота**

Тема лабораторної роботи

Розробка підпрограми для зменшення інтервалу невизначеності.

Мета лабораторної роботи.

Розробити підпрограму методу зменшення інтервалу невизначеності унімодальної цільової функції.

Порядок виконання лабораторної роботи.

1. Отримати у викладача задача на програмування одного з методів зменшення інтервалу невизначеності:

1) метод рівномірного пошуку;

2) метод дихотомії;

3) метод поділу інтервалу навпіл;

4) метод Фібоначчі;

5) метод золотого перетину.

2. На підставі алгоритму для заданого методу зменшення інтервалу невизначеності написати комп'ютерну підпрограму відповідно до вимог:

1) вхідні параметри підпрограми: $f(x)$ – цільова функція; $N$ – гранична кількість точок; $\varepsilon$ – допустима похибка, вхідні параметри алгоритму заданого методу;

2) вихідні параметри підпрограми: $X$ – масив усіх точок пошуку $x_k$; $F$ – массив відповідних значень функції $f_k = f(x_k)$; $x$ і $f_x$ – інайкраща точка пошуку та значення у ній функції;

3) на кожній ітерації виводити на екран комп'ютера рядок, що містить номер точки пошуку $k$, довжину інтервалу невизначеності $L_k$, значення функції $f_k$ та значення змінної $x_k$;

4) ітерації продовжуються доти, доки виконуються нерівності:

$$\left| L_k \right| > \varepsilon, \qquad k \le N ;$$

5) після закінчення процесу оптимізації на екрані повинна відображатися таблиця, що має відповідний заголовок і становить процес мінімізації функції;

6) під таблицею необхідно відобразити кількість обчислень цільової функції, довжину кінцевого інтервалу невизначеності, мінімальне значення функції та відповідне йому значення незалежної змінної.

3. Написати загальну програму для виконання обчислювального процесу мінімізації цільової функції однієї змінної та відображення на екрані цього процесу відповідно до вимог:

1) за основу загальної програми для мінімізації функції взяти загальну програму пошуку інтервалу невизначеності з попередньої лабораторної роботи та модифікувати цю програму;

2) задати у програмі значення $x_0$, $x^*$, $h$, $\varepsilon$, $N$;

3) шляхом виклику підпрограми методу Свенна з використанням підпрограми обчислення функції $f(x)$ та значень $x_0$, $h$, $N$ реалізувати ітераційний процес пошуку інтервалу невизначеності з побудовою таблиці процесу на екрані та отриманням масивів значень змінної $X$ та функції $F$, меж інтервалу невизначеності $a$ і $b$;

4) шляхом виклику розробленої підпрограми зменшення інтервалу невизначеності реалізувати ітераційний процес зменшення інтервалу з побудовою таблиці процесу на екрані та отриманням масивів значень змінної $X$ та функції $F$, кращої точки пошуку $x$ та значення в ній функції $f_x$;

5) за допомогою підпрограми графічного відображення ітераційного процесу по масивах $X$ і $F$ відобразити графік процесу мінімізації цільової функції однієї змінної на екрані комп'ютера.

4. Запустити загальну програму мінімізації функції однієї змінної для функцій попередніх лабораторних робіт із кількох початкових точок

$$x_{01} < x^*, \quad x_{02} = x^*, \quad x_{03} > x^*$$

при кількох значеннях кроку

$$h_{01} < \left| x_0 - x^* \right|, \quad h_{02} = \left| x_0 - x^* \right|, \quad h_{03} > \left| x_0 - x^* \right|.$$

5. Виконати аналіз одержаних результатів:

1) на підставі табличної та графічної інформації про роботу методів одновимірного пошуку визначити кінцеві точки методу Свенна та знайдений цим методом інтервал невизначеності;

2) на кількох початкових ітераціях за табличними даними та графіками зіставити початок роботи методу зменшення інтервалу з його робочими формулами;

3) оцінити ефективність методу зменшення інтервалу та зіставити її з теоретичною ефективністю.

6. Відповісти на запитання викладача на тему лабораторної роботи.

7. Оформити та здати звіт про проведену лабораторну роботу.

Зміст звіту:

1) титульний аркуш встановленого зразка із зазначенням організації, назви навчальної дисципліни, теми роботи, номера варіанта, виконавця та приймаючого, міста, року;

2) постановка задачі мінімізації цільової функції, що включає задача функції, початкову точку і точку мінімуму;

3) теоретичний опис методу зменшення інтервалу;

4) алгоритм методу;

5) роздруківки використовуваних комп'ютерних програм;

6) табличне подання процесу мінімізації та його підсумки – кількість обчислень значень цільової функції, мінімальне значення функції та відповідне йому значення точки мінімуму, досягнуту точність;

7) двовимірний графік траєкторії пошуку з усіма точками;

8) аналіз одержаних результатів;

9) висновки про виконану лабораторну роботу.

**Контрольні запитання**

1. Чим відрізняється пасивна стратегія пошуку?

2. Яке призначення методу рівномірного пошуку?

3. Опишіть метод рівномірного пошуку.

4. Як оцінюється ефективність методу рівномірного пошуку?

5. Які переваги та недоліки методу рівномірного пошуку?

6. Чим відрізняються послідовні методи одновимірного пошуку?

7. Яке призначення методу дихотомії?

8. Які основні засади методу дихотомії?

9. Опишіть метод дихотомії.

10. Як оцінюється ефективність методу дихотомії?

11. Якого порядку є метод дихотомії?

12. У чому полягають переваги та недоліки методу дихотомії?

13. Яке призначення методу поділу інтервалу навпіл?

14. Опишіть метод розподілу інтервалу навпіл.

15. Як оцінюється ефективність методу поділу інтервалу навпіл?

16. Якого порядку метод поділу інтервалу навпіл?

17. Які переваги та недоліки методу поділу інтервалу навпіл?

18. Яка задача привела до відкриття чисел Фібоначчі?

19. За якою рекурентною формулою визначаються числа Фібоначчі?

20. Запишіть 10 перших чисел Фібоначчі.

21. Вкажіть властивості чисел Фібоначчі.

22. Виведіть формулу Біне для чисел Фібоначчі.

23. Яке призначення методу Фібоначчі?

24. Опишіть метод Фібоначчі.

25. Виведіть загальну формулу для довжин інтервалів невизначеності в методі Фібоначчі.

26. Як починається пошук мінімуму у методі Фібоначчі?

27. Як оцінюється ефективність методу Фібоначчі?

28. Якого порядку метод Фібоначчі?

29. У чому полягають переваги та недоліки методу Фібоначчі?

30. Яка задача привела до відкриття золотого перетину?

31. Виведіть рівняння золотого перетину.

32. Наведіть числову постійну золотого перетину.

33. Опишіть спосіб геометричної побудови золотого перетину з використанням лінійки та циркуля.

34. Вкажіть алгебрїчні властивості золотого перетину.

35. Яка особливість золотої прогресії?

36. Який зв'язок між числами Фібоначчі та золотим перетином?

37. Яке призначення методу золотого перетину?

38. Опишіть метод золотого перетину.

39. У чому полягає схожість та відмінність методів Фібоначчі та золотого перетину.

40. Як починається пошук мінімуму у методі золотого перетину?

41. Як оцінюється ефективність методу золотого перетину?

42. Порівняйте ефективність методів Фібоначчі та золотого перетином.

43. Якого порядку є метод золотого перетину?

44. Які переваги та недоліки методу золотого перетину?

45. На які групи можна поділити методи зменшення інтервалу невизначеності?

46. Порівняйте ефективність методів зменшення інтервалу невизначеності.

# 4. МЕТОДИ ПОЛІНОМІАЛЬНОЇ ІНТЕРПОЛЯЦІЇ

Даний розділ присвячений методам одновимірної оптимізації, основаних на правилах виключення інтервалів і на апроксимації унімодальної цільової функції поліномом. Ці методи нульового, першого та другого порядків виконують інтерполяцію точки мінімуму цільової функції в точці мінімуму поліноміальної функції. Обгрунтовуються методи квадратичної інтерполяції з трьома точками, квадратичної інтерполяції з двома точками, січних, Ньютона, кубічної інтерполяції з чотирма точками, кубічної інтерполяції з двома точками. Дається порівняння швидкості асимптотичної збіжності методів апроксимації та інтерполяції. Наводяться алгоритми всіх розглянутих методів.

## 4.1. Методи квадратичної інтерполяції

Розглянуті раніше методи одновимірного пошуку зменшують інтервал невизначеності в наперед задане число разів незалежно від виду самої цільової функції. Однак якщо функції гладкі та близькі до поліномів, то більш ефективні методи поліноміальної апроксимації та інтерполяції.

Можна показати, що будь-яка гладка функція $f(x)$ у малій околиці її точки мінімуму $x^*$ близька до квадратичної функції.

*Методи квадратичної інтерполяції* основані на апроксимації, тобто наближенні до цільової функції квадратичним поліномом

$$q(x) = \alpha x^2 + \beta x + \gamma \,, \tag{4.1}$$

де коефіцієнти $\alpha$, $\beta$ і $\gamma$ – дійсні числа. Диференціюючи рівність (4.1), маємо

$$q'(x) = 2\alpha x + \beta \,. \tag{4.2}$$

За необхідною умовою існування екстремуму $q'(x) = 0$ отримаємо точку мінімуму квадратичної функції

$$\bar{x} = -\frac{\beta}{2\alpha}. \tag{4.3}$$

Це загальна формула квадратичної інтерполяції. Для обчислення точки (4.3) досить визначити лише параметри $\alpha$ і $\beta$ рівняння (4.1). Це можна зробити різними методами з використанням інтервалу невизначеності $[a, b]$, а також значень цільової функції $f(x)$ та її похідних. Таке обчислення називається *інтерполяцією*, оскільки точка з формули (4.3) повинна належати інтервалу невизначенності. В методах квадратичної інтерполяції формула (4.3) використовується ітераційно разом з правилами виключення інтервалів для зменшення інтервалу невизначенності. Ітерації продовжуються доти, доки відстань між послідовними точками пошуку не стане меншою від заданої похибки. До відомих методів квадратичної інтерполяції належать методи квадратичної інтерполяції з трьома та двома точками, метод січних і метод Ньютона.

### 4.2. Метод квадратичної інтерполяції з трьома точками

Метод квадратичної інтерполяції з трьома точками оснований на апроксимації функції $f(x)$ за трьома її точками квадратичним поліномом (4.1) і на інтерполяції шуканої точки мінімуму функції $x^*$ точкою мінімуму цього полінома за формулою (4.3) і значенням функції.

Перші три точки графіка функції $(a,\ f_a)$, $(b,\ f_b)$, $(u,\ f_u)$ отримаємо за алгоритмом методу Свенна, де $[a, b]$ – інтервал невизначеності, $a < u < b$, $f_a = f(a)$, $f_b = f(b)$, $f_u = f(u)$. Підставляючи в рівняння параболи (4.1) координати трьох відомих точок функції, складемо систему рівнянь:

$$\begin{cases} q(a) = \alpha a^2 + \beta a + \gamma = f_a, \\ q(u) = \alpha u^2 + \beta u + \gamma = f_u, \\ q(b) = \alpha b^2 + \beta b + \gamma = f_b. \end{cases}$$

Віднімаючи з другого рівняння перше, а з третього – друге, маємо:

$$\begin{cases} \alpha(u^2 - a^2) + \beta(u - a) = f_u - f_a, \\ \alpha(b^2 - u^2) + \beta(b - u) = f_b - f_u. \end{cases}$$

Розв'яжемо отриману систему за правилом Крамера:

$$D = \begin{vmatrix} u^2 - a^2 & u - a \\ b^2 - u^2 & b - u \end{vmatrix} = (u^2 - a^2)(b - u) - (b^2 - u^2)(u - a) ,$$

$$D_\alpha = \begin{vmatrix} f_u - f_a & u - a \\ f_b - f_u & b - u \end{vmatrix} = (f_u - f_a)(b - u) - (f_b - f_u)(u - a) ,$$

$$D_\beta = \begin{vmatrix} u^2 - a^2 & f_u - f_a \\ b^2 - u^2 & f_b - f_u \end{vmatrix} = (u^2 - a^2)(f_b - f_u) - (b^2 - u^2)(f_u - f_a) ,$$

$$\alpha = \frac{D_\alpha}{D}, \qquad \beta = \frac{D_\beta}{D}.$$

Для зменшення кількості арифметичних операцій введемо позначення:

$$P = (u - a)(f_b - f_u), \qquad Q = (b - u)(f_a - f_u). \tag{4.4}$$

Тоді отримаємо вирази для визначників:

$$D_\alpha = -(P + Q), \qquad D_\beta = P(u + a) + Q(b + u).$$

Підставивши визначники у формулу (4.2) для точки мінімуму параболи, остаточно отримаємо нову точку

$$v = \frac{P(a+u)+Q(b+u)}{2(P+Q)}. \tag{4.5}$$

Це формула квадратичної інтерполяції за трьома точками. Початковий інтервал невизначеності $[a, b]$ ділиться двома точками $u$ і $v$ на три частини. За правилом двох точок на підставі порівняння значень цільової функції у внутрішніх точках інтервалу невизначеності $f_u = f(u)$ і $f_v = f(v)$ та частина, яка не містить мінімум, відкидається, й інтервал $[a, b]$ зменшується. Для інтервалу, що залишився, з однією відомою внутрішньою точкою знову застосовується квадратична інтерполяція за формулами (4.4) і (4.5). Такі ітерації продовжуються доти, доки відстань між внутрішніми точками $u$ і $v$ залишається більше допустимої помилки $\varepsilon$.

Внаслідок помилок округлення при обчисленнях може виявитися, що значення $v$, обчислене за формулою (4.5), виявиться поза інтервалом невизначеності. У такому разі необхідно передбачити вихід з циклу, незважаючи на те, що задана точність не була досягнута.

Розглянутий метод виключення інтервалів називається *методом квадратичної інтерполяції з трьома точками*. На підставі робочих формул (4.4) та (4.5) складено алгоритм методу квадратичної інтерполяції з трьома точками.

Алгоритм методу квадратичної інтерполяції з трьома точками.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях интервалу; $u$ і $f_u$ – внутрішня точка інтервалу і значення функції в ній; $\varepsilon$ – допустима похибка.

Вихідні параметри: $v$ і $f_v$ – кінцеві точка і значення функції.

1. Знайти $P = (u-a)(f_b - f_u)$, $Q = (b-u)(f_a - f_u)$, $S = P + Q$.

2. Обчислити $v = 0{,}5 \cdot [P \cdot (a+u) + Q \cdot (b+u)]/S$, $f_v = f(v)$.

3. Якщо $u > v$, то покласти $c = u$, $f_c = f_u$, $u = v$, $f_u = f_v$, $v = c$,

$f_v = f_c$.

4. Покласти $d = v - u$.

5. Якщо $f_u < f_v$, то покласти $b = v$, $f_b = f_v$, $v = u$, $f_v = f_u$, інакше покласти $a = u$, $f_a = f_u$, $u = v$, $f_u = f_v$.

6. Якщо $d > \varepsilon$, то перейти до кроку 1.

7. Вийти.

Метод квадратичної інтерполяції з трьома точками – це метод нульового порядку.

## 4.3. Метод квадратичної інтерполяції з двома точками

Метод квадратичної інтерполяції з двома точками оснований на апроксимації функції $f(x)$ за двома її точками квадратичним поліномом (4.1) і на інтерполяції шуканої точки мінімуму функції $x^*$ точкою мінімуму цього полінома за формулою (4.3) і значенням функції та її першої похідної.

Перші дві точки графіка функції $(a, f_a)$ і $(b, f_b)$ на кінцях інтервалу невизначеності $[a, b]$ знайдемо за алгоритмом методу Свенна. Обчислимо значення похідної функції одному з кінців інтервалу $f'_a = f'(a)$ чи $f'_b = f'(b)$. Підставляючи в рівняння параболи та її похідної (4.1) та (4.2) координати двох відомих точок функції та одне значення похідної, складемо системи рівнянь:

$$\begin{cases} q(a) = \alpha a^2 + \beta a + \gamma = f_a, \\ q(b) = \alpha b^2 + \beta b + \gamma = f_b, \\ q'(a) = 2\alpha a + \beta = f'_a, \end{cases} \qquad \begin{cases} q(a) = \alpha a^2 + \beta a + \gamma = f_a, \\ q(b) = \alpha b^2 + \beta b + \gamma = f_b, \\ q'(b) = 2\alpha b + \beta = f'_b. \end{cases}$$

Розв'язуючи ці системи і підставляючи знайдені значення коефіцієнтів $\alpha$ і $\beta$ в рівність (4.3), отримаємо формули для обчислення точки мінімуму апроксимуючої параболи. Розв'язуючи ліву систему лінійних алгебраїчних рівнянь, маємо:

$$\alpha(b^2-a^2)+\beta(b-a)=f_b-f_a,$$

$$\beta=f'_a-2\alpha a,$$

тобто

$$\alpha(b^2-a^2)+(f'_a-2\alpha a)(b-a)=f_b-f_a,$$

$$\alpha(b-a)^2=f_b-f_a-f'_a(b-a).$$

Позначимо довжину інтервалу $L=b-a$ і отримаємо

$$\alpha L^2=f_b-f_a-f'_aL.$$

Звідси

$$\alpha=\frac{f_b-f_a-Lf'_a}{L^2}.$$

Тоді

$$\bar{x}=-\frac{\beta}{2\alpha}=a-\frac{f'_a}{2\alpha}=a-\frac{1}{2}\frac{f'_aL^2}{f_b-f_a-Lf'_a}.$$

Таким чином отримані формули для обчислення точки мінімуму апроксимуючої параболи:

$$u=a+\frac{1}{2}\frac{f'_aL^2}{f'_aL-f_b+f_a},\qquad u=b-\frac{1}{2}\frac{f'_bL^2}{f'_bL-f_b+f_a},\tag{4.6}$$

де $L=b-a$ – довжина інтервалу невизначеності. Це формули квадратичної інтерполяції з двома точками.

Залежно від відомого значення похідної використовується одна з формул (4.6). Потім обчислимо $f_u=f(u)$ і $f'_u=f'(u)$ та примінимо правило однієї точки. Якщо $f'_u=0$, то знайдено точку мінімуму і обчислення припиняються. Якщо $f'_u<0$, то виключимо інтервал $[a,u)$, який не містить точку мінімуму, вважаючи $a=u$, $f_a=f_u$, $f'_a=f'_u$, і застосуємо першу формулу (4.6). Якщо ж $f'_u > 0$, то виключимо інтервал $(u, b]$, вважаючи $b = u$, $f_b = f_u$, $f'_b = f'_u$, і застосуємо другу формулу (4.6). Обчислення продовжуються, поки відстань між двома послідовними точками інтерполяції перевищує допустиму похибку $\varepsilon$.

Очевидно, формули (4.6) простіші, ніж формули квадратичної інтерполяції з трьома точками (4.4) та (4.5).

Метод квадратичної інтерполяції з двома точками, оснований на ітеративних формулах (4.6), є ефективним при мінімізації опуклих цільових функцій.

Алгоритм методу квадратичної інтерполяції з двома точками.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях интервалу; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – кінцеві точка і значення функції.

1. Обчислити $f'_a = f'(a)$.

2. Обчислити $L = b - a$, $d_f = f_b - f_a$.

3. Обчислити $p = f'_a \cdot L$, $u = a + 0{,}5 \cdot p \cdot L / (p - d_f)$.

4. Якщо $u < a \vee u > b$, то обчислити $u = 0{,}5 \cdot (a + b)$.

5. Обчислити $f_u = f(u)$, $f'_u = f'(u)$.

6. Якщо $f'_u = 0$, то перейти до кроку 18.

7. Покласти $v = u$.

8. Якщо $f'_u > 0$, то перейти до кроку 11.

9. Покласти $a = u$, $f_a = f_u$, $f'_a = f'_u$, $L = b - a$, $d_f = f_b - f_a$.

10. Обчислити $p = f'_a \cdot L$, $u = a + 0{,}5 \cdot p \cdot L / (p - d_f)$ і перейти до кроку 13.

11. Покласти $b = u$, $f_b = f_u$, $f'_b = f'_u$, $L = b - a$, $d_f = f_b - f_a$.

12. Обчислити $p = f'_b \cdot L$, $u = b - 0{,}5 \cdot p \cdot L / (p - d_f)$.

13. Якщо $u < a \vee u > b$, то обчислити $u = 0{,}5 \cdot (a + b)$.

14. Обчислити $h = u - v$, $f_u = f(u)$, $f'_u = f'(u)$.

15. Якщо $|h| > \varepsilon$, то перейти до кроку 6.

16. Вийти.

Метод квадратичної інтерполяції з двома точками – це метод першого порядку.

### 4.4. Метод січних

*Метод січних* також ґрунтується на наближенні функції $f(x)$ за двома її точками квадратичним поліномом (4.1) та на інтерполяції шуканої точки мінімуму функції $x^*$ точкою мінімуму цього полінома за формулою (4.3) та значенням перших похідних функції.

Перші дві точки $a$ і $b$ на кінцях інтервалу невизначеності знайдемо за алгоритмом методу Свенна. Обчислимо значення похідних в цих точках $f'_a = f'(a)$ і $f'_b = f'(b)$. У формулі кінцевої різниці при обчисленні похідної використовуються значення функції $f_a$ і $f_b$. Підставимо в рівняння похідної (4.2) ці значеня похідних і складемо систему рівнянь:

$$\begin{cases} q'(a) = 2\alpha a + \beta = f'_a, \\ q'(b) = 2\alpha b + \beta = f'_b. \end{cases}$$

Тоді

$$\alpha = \frac{f'_b - f'_a}{2(b-a)} = \frac{f'_b - f'_a}{2L},$$

де $L = b - a$ – довжина інтервалу невизначеності. За формулою (4.3)

$$\bar{x} = a - \frac{L}{f'_b - f'_a} f'_a .$$

Отримали формулу для обчислення точки мінімуму апроксимуючої параболи

$$u = a - \frac{f_a' L}{f_b' - f_a'}, \tag{4.7}$$

Це формула квадратичної інтерполяції для методу січних, вона геометрично відповідає розташуванню перетину з віссю абсцис $Ox$ січної, проведеної через дві точки графіка функції $f'(x)$, і дає наближення для розв'язання рівняння $f'(x) = 0$.

Обчислимо $f_u = f(u)$ і $f_u' = f'(u)$ і застосуємо правило однієї точки. Якщо $f_u' = 0$, то знайдена точка мінімуму $u$, і розрахунки припиняються. Якщо $f_u' < 0$, тоді виключимо інтервал $[a, u)$, який не містить точки мінімуму, вважаючи $a = u$, $f_a' = f_u'$. Якщо ж $f_u' > 0$, то виключимо інтервал $(u, b]$, вважаючи $b = u$, $f_b' = f_u'$. Обчислення тривають до тих пір, поки відстань між двома послідовними точками інтерполяції перевищує допустиму помилку $\varepsilon$.

Метод січних також називається *методом хорд*. Перевага цього методу – простота його формул. Недоліком є повільна збіжність при мінімізації неопуклих функцій.

Алгоритм методу січних.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях інтервалу; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – кінцева точка та значення функції в ній.

1. Обчислити $f_a' = f'(a)$, $f_b' = f'(b)$, $L = b - a$.

2. Покласти $h = L$, $v = a$.

3. Обчислити $u = a - f_a' \cdot L/(f_b' - f_a')$, $f_u = f(u)$, $f_u' = f'(u)$.

4. Якщо $f_u' = 0$, то перейти до кроку 9.

5. Обчислити $h = u - v$.

6. Якщо $f_u' < 0$, то покласти $a = u$, $f_a' = f_u'$, інакше покласти $b = u$ і $f_b' = f_u'$.

7. Обчислити $L = b - a$ і покласти $v = u$.

8. Якщо $|h| > \varepsilon$, то перейти до кроку 3.

9. Вийти.

У цьому алгоритмі значення похідної цільової функції обчислюються на кроках 1 і 3.

## 4.5. Метод Ньютона

*Метод Ньютона* для мінімізації унімодальної двічі диференційованої цільової функції $f(x)$ використовує значення її першої та другої похідних $f'(x)$ і $f''(x)$. Це метод другого порядку. Він оснований на апроксимації функції $f(x)$ за однією її точкою квадратичним поліномом (4.1) та на наближенні шуканої точки мінімуму функції $x^*$ точкою мінімуму цього полінома за формулою (4.3) та значенням першої та другої похідних функції.

Першу точку $a$ на кінці інтервалу невизначеності $[a, b]$ зі значенням цільової функції $f_a = f(a)$ знайдемо за алгоритмом методу Свенна. Обчислимо значення першої та другої похідних у цій точці $f'_a = f'(a)$ і $f''_a = f''(a)$. У формулах кінцевих різниць при обчисленні похідних використовується значення функції $f_a$. Диференціюванням рівності (4.2) маємо $q''(x) = 2\alpha$. Підставимо у цю рівність і в рівняння першої похідної (4.2) значення похідних цільової функції та складемо систему рівнянь:

$$\begin{cases} q'(a) = 2\alpha a + \beta = f'_a, \\ q''(a) = 2\alpha = f''_a. \end{cases}$$

Розв’яжуючи цю систему та підставляючи знайдені значення коефіцієнтів $\alpha$ і $\beta$ в рівність (4.3), отримаємо формулу для обчислення точки мінімуму апроксимуючої параболи

$$u = a - \frac{f'_a}{f''_a}. \tag{4.8}$$

Це формула квадратичної інтерполяції для методу Ньютона, геометрично вона відповідає знаходженню точки перетину з віссю абсцис $Ox$ дотичної, проведеною з точки $a$ до графіка функції $f'(x)$, тому вона називається *формулою дотичної*.

Обчислимо $f_u = f(u)$, $f'_u = f'(u)$ і $f''_u = f''(u)$. Якщо $f'_u = 0$, то знайдена точка мінімуму $u$ і розрахунки припиняються. В іншому випадку покладемо $a = u$, $f_a = f_u$, $f'_a = f'_u$, $f''_a = f''_u$ і знову застосуємо формулу (4.8). Розрахунки тривають до тих пір, поки відстань між двома послідовними точками інтерполяції більше допустимої помилки $\varepsilon$.

Метод Ньютона також називається *методом дотичних*. Перевагою цього методу є простота його формули та висока швидкість збіжності при мінімізації опуклих функцій.

Якщо функція не опукла, то точка інтерполяції $u$, обчислена за формулою (4.8), може бути поза інтервалом невизначеності, і це призведе до розбіжності процесу одновимірного пошуку. У цьому випадку застосовують метод бісекції: якщо $u \notin [a, b]$, тобто $u < a \vee u > b$, то покладемо $u = (a+b)/2$, і лише після цього обчислимо $f_u = f(u)$, $f'_u = f'(u)$ і $f''_u = f''(u)$. Якщо $f'_u = 0$, то знайдена точка мінімуму $u$ і розрахунки припиняються. Інакше використовується правило однієї точки. Якщо $f'_u < 0$, то виключимо інтервал $[a, u)$, який не містить точки мінімуму, покладаючи $a = u$. Якщо ж $f'_u > 0$, то виключимо інтервал $(u, b]$, покладаючи $b = u$. Після цього покладемо $u = u - f'_u / f''_u$. Розрахунки тривають до тих пір, поки відстань між двома послідовними точками пошуку перевищує допустиму помилку $\varepsilon$. Такий комбінований метод Ньютона та бісекції також ефективний і для неопуклих функцій.

На основі формули (4.8) складено алгоритм модифікованого методу Ньютона з методом бісекції.

А л г о р и т м  м е т о д у  Н ь ю т о н а .

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ – значення функції на кінці інтервалу; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – кінцева точка та значення функції в ній.

1. Покласти $u = a$, $v = a$.

2. Обчислити $f_a' = f'(a)$, $f_a'' = f''(a)$.

3. Обчислити $u = a - f_a' / f_a''$.

4. Якщо $u < a \vee u > b$, то обчислити $u = 0{,}5 \cdot (a + b)$.

5. Обчислити $f_u = f(u)$, $f_u' = f'(u)$, $f_u'' = f''(u)$.

6. Покласти $h = u - v$, $v = u$.

7. Якщо $f_u' = 0$, то перейти до кроку 11.

8. Якщо $f_u' < 0$, то покласти $a = u$, інакше покласти $b = u$.

9. Обчислити $u = u - f_u' / f_u''$.

10. Якщо $|h| > \varepsilon$, то перейти до кроку 4.

11. Вийти.

У цьому алгоритмі на кроках 2 та 5 обчислюються значення першої та другої похідних цільової функції за формулами кінцевих різниць. Обчислення значень першої та другої похідних доцільно виконувати за допомогою однієї процедури, яка використовує два додаткові обчислення функції.

### 4.6. Методи кубічної інтерполяції

*Методи кубічної інтерполяції* – це методи виключення інтервалів, які ґрунтуються на наближенні цільової функції кубічном поліномом:

$$p(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta , \tag{4.9}$$

де коефіцієнти $\alpha$, $\beta$, $\gamma$ і $\delta$ – дійсні числа. Диференціюючи функцію (4.9), маємо

$$p'(x) = 3\alpha x^2 + 2\beta x + \gamma \ . \tag{4.10}$$

З необхідної умови екстремуму $p'(x) = 0$ маємо квадратне рівняння

$$3\alpha x^2 + 2\beta x + \gamma = 0 \ .$$

Це рівняння має корені

$$\bar{x}_{1,2} = \frac{-\beta \pm \sqrt{\beta^2 - 3\alpha\gamma}}{3\alpha} \ .$$

Мінімуму поліному відповідає корінь, для якого

$$p''(x) = 6\alpha x + 2\beta > 0 \, .$$

Маємо

$$p''(x_{1,2}) = 6\alpha \frac{-\beta \pm \sqrt{\beta^2 - 3\alpha\gamma}}{3\alpha} + 2\beta = \pm 2\sqrt{\beta^2 - 3\alpha\gamma} \ .$$

Умові мінімуму задовольняє корінь

$$\bar{x} = \frac{\sqrt{\beta^2 - 3\alpha\gamma} - \beta}{3\alpha} \ . \tag{4.11}$$

Однак ця формула втрачає сенс, якщо внаслідок апроксимації виявиться $\alpha = 0$. В цьому випадку

$$p(x) = \beta x^2 + \gamma x + \delta \, , \qquad p'(x) = 2\beta x + \gamma \, , \qquad \bar{x} = \frac{-\gamma}{2\beta} \ .$$

Для того щоб врахувати обидва випадки $\alpha \neq 0$ і $\alpha = 0$, помножимо чисельник і знаменник правої частини формули (4.11) на вираз, спряжений для чисельника $\sqrt{\beta^2 - 3\alpha\gamma} + \beta$, та після перетворень отримаємо

$$\bar{x} = \frac{-\gamma}{\sqrt{\beta^2 - 3\alpha\gamma} + \beta} \ . \tag{4.12}$$

Це загальна формула кубічної інтерполяції. Для обчислення точки (4.12) досить визначити лише параметри $\alpha$, $\beta$ і $\gamma$ рівняння (4.9). Це можна зробити різними методами з використанням інтервалу невизначеності $[a, b]$, а також значень цільової функції $f(x)$ та її похідних. Точка з формули (4.12) повинна належати поточному інтервалу невизначеності. В методах кубічної інтерполяції формула (4.12) використовується ітераційно разом з правилами виключення інтервалів для зменшення інтервалу невизначеності. Ітерації продовжуються доти, доки відстань між послідовними точками пошуку більша допустимої помилки. До відомих методів кубічної інтерполяції належать методи кубічної інтерполяції з чотирма та двома точками.

### 4.7. Метод кубічної інтерполяції з чотирма точками

*Метод кубічної інтерполяції з чотирма точками* базується на апроксимації цільової функції за чотирма точками кубічним поліномом (4.9) та на інтерполяції шуканої точки мінімуму функції точкою мінімуму цього полінома за формулою ( 4.12) та значеннями функції.

Перші три точки графіка функції $(a,\ f_a)$, $(b,\ f_b)$, $(u,\ f_u)$ отримаємо за алгоритмом методу Свенна, де $[a, b]$ – інтервал невизначеності, $a < u < b$, $f_a = f(a)$, $f_b = f(b)$, $f_u = f(u)$. Четверту точку $(v,\ f_v)$ знайдемо за формулами квадратичної інтерполяції (4.4) та (4.5) з використанням перших трьох точок. Підставимо в рівняння кубічного полінома (4.9) координати чотирьох відомих точок функції і складемо систему рівнянь:

$$\begin{cases} p(a) = \alpha a^3 + \beta a^2 + \gamma a + \delta = f_a, \\ p(u) = \alpha u^3 + \beta u^2 + \gamma u + \delta = f_u, \\ p(v) = \alpha v^3 + \beta v^2 + \gamma v + \delta = f_v, \\ p(b) = \alpha b^3 + \beta b^2 + \gamma b + \delta = f_b. \end{cases}$$

Розв’язок цієї системи для трьох параметрів формули (4.12) отримаємо за правилом Крамера:

$$\alpha = \frac{D_\alpha}{D}, \quad \beta = \frac{D_\beta}{D}, \quad \gamma = \frac{D_\gamma}{D},$$

де

$$D = \begin{vmatrix} a^3 & a^2 & a & 1 \\ u^3 & u^2 & u & 1 \\ v^3 & v^2 & v & 1 \\ b^3 & b^2 & b & 1 \end{vmatrix}, \quad D_\alpha = \begin{vmatrix} f_a & a^2 & a & 1 \\ f_u & u^2 & u & 1 \\ f_v & v^2 & v & 1 \\ f_b & b^2 & b & 1 \end{vmatrix},$$

$$D_\beta = \begin{vmatrix} a^3 & f_a & a & 1 \\ u^3 & f_u & u & 1 \\ v^3 & f_v & v & 1 \\ b^3 & f_b & b & 1 \end{vmatrix}, \quad D_\gamma = \begin{vmatrix} a^3 & a^2 & f_a & 1 \\ u^3 & u^2 & f_u & 1 \\ v^3 & v^2 & f_v & 1 \\ b^3 & b^2 & f_b & 1 \end{vmatrix}.$$

Для раціонального розрахунку цих визначників введемо позначення, і після перетворення отримаємо формулу для обчислення точки мінімуму апроксимуючого полінома

$$w = \frac{-D_\gamma}{\sqrt{D_\beta^2 - 3D_\alpha D\gamma} + D_\beta}, \tag{4.13}$$

де

$$D_\alpha = -p_a + p_u - p_v + p_b, \tag{4.14}$$

$$D_\beta = p_a(u+v+b) - p_u(a+v+b) + p_v(a+u+b) - p_b(a+u+v), \tag{4.15}$$

$$D_\gamma = -p_a(uv+ub+vb) + p_u(av+ab+vb) -$$

$$-p_v(au+ab+ub) + p_b(au+av+uv), \tag{4.16}$$

$$p_a = f_a(v-u)(b-u)(b-v), \quad p_u = f_u(v-a)(b-a)(b-v), \tag{4.17}$$

$$p_v = f_v(u-a)(b-a)(b-u), \quad p_b = f_b(u-a)(v-a)(v-u). \tag{4.18}$$

Таким чином, інтервал невизначеності ділиться трьома внутрішніми точками $u$, $v$ і $w$ на чотири частини. За правилом виключення інтервалів одна з крайніх частин відкидається, а кубічна інтерполяція знову застосовується до нового інтервалу з двома відомими внутрішніми точками. Ітерації тривають до тих пір, поки відстань між двома послідовними точками інтерполяції перевищує допустиму помилку.

Метод кубічної інтерполяції з чотирма точками не використовує похідні цільової функції, й тому є методом нульового порядку.

На основі формул (4.13)–(4.18) складено алгоритм методу кубічної інтерполяції з чотирма точками.

Алгоритм методу кубічної інтерполяції з чотирма точками.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях інтервалу; $u$ і $f_u$ – внутрішня точка інтервалу та значення функції в ній; $\varepsilon$ – допустима помилка.

Вихідні параметри: $w$ і $f_w$ – кінцева точка та значення функції в ній.

1. Обчислити $p = (u-a)(f_b - f_u)$, $q = (b-u)(f_a - f_u)$, $s = p+q$.

2. Обчислити $v = 0{,}5 \cdot [p \cdot (a+u) + q \cdot (b+u)]/s$, $f_v = f(v)$.

3. Покласти $z = v$.

4. Якщо $v < u$, то покласти $c = u$, $f_c = f_u$, $u = v$, $f_u = f_v$, $v = c$, $f_v = f_c$.

5. Обчислити $d_{ua} = u - a$, $d_{va} = v - a$, $d_{ba} = b - a$, $d_{vu} = v - u$,

$$d_{bu} = b - u,\ d_{bv} = b - v,\ s_{au} = a + u,\ s_{vb} = v + b,\ r_{au} = a \cdot u,\ r_{vb} = v \cdot b,$$

$$p_a = f_a \cdot d_{vu} \cdot d_{bu} \cdot d_{bv},\quad p_u = f_u \cdot d_{va} \cdot d_{ba} \cdot d_{bv},\quad p_v = f_v \cdot d_{ua} \cdot d_{ba} \cdot d_{bu},$$

$$p_b = f_b \cdot d_{ua} \cdot d_{va} \cdot d_{vu},\quad D_\alpha = p_u - p_a - p_v + p_b,$$

$$D_\beta = p_a \cdot (u + s_{vb}) - p_u \cdot (a + s_{vb}) + p_v \cdot (s_{au} + b) - p_b \cdot (s_{au} + v),$$

$$D_\gamma = p_u(s_{vb} \cdot a + r_{vb}) - p_a(s_{vb} \cdot u + r_{vb}) - p_v(r_{au} + s_{au} \cdot b) + p_b(r_{au} + s_{au} \cdot v).$$

6. Обчислити $w = \dfrac{-D_\gamma}{\sqrt{D_\beta^2 - 3 \cdot D_\alpha \cdot D_\gamma} + D_\beta}$, $f_w = f(w)$, $h = w - z$.

7. Покласти $z = w$.

8. Якщо $w < u$, то покласти $c = u$, $f_c = f_u$, $u = w$, $f_u = f_w$, $w = c$, $f_w = f_c$.

9. Якщо $w < v$, то покласти $c = v$, $f_c = f_v$, $v = w$, $f_v = f_w$, $w = c$, $f_w = f_c$.

10. Якщо $f_u < f_w$, то покласти $b = w$, $f_b = f_w$, інакше покласти $a = u$, $f_a = f_u$, $u = v$, $f_u = f_v$, $v = w$, $f_v = f_w$.

11. Якщо $|h| > \varepsilon$, то перейти до кроку 5.

12. Вийти.

На кроках 1–4 проводиться квадратична інтерполяція, а на кроках 5–10 виконується ітерація методу кубічної інтерполяції.

Через велику кількість операцій у формулах кубічної інтерполяції цей метод підлягає великим обчислювальним помилкам і не дозволяє знайти точку мінімуму з високою точністю. Однак в алгоритмах багатовимірної оптимізації, де інтервал невизначеності повинен бути зменшений у 100–1000 разів, він дозволяє істотно зменшити кількість обчислень цільової функції порівняно з методом квадратичної інтерполяції.

### 4.8. Метод кубічної інтерполяції з двома точками

*Метод кубічної інтерполяції з двома точками* базується на апроксимації цільової функції за двома точками кубічним поліномом (4.9) та на інтерполяції шуканої точки мінімуму функції точкою мінімуму цього полінома за формулою (4.12) з використанням значень функції та її першої похідної.

Перші дві точки графіка функції $(a,\ f_a)$ і $(b,\ f_b)$ на кінцях інтервалу невизначеності $[a, b]$ знайдемо за алгоритмом методу Свенна. Обчислимо значення похідної функції на кінцях інтервалу $f'_a = f'(a)$ і $f'_b = f'(b)$.

Підставляючи в рівняння кубічного полінома та його похідної (4.9) і (4.10) відомі значення функції та її похідної, складемо систему рівнянь:

$$\begin{cases} p(a) = \alpha a^3 + \beta a^2 + \gamma a + \delta = f_a, \\ p(b) = \alpha b^3 + \beta b^2 + \gamma b + \delta = f_b, \\ p'(a) = 3\alpha a^2 + 2\beta a + \gamma = f'_a, \\ p'(b) = 3\alpha b^2 + 2\beta b + \gamma = f'_b. \end{cases}$$

Розв'язуючи цю систему та підставляючи значення знайдених коефіцієнтів $\alpha$, $\beta$ і $\gamma$ в рівність (4.12), після перетворень отримаємо формулу для обчислення точки мінімуму апроксимуючого полінома

$$u = a + L\frac{w - f'_a - z}{f'_b - f'_a + 2w}, \qquad (4.19)$$

де

$$z = 3(f_b - f_a)/L - f'_a - f'_b, \quad w = \sqrt{z^2 - f'_a f'_b}, \qquad (4.20)$$

$L = b - a$ – довжина інтервалу невизначеності. Потім обчислимо $f_u = f(u)$ і $f'_u = f'(u)$ та застосовуємо правило однієї точки. Якщо $f'_u = 0$, то знайдена точка мінімуму $u$ і розрахунки припиняються. Якщо $f'_u < 0$, то покладемо $a = u$, $f_a = f_u$, $f'_a = f'_u$, інакше покладемо $b = u$, $f_b = f_u$, $f'_b = f'_u$. Розрахунки тривають до тих пір, поки відстань між двома послідовними точками інтерполяції залишається більшою, ніж допустима помилка $\varepsilon$.

За формулами (4.19) і (4.20) складено алгоритм методу кубічної інтерполяції з двома точками.

А л г о р и т м   м е т о д у   к у б і ч н о ї   і н т е р п о л я ц і ї   з   д в о м а   т о ч к а м и.

Вхідні параметри: $a$ і $b$ – кінці інтервалу невизначеності; $f_a$ і $f_b$ – значення функції на кінцях інтервалу; $\varepsilon$ – допустима похибка.

Вихідні параметри: $u$ і $f_u$ – кінцева точка і значення функції в ній.

1. Покласти $v = a$.
2. Обчислити $f'_a = f'(a)$, $f'_b = f'(b)$.
3. Покласти $L = b - a$.
4. Обчислити $z = 3 \cdot (f_b - f_a)/L - f'_a - f'_b$, $w = \sqrt{z^2 - f'_a \cdot f'_b}$.
5. Обчислити $u = a + L \cdot (w - f'_a - z)/(f'_b - f'_a + 2 \cdot w)$.
6. Обчислити $h = u - v$, $f_u = f(u)$, $f'_u = f'(u)$.
7. Покласти $v = u$.
8. Якщо $f'_u = 0$, то перейти до кроку 11.
9. Якщо $f'_u < 0$, то покласти $a = u$, $f_a = f_u$, $f'_a = f'_u$, інакше покласти $b = u$, $f_b = f_u$, $f'_b = f'_u$.
10. Якщо $|h| > \varepsilon$, то перейти до кроку 3.
11. Вийти.

У цьому алгоритмі на кроках 2 та 6 значення похідної цільової функції обчислюються числовим методом диференціювання.

### 4.9. Порівняння методів поліноміальної інтерполяції

Ми розглянули різні методи поліноміальної інтерполяції, розроблені для обчислення точки мінімуму $x^*$ унімодальної цільової функції однієї змінної $f(x)$ як з використанням значень похідних функції, так і без використання похідних. За цією ознакою розглянуті методи можна розділити на три групи.

1. Методи нульового порядку. Це метод квадратичної інтерполяції з трьома точками та метод кубічної інтерполяції з чотирма точками, основані на обчисленні значень самої цільової функції без використання значень її похідних.

2. Методи першого порядку. Це метод квадратичної інтерполяції з двома точками, метод січних та метод кубічної інтерполяції з двома точками, які використовують значення першої похідної цільової функції.

3. Метод другого порядку. Це метод Ньютона, який використовує значення першої та другої похідних цільової функції.

Ці методи одновимірної оптимізації можна порівняти за швидкістю збіжності. Методи одновимірної оптимізації повинні генерувати послідовність точок $\{x_k\}$, для якої

$$\lim_{k\to\infty} x_k = x^*.$$

Тоді за властивістю границі

$$\left| x_{k+1} - x^* \right| < \left| x_k - x^* \right|,$$

тому виконується нерівність

$$\frac{\left| x_{k+1} - x^* \right|}{\left| x_k - x^* \right|} < 1.$$

Це означає *глобальну збіжність* методу. Якщо глобальна збіжність методу встановлена, то викликає інтерес оцінка її ефективності.

З практичної точки зору ефективність алгоритму залежить від числа ітерацій, необхідних для отримання наближення оптимальної точки $x^*$ з допустимою похибкою $\varepsilon$. Якщо припустити, що час обчислення ітерацій є однаковим для всіх алгоритмів, то найкращим серед них буде той, який вимагає найменшої кількості ітерацій.

Поведінка послідовності точок $\{x_k\}$ в околиці оптимальної точки $x^*$ дозволяє встановити характер *асимптотичної збіжності*. Якщо виконується нерівність

$$\lim_{k\to\infty} \frac{\left| x_{k+1} - x^* \right|}{\left| x_k - x^* \right|} = \alpha < 1,$$

то кажуть, що має місце *лінійна збіжність* і що $\alpha$ – відповідний *коефіцієнт збіжності*. Розглянуті в попередньому розділі прямі методи виключення інтервалів для зменшення інтервалу невизначеності мають лінійну збіжність.

Якщо виконується нерівність

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

то кажуть, що має місце *надлінійна збіжність*. При цьому якщо існує таке $\gamma>1$, що

$$\lim_{k\to\infty}\frac{|\,x_{k+1}-x^*\,|}{|\,x_k-x^*\,|^{\gamma}}<+\infty\,,$$

то кажуть, що має місце *збіжність порядку* $\gamma$. Зокрема, якщо

$$\lim_{k\to\infty}\frac{|\,x_{k+1}-x^*\,|}{|\,x_k-x^*\,|^{2}}<+\infty\,,$$

то кажуть, що має місце *квадратична збіжність*.

Метод квадратичної інтерполяції нульового порядку з трьома точками має надлінійну швидкість збіжності порядку $\gamma=1{,}32\ldots$, якщо функція $f(x)$ має безперервні похідні четвертого порядку і $f'(x^*)=0$, $f''(x^*)\neq 0$.

Метод січних першого порядку з двома точками має надлінійну швидкість збіжності порядку $\gamma=\tau=1{,}618\ldots$, якщо функція $f(x)$ тричі безперервно диференційована і $f'(x^*)=0$, $f''(x^*)\neq 0$.

Метод Ньютона та метод кубічної інтерполяції першого порядку з двома точками мають квадратичну швидкість збіжності, якщо функція

$f(x)$ двічі безперервно диференційована і $f'(x^*) = 0$, $f''(x^*) \neq 0$.

Отже, в малому околі оптимальної точки метод Ньютона та метод кубічної інтерполяції першого порядку з двома точками при вказаних припущеннях щодо цільової функції $f(x)$ сходяться швидше за інші методи.

Однак у загальному випадку при роз'язанні конкретної задачі оптимізації необхідні обчислювальні експерименти для вибору найефективнішого методу, що дозволяє роз'язати задачу оптимізації з допустимою похибкою при використанні мінімальних обчислювальних ресурсів.

**Лабораторна робота**

Тема лабораторної роботи

Розробка підпрограми для методів поліноміальної інтерполяції.

Мета лабораторної роботи.

Розробити підпрограму методу поліноміальної інтерполяції для оптимізації унімодальної цільової функції.

Порядок виконання лабораторної роботи.

1. Отримати у викладача задачу на програмування одного з методів поліноміальної інтерполяції:

1) метод квадратичної інтерполяції з трьома точками;

2) квадратичної інтерполяції з двома точками;

3) метод січних;

4) метод Ньютона;

5) метод кубічної інтерполяції з чотирма точками;

6) метод кубічної інтерполяції з двома точками.

2. На підставі алгоритму для заданого методу поліноміальної інтерполяції написати комп'ютерну підпрограму відповідно до вимог:

1) вхідні параметри підпрограми: $f(x)$ – цільова функція; $N$ – гранична кількість точок; $\varepsilon$ – допустима похибка, вхідні параметри алгоритму заданого методу;

2) вихідні параметри підпрограми: $X$ – масив усіх точок пошуку $x_k$; $F$ – массив відповідних значень функції $f_k = f(x_k)$; $x$ і $f_x$ – інайкраща точка пошуку та значення у ній функції;

3) на кожній ітерації виводити на екран комп'ютера рядок, що містить номер точки пошуку $k$, довжину інтервалу невизначеності $L_k$, значення функції $f_k$ та значення змінної $x_k$;

4) ітерації продовжуються доти, доки виконуються нерівності:

$$\left| L_k \right| > \varepsilon, \quad k \le N;$$

5) після закінчення процесу оптимізації на екрані повинна відображатися таблиця, що має відповідний заголовок і представляє процес мінімізації функції;

6) під таблицею необхідно відобразити кількість обчислень цільової функції, довжину кінцевого інтервалу невизначеності, мінімальне значення функції та відповідне йому значення незалежної змінної.

3. Написати загальну програму для виконання обчислювального процесу мінімізації цільової функції однієї змінної та відображення на екрані цього процесу відповідно до вимог:

1) за основу загальної програми для мінімізації функції взяти загальну програму пошуку інтервалу невизначеності з попередньої лабораторної роботи та модифікувати цю програму;

2) задати у програмі значення $x_0$, $x^*$, $h$, $\varepsilon$, $N$;

3) шляхом виклику підпрограми методу Свенна з використанням підпрограми обчислення функції $f(x)$ та значень $x_0$, $h$, $N$ реалізувати ітераційний процес пошуку інтервалу невизначеності з побудовою таблиці процесу на екрані та отриманням масивів значень змінної $X$ та функції $F$, меж інтервалу невизначеності $a$ і $b$;

4) шляхом виклику розробленої підпрограми поліноміальної інтерполяції реалізувати ітераційний процес зменшення інтервалу з побудовою таблиці процесу на екрані та отриманням масивів значень змінної $X$

та функції $F$, кращої точки пошуку $x$ та значення в ній функції $f_x$;

5) за допомогою підпрограми графічного відображення ітераційного процесу по масивах $X$ і $F$ відобразити графік процесу мінімізації цільової функції однієї змінної на екрані комп'ютера.

4. Запустити загальну програму мінімізації функції однієї змінної для функцій попередніх лабораторних робіт із кількох початкових точок

$$x_{01} < x^*, \quad x_{02} = x^*, \quad x_{03} > x^*$$

при кількох значеннях кроку

$$h_{01} < \left| x_0 - x^* \right|, \quad h_{02} = \left| x_0 - x^* \right|, \quad h_{03} > \left| x_0 - x^* \right|.$$

5. Виконати аналіз одержаних результатів:

1) на підставі табличної та графічної інформації про роботу методів одновимірного пошуку визначити кінцеві точки методу Свенна та знайдений цим методом інтервал невизначеності;

2) на кількох початкових ітераціях за табличними даними та графіками зіставити початок роботи методу поліноміальної інтерполяції з його робочими формулами;

3) оцінити ефективність методу поліноміальної інтерполяції та зіставити її з теоретичною ефективністю.

6. Відповісти на запитання викладача на тему лабораторної роботи.

7. Оформити та здати звіт про проведену лабораторну роботу.

З м і с т  з в і т у:

1) титульний аркуш установленого зразка із зазначенням організації, назви навчальної дисципліни, теми роботи, номера варіанта, виконавця та приймаючого, міста, року;

2) постановка задачі мінімізації цільової функції, що включає задачу мінімізації функції, початкову точку і точку мінімуму;

3) теоретичний опис методу поліноміальної інтерполяції;

4) алгоритм методу;

5) роздруківки використовуваних комп'ютерних програм;

6) табличне подання процесу мінімізації та його підсумки – кількість обчислень значень цільової функції, мінімальне значення функції та відповідне йому значення точки мінімуму, досягнуту точність;

7) двовимірний графік траєкторії пошуку з усіма точками;

8) аналіз одержаних результатів;

9) висновки про виконану лабораторну роботу.

**Контрольні запитання**

1. Яке призначення методу квадратичної інтерполяції з трьома точками?
2. Опишіть метод квадратичної інтерполяції з трьома точками.
3. Як починається пошук мінімуму в методі квадратичної інтерполяції з трьома точками?
4. Якого порядку є метод квадратичної інтерполяції з трьома точками?
5. Які переваги та недоліки методу квадратичної інтерполяції з трьома точками?
6. Яке призначення методу квадратичної інтерполяції з двома точками?
7. Опишіть метод квадратичної інтерполяції з двома точками.
8. Які переваги та недоліки методу квадратичної інтерполяції з двома точками?
9. Яке призначення методу січних?
10. Опишіть метод січних.
11. Якого порядку метод січних?
12. Які переваги та недоліки методу січних?
13. Яке призначення методу Ньютона?
14. Виведіть ітераційну формулу методу Ньютона.
15. Опишіть метод Ньютона.
16. Якого порядку метод Ньютона?
17. Які переваги та недоліки методу Ньютона?

18. Яке призначення методу кубічної інтерполяції з чотирма точками?

19. Опишіть метод кубічної інтерполяції з чотирма точками.

20. Як починається пошук мінімуму в методі кубічної інтерполяції з чотирма точками?

21. Якого порядку метод кубічної інтерполяції з чотирма точками?

22. Які переваги та недоліки методу кубічної інтерполяції з чотирма точками?

23. Яке призначення методу кубічної інтерполяції з двома точками?

24. Опишіть метод кубічної інтерполяції з двома точками.

25. Які переваги та недоліки методу кубічної інтерполяції з двома точками?

26. На які групи можна поділити методи поліноміальної інтерполяції за порядком похідних?

27. Дайте визначення глобальної збіжності методу одновимірної оптимізації.

28. Дайте визначення асимптотичної збіжності методу одновимірної оптимізації.

29. Як визначається лінійна збіжність методу одновимірної оптимізації?

30. Наведіть приклади методів одновимірної оптимізації, які мають властивість лінійної збіжності.

31. Як визначається надлінійна збіжність методу одновимірної оптимізації?

32. Наведіть приклади методів одновимірної оптимізації, які мають властивість надлінійної збіжності.

33. Як визначається квадратична збіжність методу одновимірної оптимізації?

34. Наведіть приклади методів одновимірної оптимізації, які мають властивість квадратичної збіжності.

## СПИСОК ЛІТЕРАТУРИ

1. Григорків В. С. Оптимізаційні методи та моделі : підручник / В. С. Григорків, М. В. Григорків. – Чернівці : ЧНУ, 2016. – 400 с.

2. Задачин В. М. Чисельні методи : навч. посіб. / В. М. Задачин, І. Г. Конюшенко. – Харків : Вид-во ХНЕУ ім. С. Кузнеця, 2014. – 180 с.

3. Зайченко Ю. П. Дослідження операцій : підручник. – 7-ме вид. / Ю. П. Зайченко. – Київ : Видавничий дім «Слово», 2006. – 816 с.

4. Катренко А. В. Дослідження операцій : підручник / А. В. Катренко. – Львів : «Магнолія Плюс», 2005. – 549 с.

5. Кісельова О. М. Чисельні методи оптимізації : навч. посіб. / О. М. Кісельова, А. Є. Шевельова. – Дніпропетровськ : Вид-во ДНУ, 2008. – 208 с.

6. Кочевський О. М. Конспект лекцій з дисципліни «Моделі і методи оптимізації» / О. М. Кочевський, О. Г. Гусак. – Суми : Вид-во СУМДУ, 2005. – 74 с.

7. Мовчан А. П. Методи статичної оптимізації : навч. посіб. / А. П. Мовчан, О. В. Степанець. – Київ : НТУУ «КПІ», 2012. – 138 с.

8. Ладогубець Т. С. Методи оптимізації без використання похідних : практикум з дисципліни «Дослідження операцій» / Т. С. Ладогубець, О. Д. Фіногенов. – Київ : КПІ ім. Ігоря Сікорського, 2020. – 45 с.

9. Наконечний С. І. Математичне програмування : навч. посіб. / С. І. Наконечний, С. С. Савіна. – Київ : КНЕУ, 2003. – 452 с.

10. Нефьодов Ю. М. Методи оптимізації в прикладах і задачах : навч. посіб. / Ю. М. Нефьодов, Т. Ю. Балицька. – Київ : Кондор, 2011. – 324 с.

11. Нікуліна О. М. Чисельні методи моделювання та оптимізації управління динамічними системами : навч. посібник / О. М. Нікуліна, В. П. Северин. – Харків : НТУ «ХПІ», 2024. – 144 с.

12. Основи математичних методів дослідження операцій / Лавров Є. А. , Клименко Н. А. , Перхун Л. П. та ін.; за ред. Н. А. Клименко. – Київ : ЦК «Компринт», 2015. – 452 с.

13. Ржевський С. В. Дослідження операцій : підручник / С. В. Ржевський, В. М. Александрова. – Київ : Академвидав, 2006. – 560 с.

14. Северин В. П. Методи та алгоритми багатовимірної безумовної оптимізації : навч. посіб. / В. П. Северин, О. М. Нікуліна. – Харків : НТУ «ХПІ», 2023. – 160 с.

15. Сікора Я. Б. Методи оптимізації : навч.-метод. посібник / Я. Б. Сікора. – Житомир : Вид-во ЖДУ ім. Івана Франка, 2012. – 82 с.

16. Сікора Я. Б. Методи оптимізації та дослідження операцій : навч. посіб. / Я. Б. Сікора, А. Й. Щехорський, Б. Л. Якимчук. – Житомир : Вид-во ЖДУ ім. Івана Франка, 2019. – 148 с.

17. Шелудько Г. А. Методи розв'язання задач оптимізації : конспект лекцій / Г. А. Шелудько, В. В. Науменко. – Харків : УкрДАЗТ, 2014. – 50 с.

18. Bazaraa M. S. Nonlinear programming: theory and algorithms / M. S. Bazaraa, H. D. Sherali, C. M. Shetty. – 3-d edition. – Hoboken, New Jersey : John Wiley & Sons Inc., 2006. – 853 p.

19. Bunday B. D. Basic optimization methods / B. D. Bunday. – London : Edward Arnold Limited, 1984. – 136 p.

20. Dennis J. E. Numerical methods for unconstrained optimization and nonlinear equations / J. E. Dennis, R. B. Schnabel. – New Jersey : Prentice-Hall Inc, Englewood Cliffs, 1983. – 378 p.

21. Fletcher R. Practical methods of optimization / R. Fletcher. – 2-d edition. – Chichester : John Wiley & Sons Ltd, 2000. – 436 p.

22. Forsythe G. E. Computer methods for mathematical computations / G. E. Forsythe, M. A. Malcolm, C. B. Moler. – New York : Prentice-Hall Inc., Englewood Cliffs, 1977. – 270 p.

23. Gill P. E. Practical optimization / P. E. Gill, W. Murray, M. H. Wright. – London, New York, Toronto, Sydney, San Francisco: Academic Press Inc. Ltd., 1981. – 401 p.

24. Himmelblau D. M. Applied nonlinear programming / D. M. Himmelblau. – Texas, Austin: McGraw-Hill Book Company, 1972. – 498 p.

25. Kahaner D. Numerical methods and software / D. Kahaner,

C. Moler, S. Nash. – Englewood Cliffs: Prentice-Hall Inc., 2001. – 495 p.

26. Luenberger D. G. Linear and nonlinear programming / D. G. Luenberger, Y. Yinyu. – 3-d edition. – New York: Springer, 2008. – 546 p.

27. Mathews J. H. Numerical methods using MATLAB / J. H. Mathews, K. D. Fink. – 3-d edition. – New York: Prentice-Hall Inc., 1999. – 662 p.

28. Minoux M. Programmation mathématique: théorie et algorithms / M. Minoux. – Paris : Dunod, 1989. – 488 p.

29. Polak E. Computational methods in optimization: a unified approach / E. Polak. – New York, London: Academic Press, 1971. – 329 p.

30. Reklaitis G. V. Engineering optimization: methods and applications / G. V. Reklaitis, A. Ravindran, K. M. Ragsdell. – New York, Chichester, Brisbane, Toronto, Singapore: A Willey-Interscience Publication, John Willey and Sons Inc., 1983. – 684 p.

31. Sun W. Optimization theory and methods. Nonlinear programming / W. Sun, Y. X. Yuan. – New York : Springer, 2006. – 687 p.

32. Yang W. Y. Applied numerical methods using MATLAB / W. Y. Yang, W. Cao, T.-S. Chung, J. Morris. – New Jersey : John Wiley & Sons Inc., 2005. – 662 p.

Навчальне видання

СЕВЕРИН Валерій Петрович
НІКУЛІНА Олена Миколаївна

# МЕТОДИ ТА АЛГОРИТМИ ОДНОВИМІРНОЇ ОПТИМІЗАЦІЇ

Навчальний посібник з навчальної дисципліни
«Дослідження операцій» для студентів
напряму «Інформаційні технології»,

Відповідальний за випуск доц. Хацько Н. Є.
Роботу до видання рекомендував доц. Хацько Н. Є.
Редактор О. С. Самініна

План 2024 р., поз. 106

Підп. до друку 2025 р.
Гарнітура Times New Roman. Ум. друк. арк. 5,2

---

Видавничий центр НТУ «ХПІ».
Свідоцтво про державну реєстрацію ДК № 5478 від 21.08.2017 р.
61002, Харків, вул. Кирпичова, 2

---

Електронне видання