Чисельні методи оптимізації функцій

Лекція 11

В основі проектувати нових, ефективніших та дешевших технічних систем, а також для покращення якості функціонування існуючих систем лежить процес оптимізації.

Ефективність методів оптимізації, які дозволяють вибирати найкращий варіант без перевірки всіх можливих варіантів, тісно пов'язана із широким використанням досягнень у галузі математики шляхом реалізації ітераційних обчислювальних схем.

Задача одновимірної оптимізації

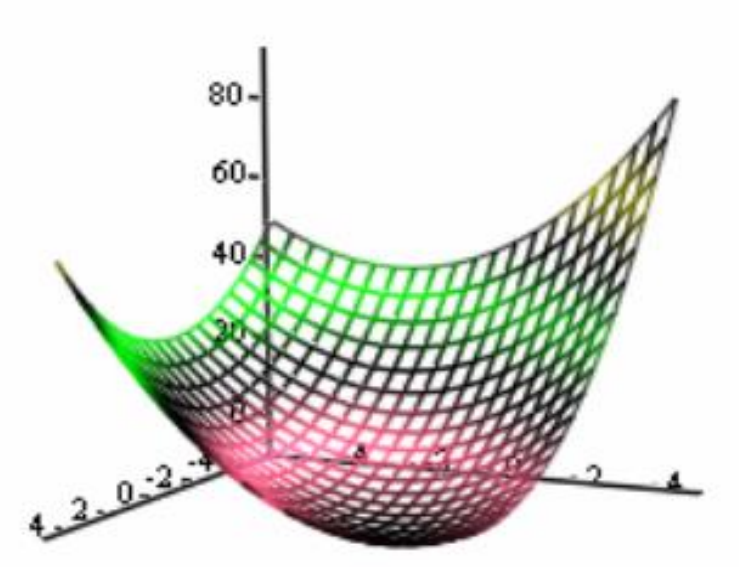
Цільова функція *(функція мети) – це вираз (функція), значення якого необхідно мінімізувати або максимізувати.*

Задана функція F(x). Треба знайти мінімум (максимум) цієї функції на інтервалі [a,b] з заданою точністю ε.

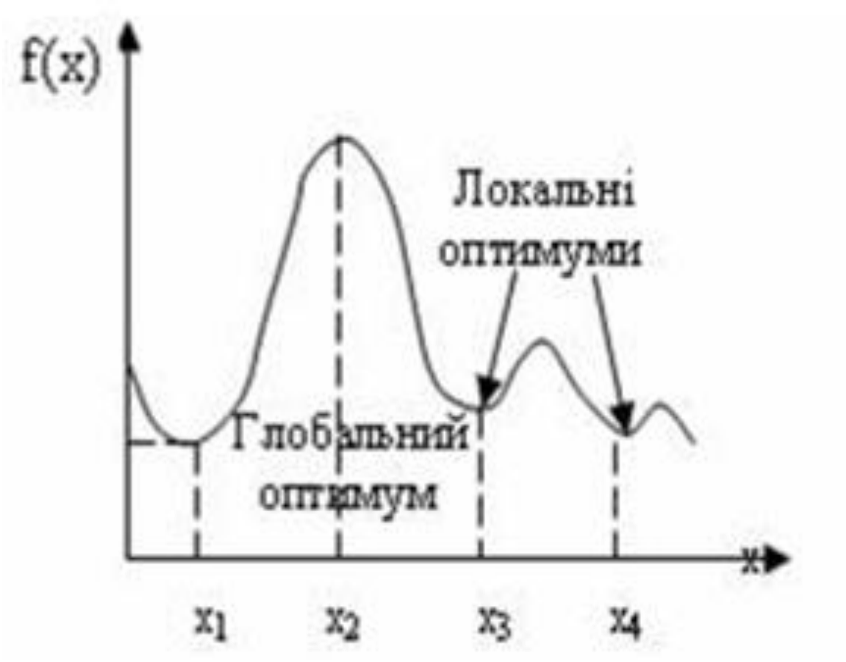
Цільова функція F(x) є ***унімодальною****, якщо вона має єдиний екстремум на [a,b].*

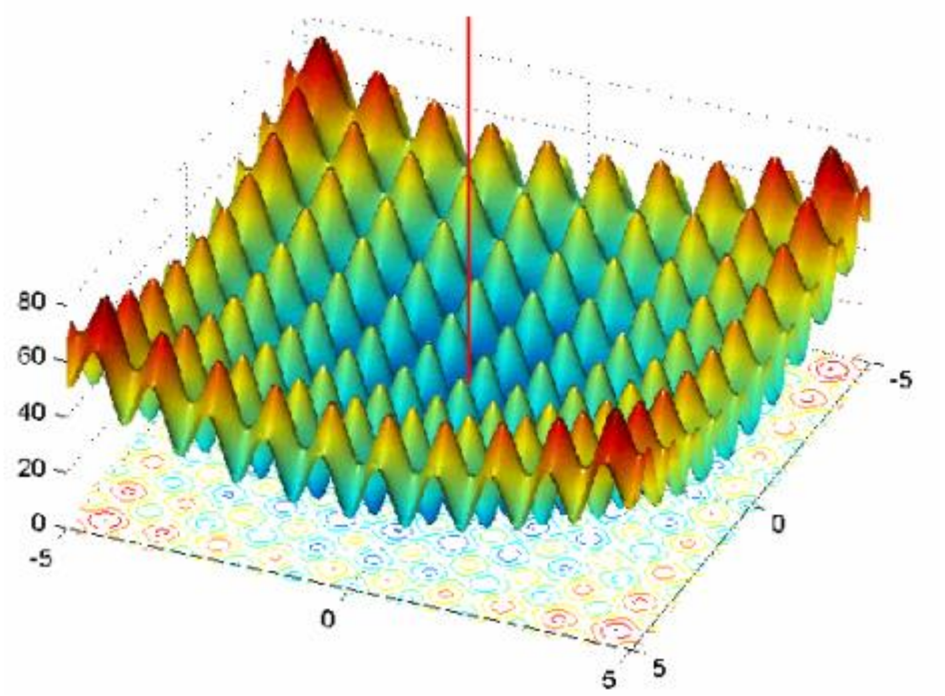
Деякі алгоритми оптимізації пристосовані до пошуку максимуму, а інші – для пошуку мінімуму.

Унімодальні функції



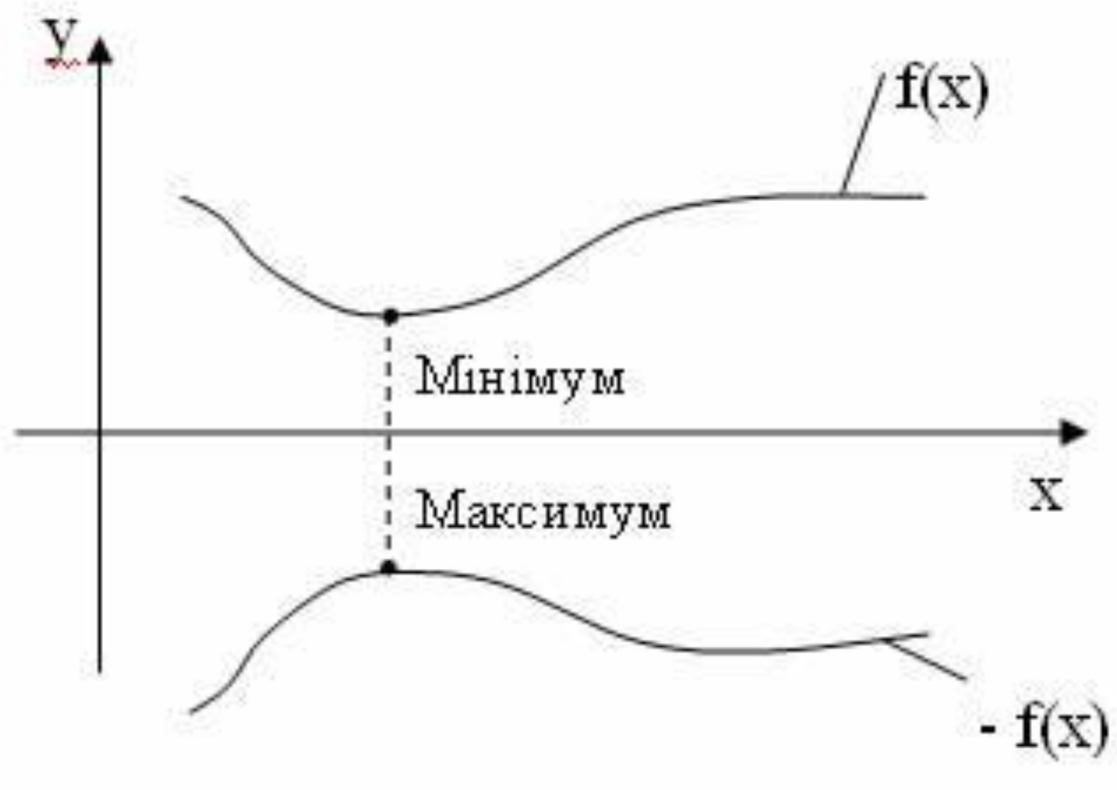
Функція з локальними та глобальними оптимумами





Зведення задачі мінімізації до задачі пошуку максимуму

(зміна знаку цільової функції на протилежний)



*Надалі вважаємо, що розв'язується **задача на мінімум**.*

Методи оптимізації

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

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

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

Формулювання задачі одновимірної оптимізації

Знайти значення ***проектного параметра*** x *цільової функції* $f(x)$ *на* ***інтервалі невизначеності [a,b].***

В процесі пошуку оптимуму цільової функції цей інтервал постійно зменшується (звужується).

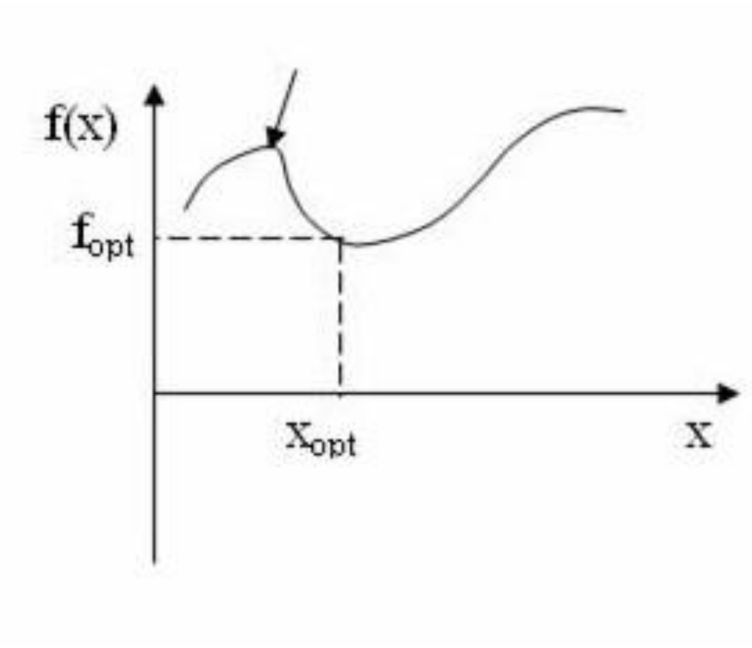
Тому методи одновимірної оптимізації називають ***методами звуження інтервалу невизначеності.***

Основні чисельні методи одновимірної оптимізації

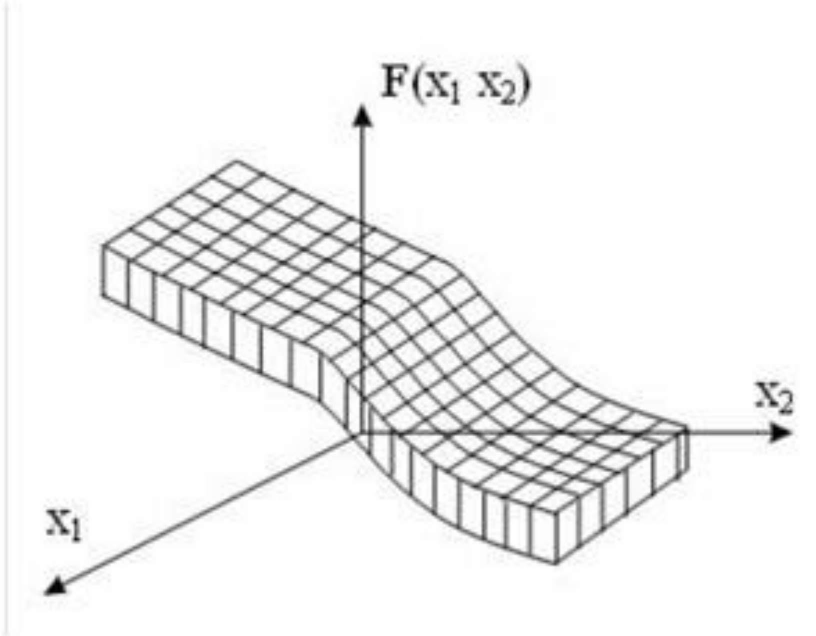
- *Метод рівномірного пошуку,*
- *метод дихотомії (поділу відрізку навпіл) ,*
- *метод золотого перерізу,*
- *метод Фібоначчі.*

Вид цільової функції

Однопараметрична



Багатопараметрична



Етапи комп'ютерної одновимірної оптимізації

- ***Локалізація*** *- встановлення меж відрізка, на якому реалізується процедура пошуку оптимуму.*
- ***Уточнення*** *- зменшення відрізка до заданої похибки обчислення за допомогою ітераційних алгоритмів.*

Умова закінчення ітераційного процесу:

$$|b-a| \le \varepsilon.$$

Процедура пошуку мінімуму

- *розбиття відрізка на* ***відрізки унімодальності*** *- в кожному з яких функція унімодальна;*
- *уточнення точки мінімуму на кожному відрізку, тобто* ***знаходження локального мінімуму*** *із заданою точністю;*
- ***вибір точки мінімуму*** *серед точок локального мінімуму шляхом виділення тієї, в якій значення функції мети найменше.*

Відділення відрізків унімодальності

1. Виберемо n, визначимо крок

$$h_0 = (b-a)/n.$$

$x_0 = a$, $x_1 = x_0 + h_0$, $x_2 = x_0 + 2h_0$, $\ldots$, $x_n = x_0 + n \cdot h = b$.

Обчислимо

$$y_0 = f(x_0), \quad y_1 = f(x_1), \ldots, y_N = f(x_n).$$

2. Розглянемо відрізки $[x_{i-1}, x_{i+1}]$, $i = 1, 2, \ldots, n-1$.

Вибираємо такі: $y_i \le y_{i-1}$, $y_i \le y_{i+1}$.

До них приєднаємо:

відрізок $[a, x_1]$, якщо $y_0 < y_1$,

відрізок $[x_{n-1}, b]$, якщо $y_{n-1} > y_n$.

Число виділених відрізків - K_0.

3. Процедуру виділення відрізків повторюємо для

$$h_1 = h_0 / 2, \quad h_2 = h_0 / 4, \ldots, h_m = h_0 / 2^m .$$

Якщо з деякого кроку h_i числа K_i, K_{i+1} . K_{i+k-1} співпадають підряд k-разів, то процес відділення закінчуємо, і як відрізки унімодальності, вибираємо відрізки, виділені для кроку h_{i+k-1}. На практиці число n вибирається таким, щоб крок h_0 був у декілька разів більше заданої точності ε, а число k в проміжку від 2 до 5.

Приклад 1. *Відділити точки локального мінімуму для функції*

$$f(x) = 5\cos x - 3\sin x$$

на відрізку $[2;9{,}2]$ *при* $n = 12$ *і* $k = 2$.

Крок

$$h_0 = (9{,}2 - 2)/12 = 7{,}2/12 = 0{,}6.$$

$$x_0 = a = 2$$

$$x_1 = a + h_0 = 2 + 0{,}6 = 2{,}6$$

$$x_2 = a + 2h_0 = 2 + 1{,}2 = 3{,}2$$

…………………………………

$$x_{12} = a + 12h_0 = 2 + 7{,}2 = 9{,}2$$

i	0	1	2	3	4	5	6
x_i	2	2,6	3,2	3,8	4,4	5,0	5,6
$f(x_i)$	-4,807	-5,883	-4,816	-2,119	1,321	4,295	5,773
i	7	8	9	10	11	12	
x_i	6,2	6,8	7,4	8	8,6	9,2	
$f(x)$	5,229	2,863	-0,502	-3,697	-5,597	-5,542	

Розглядаємо відрізки $[2;3{,}2]$, $[2{,}6;3{,}8]$, $[3{,}2;4{,}4]$

$\underline{[2;3{,}2]}$

$$-5{,}883 = f(2{,}6) < f(2) = -4{,}807$$

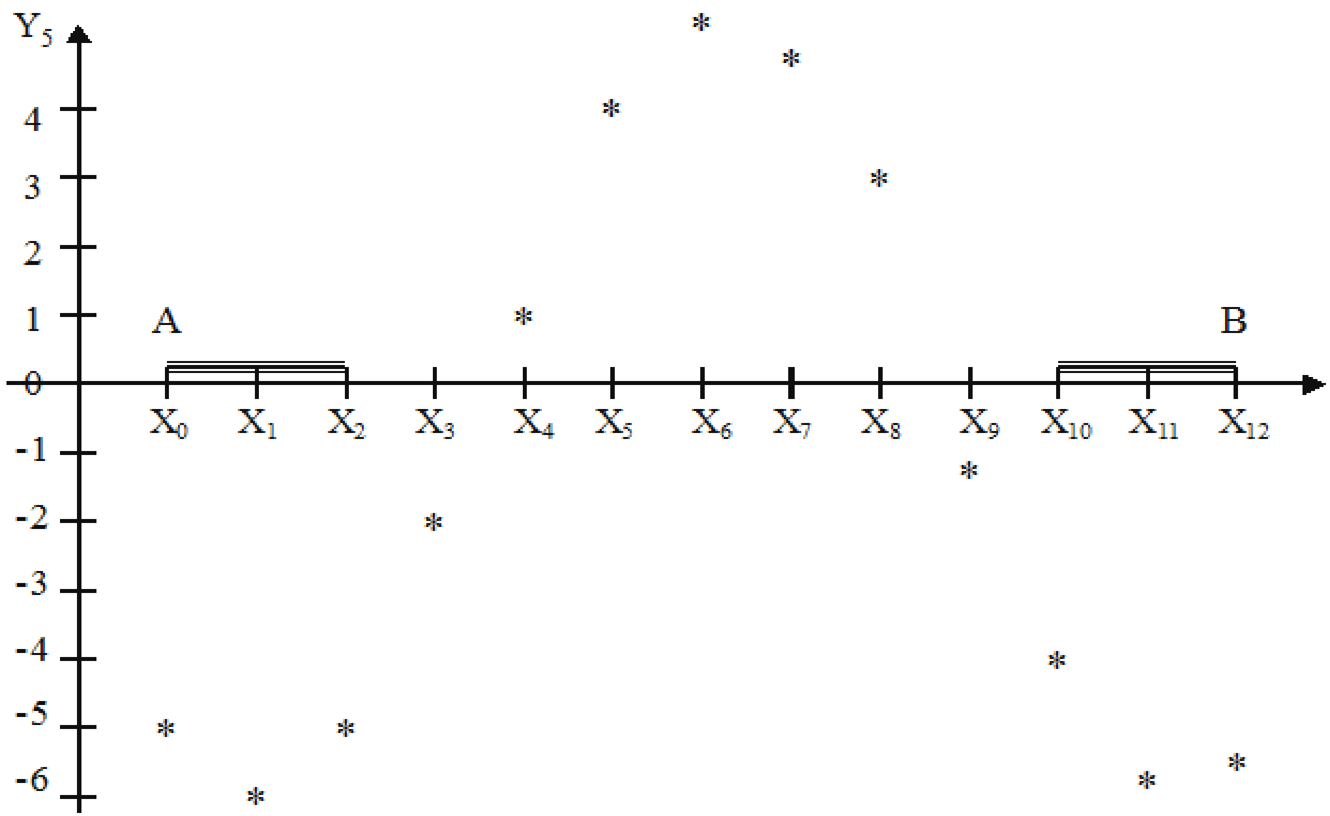
$$-5{,}883 = f(2{,}6) < f(3{,}2) = -4{,}816$$

1) $[2;3{,}2]$ *з центром в точці* $x = 2{,}6$ *і*

2) $[8;9{,}2]$ *з центром в точці* $x = 8{,}6$.

Виділення відрізків унімодальності

$K_0 = 2.$



Крок $h_1 = 0,6/2 = 0,3$, $K_1 = 2$

1) $[2,3; 2,9]$ з центром в точці $x = 2,6$ і

2) $[8,6; 9,2]$ з центром в точці $x = 8,9$.

Оскільки $K_0 = K_1 = 2$, то завершуємо процес.

i	x_i	$f(x_i)$	i	x_i	$f(x_i)$
0	2	-4,807	12	5,6	5,773
1	2,3	-5,568	13	5,9	5,757
2	2,6	-5,883	14	6,2	5,229
3	2,9	-5,572	15	6,5	4,24
4	3,2	-4,816	16	6,8	2,863
5	3,5	-3,627	17	7,1	1,248
6	3,8	-2,119	18	7,4	-0,502
7	4,1	-0,421	19	7,7	-2,199
8	4,4	1,321	20	8,0	-3,697
9	4,7	2,937	21	8,3	-4,863
10	5	4,295	22	8,6	-5,597
11	5,3	5,266	23	8,9	-5,83
			24	9,2	-5,542

Метод рівномірного пошуку

Виконується табуляція функції F(x) з деяким кроком зміни x і визначається її найменше (найбільше) значення.

Метод рівномірного пошуку – це найбільш непродуктивний метод розв’язання задачі оптимізації.

Метод дихотомії

1. Ввести значення **A, B,** ε .

2. Обчислити x_m**=(A+B)/2 .**

3. Обчислити значення функції $f_m = f(x_m)$.

4. Обчислити значення **L=B-A** .

5. Якщо **L<**ε процес завершити, за точку мінімуму вважати x_m .

6. Обчислити значення

$x_1 = a+L/4$, $f_1 = F(x_1)$,

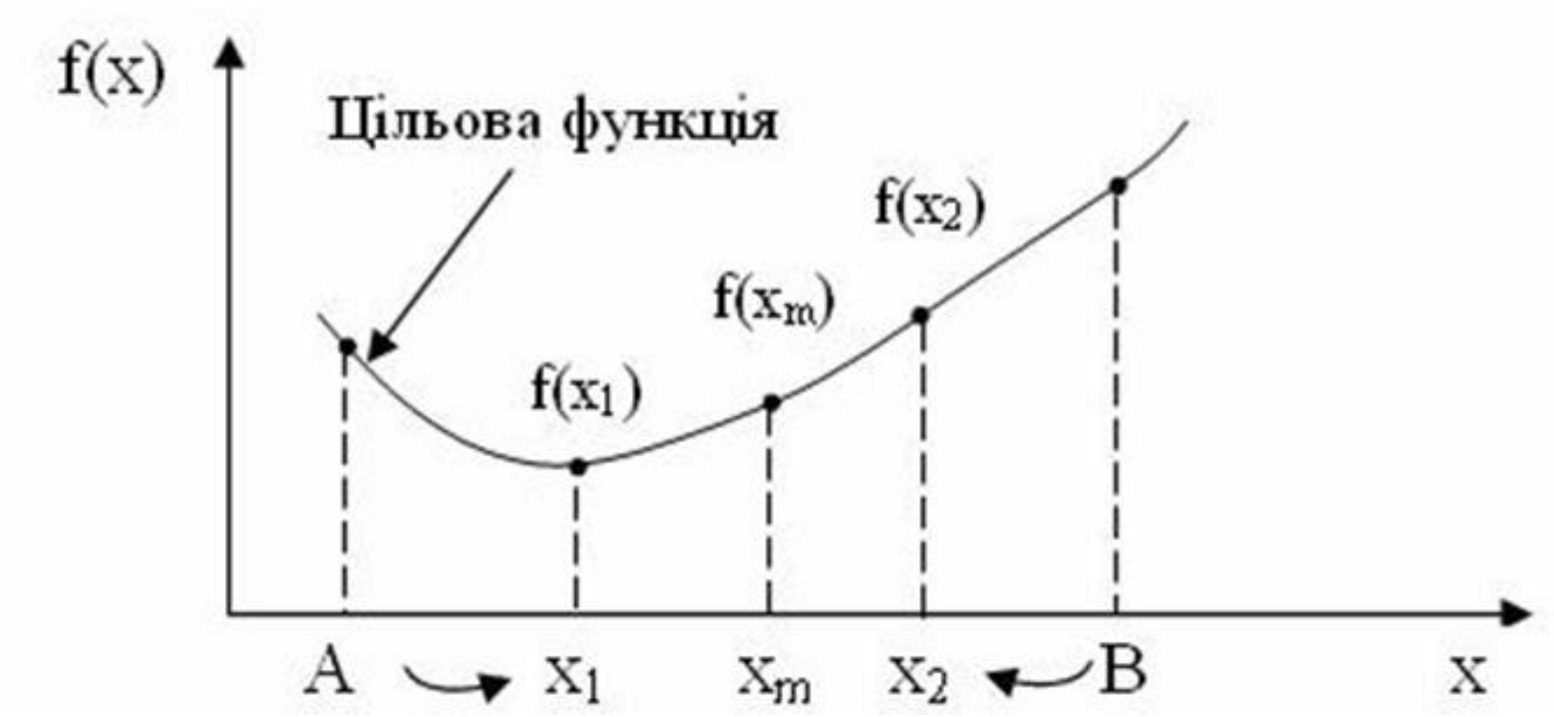
$x_2 = b-L/4$, $f_2 = F(x_2)$.

7. Якщо $f_1 < f_m$, тоді $B = x_m$, $x_m = x_1$, $f_m = f_1$; перейти до п.4;

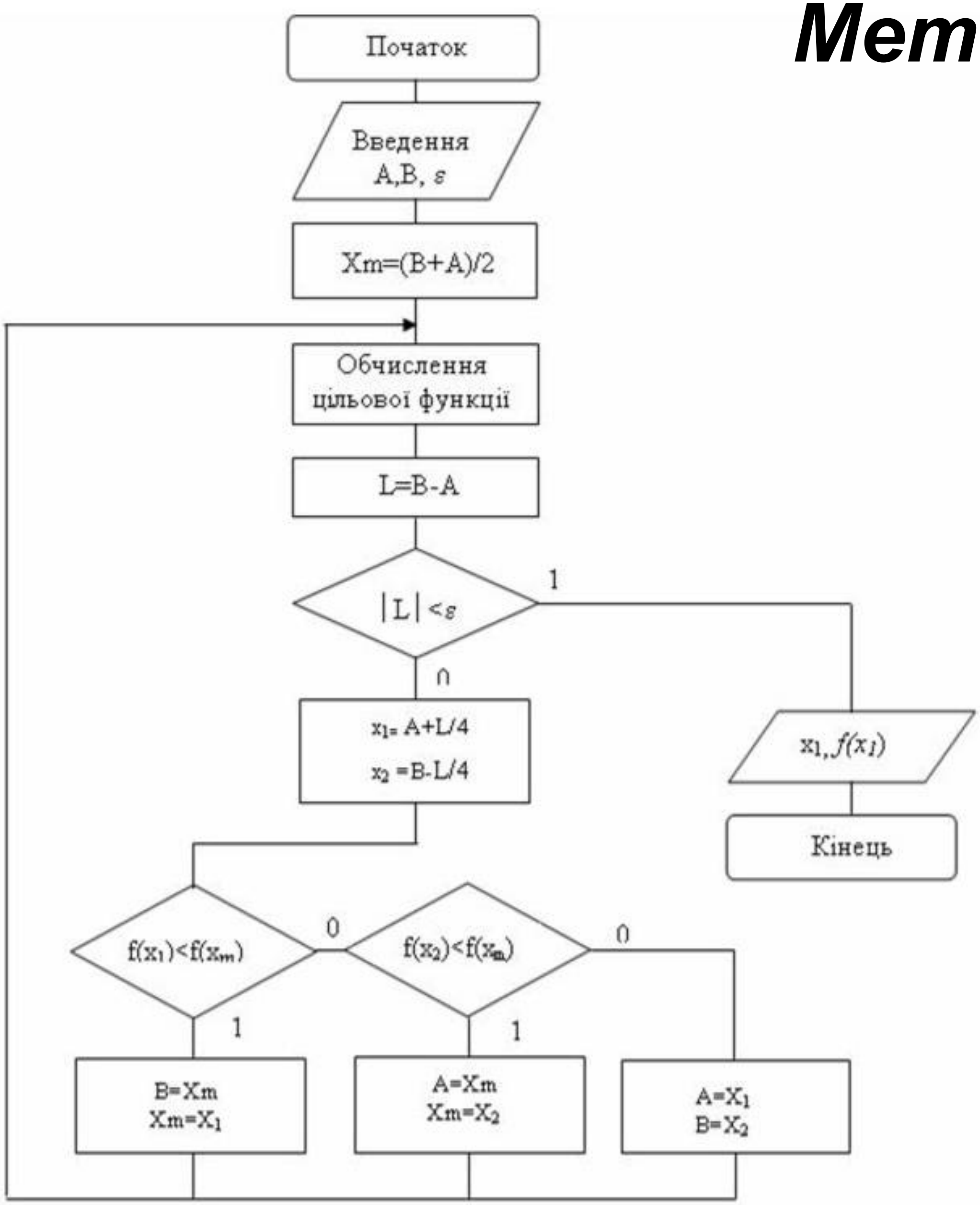
якщо $f_2 < f_m$, тоді $A = x_m$, $x_m = x_2$, $f_m = f_2$; перейти до п.4.

8. Виконати $A = x_1$, $B = x_2$; перейти до п.4.

Метод дихотомії



Метод дихотомії



Метод золотого перетину

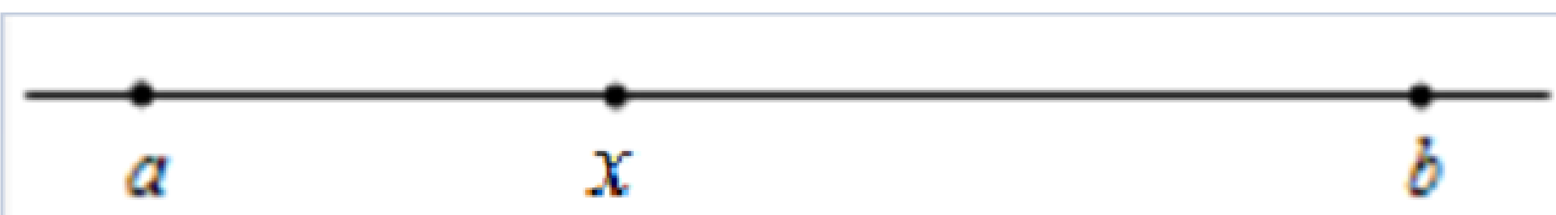
Цей метод є **найбільш ефективним.**

$$\frac{b-a}{b-x} = \frac{b-x}{x-a},$$

$$x_1 = a + (1-\alpha)(b-a) \qquad x_2 = a + \alpha(b-a),$$

де $\alpha = \dfrac{\sqrt{5}-1}{2} \approx 0{,}618034$.

Ілюстрація методу золотого перетину



Точка x_1 виконує золотий перетин відрізка $[a, x_2]$, а точка x_2 - золотий перетин відрізка $[x_1, b]$.

Параметри, що задаються: a, b, ε. У алгоритмі застосовується також параметр α, він дорівнює: $\alpha = \left(\sqrt{5} - 1\right)/2$.

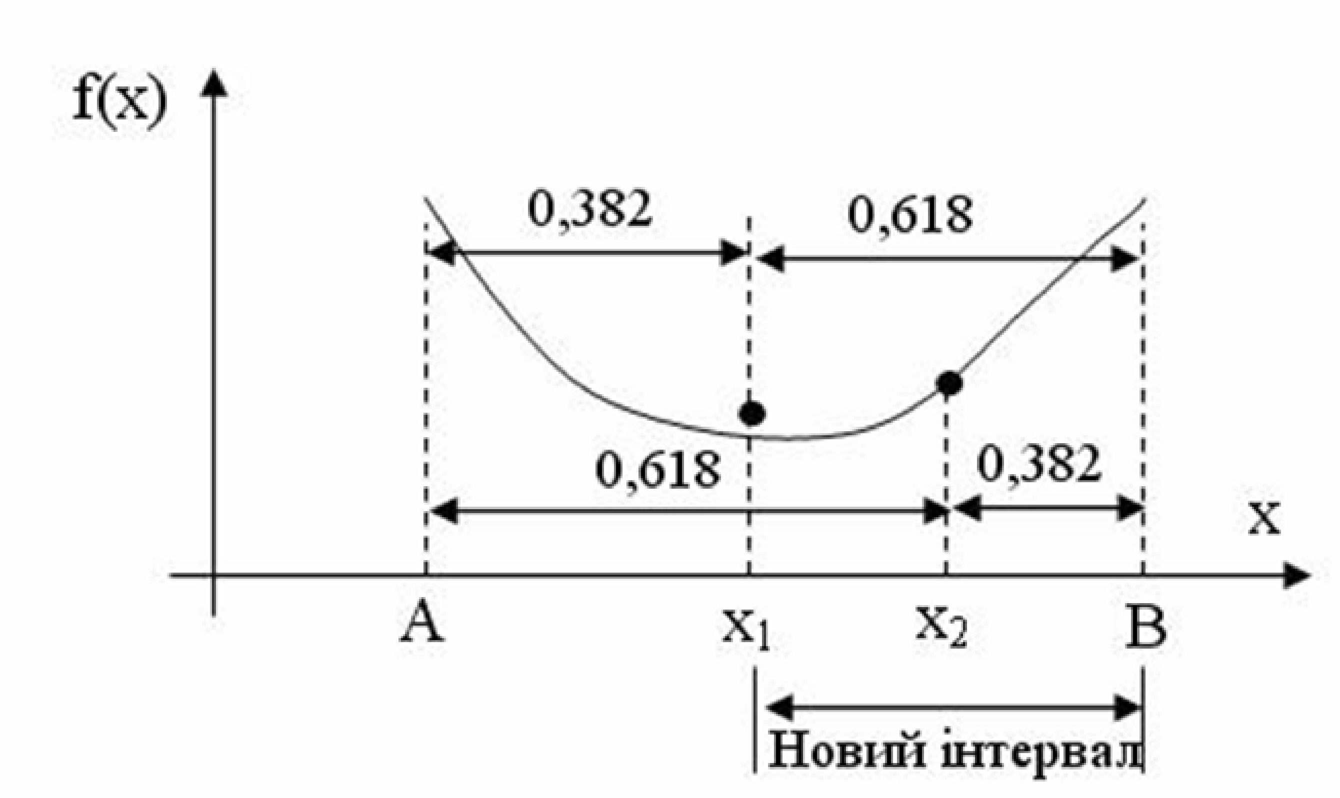


Схема алгоритму:

1. Задати параметри: a, b, ε. Обчислити $l = b-a$.

2. Якщо $l<\varepsilon$, то завершити процес. За точку мінімуму вважати значення $x = (a+b)/2$.

3. Обчислити $x_1 = a+(1-\alpha)\, l$, $x_2 = a+\alpha\, l$.

4. Обчислити нове значення l за правилом: $l = \alpha\, l$.

5. Якщо $f(x_1)<f(x_2)$, то продовжуємо пошук мінімуму на проміжку $[a, x_2]$, тобто $b=x_2$, $x_2=x_1$, $x_1=a+(1-\alpha)l$,

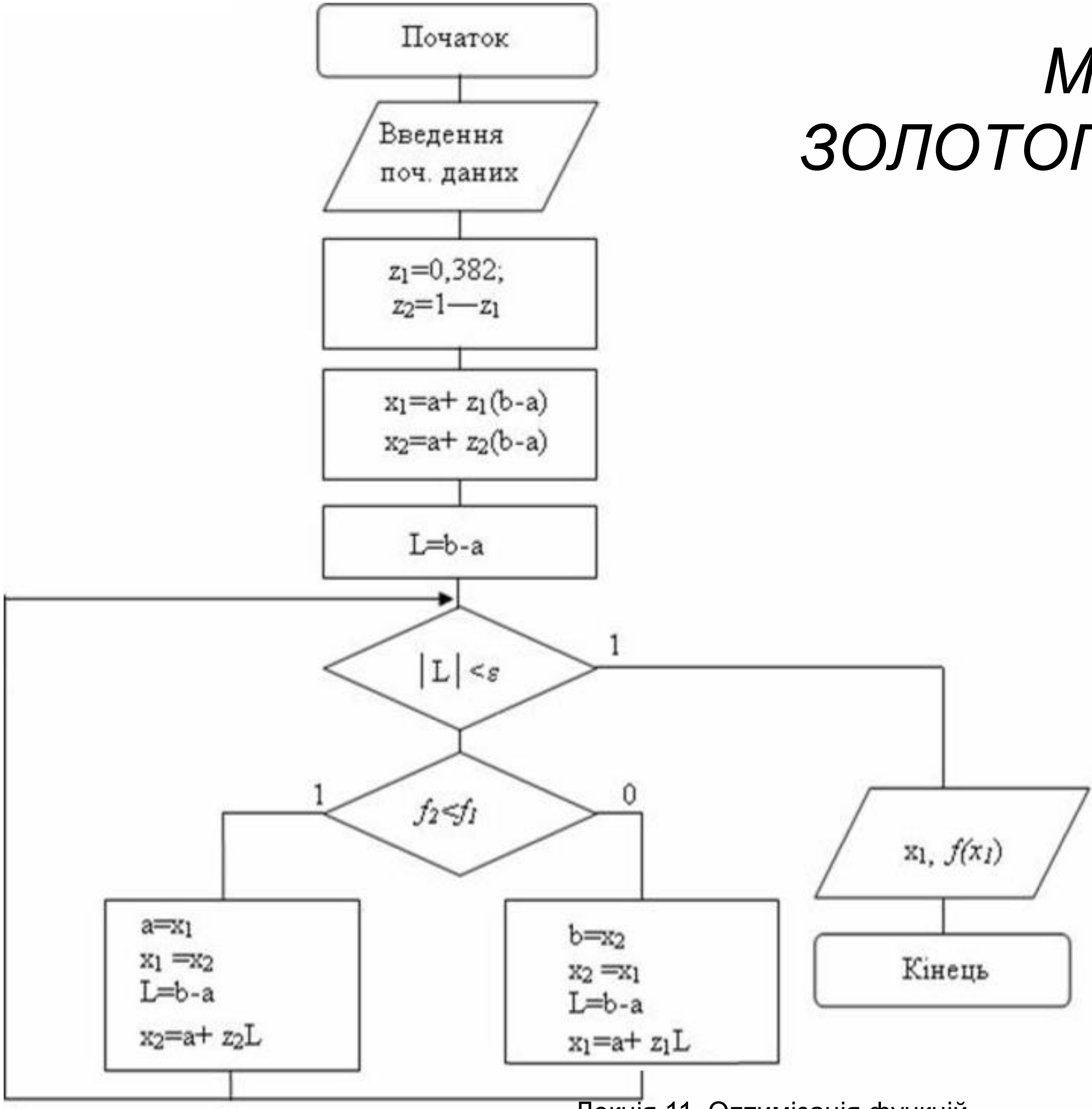
якщо ні, то на проміжку $[x_1, b]$, тобто $a=x_1$, $x_1=x$, $x_2=a+\alpha\, l$.

На цих проміжках вже є одна точка, що робить золотий перетин, потрібно визначити іншу.

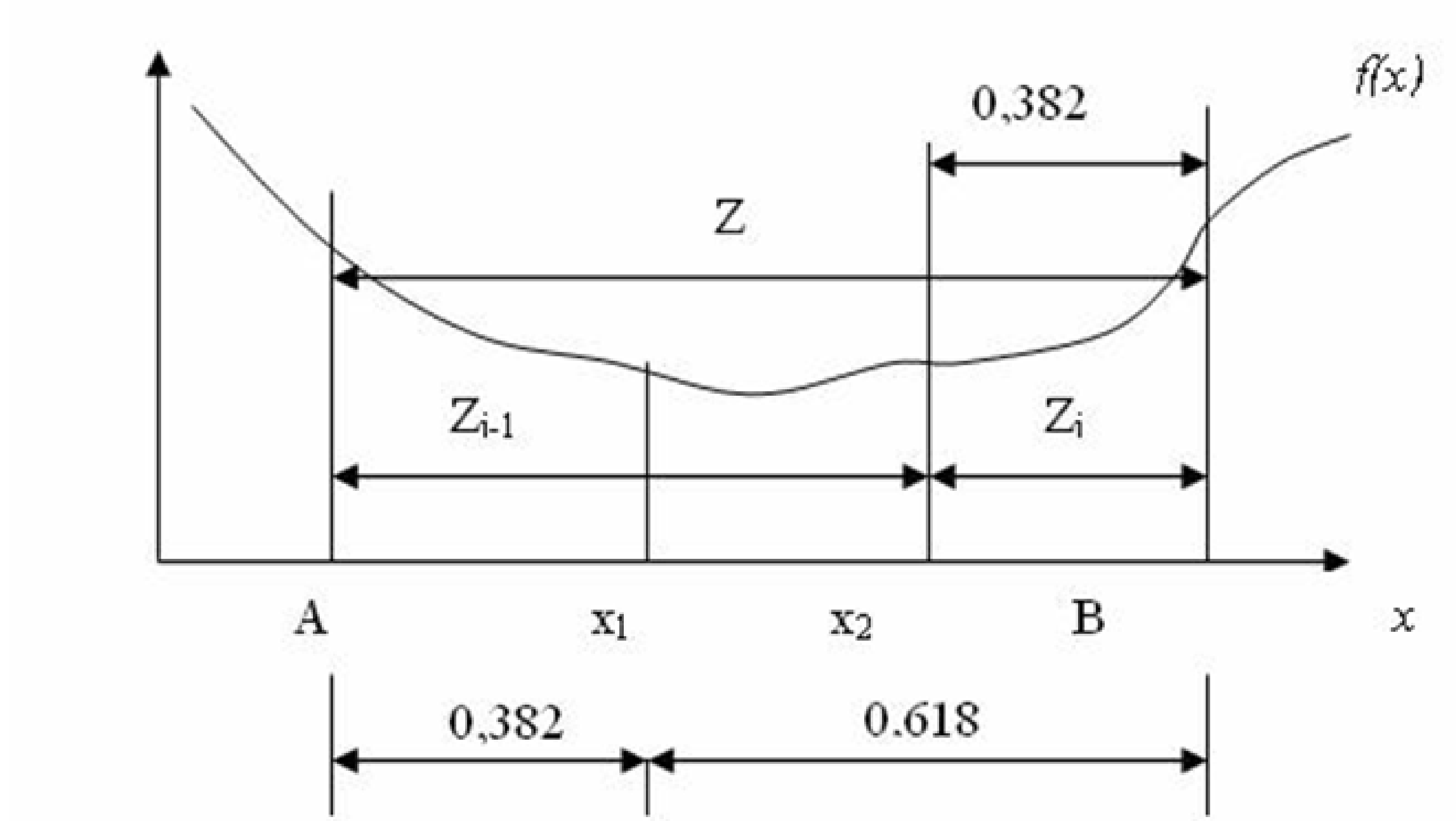
6. Перейти до п.2.

Продовжуємо описаний алгоритм доти, поки інтервал невизначеності $l<\varepsilon$.

МЕТОД ЗОЛОТОГО ПЕРЕТИНУ



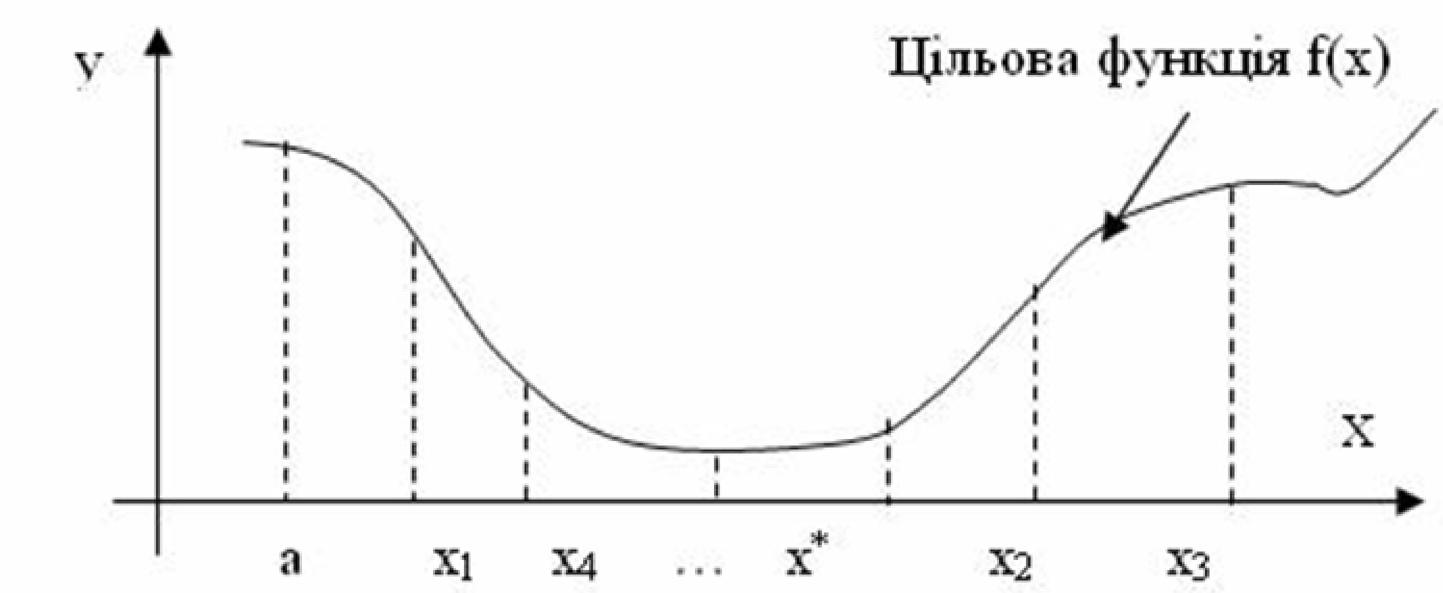
Приклад 1. *Знайти мінімум функції* $f(x) = -e^{-x}\ln(x)$ *на інтервалі* $[0,2]$.



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

Метод Фібоначчі

Задано $[x_1, x_3]$, $f(x_2)$

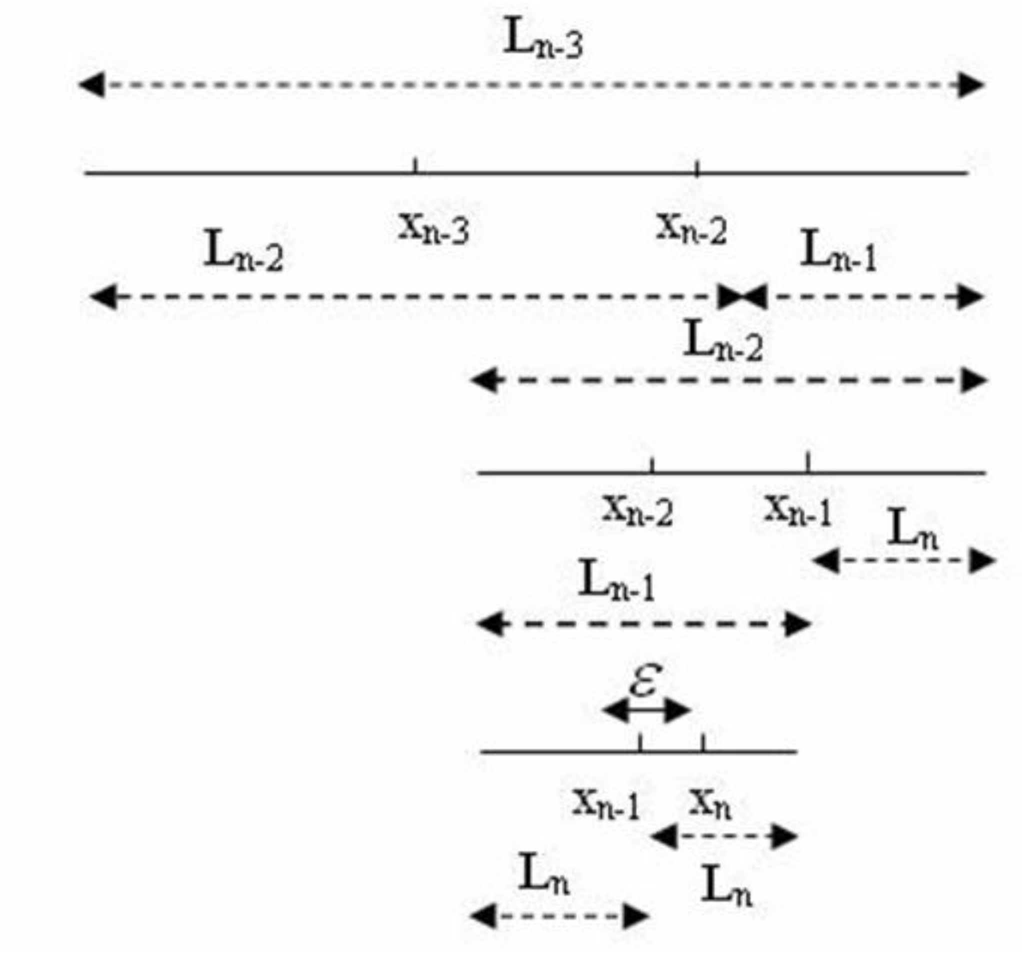


Припустимо $x_2 - x_1 = L$ і $x_3 - x_2 = R$, $L > R$.

Якщо x_4 знаходиться на інтервалі (x_1, x_2), то:

1. Якщо $f(x_4) < f(x_2)$, то новим інтервалом невизначеності буде (x_1, x_2) довжиною $x_2 - x_1 = L$.
2. Якщо $f(x_4) > f(x_2)$, те новим інтервалом невизначеності буде (x_4, x_3) довжиною $x_3 - x_4$.

Геометрична інтерпретація ітераційного процесу Фібоначчі



$$L_{n-1} = 2L_n - \varepsilon,$$
$$L_{n-2} = L_{n-1} + L_n = 3L_n - \varepsilon,$$
$$L_{n-3} = L_{n-2} + L_{n-1} = 5L_n - 2\varepsilon,$$
$$L_{n-4} = L_{n-3} + L_{n-2} = 8L_n - 3\varepsilon \text{ і т.д.}$$

Послідовність чисел Фібоначчі:

$$F_0 = 1,\ F_1 = 1, \quad F_k = F_{k-1} + F_{k-2}$$

для $k = 2,3,\ldots,$ *тоді*

$$L_{n-j} = F_{j+1}L_n - F_{j-1}\varepsilon, \quad j = 1,2,\ldots,n-1.$$

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

$$L_1 = F_n L_n - \varepsilon \cdot F_{n-2}.$$

Тобто

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

Метод Фібоначчі, названий так через появу при пошуку чисел Фібоначчі, є ітераційною процедурою.

В процесі пошуку інтервалу (x_1, x_2) *з точкою* x_2, *що вже лежить в цьому інтервалі, наступна точка* x_4 *завжди вибирається такою, що*

$$x_3 - x_4 = x_2 - x_1 \text{ або } x_4 - x_1 = x_3 - x_2,$$

тобто

$$x_4 = x_1 - x_2 + x_3.$$

ПОРІВНЯННЯ МЕТОДІВ ОДНОВИМІРНОГО ПОШУКУ

Ефективність алгоритму *- число обчислень функції, необхідне для досягнення необхідного звуження інтервалу невизначеності.*

Найкращий - метод Фібоначчі, (зол. перетину).

найгірший – метод загального пошуку.

Універсальність алгоритму *- можливість його застосування для розв'язку різноманітних задач.*

Метод Фібоначчі поступається іншим.

Висновок: *не існує універсального алгоритму, який дозволяв би розв'язувати будь-які задачі.*

Багатовимірна оптимізація

Зміст задачі багатовимірної оптимізації - пошук мінімуму (максимуму) функції від n аргументів:

$$F(x_1, x_2, \dots, x_n) \longrightarrow \min . \qquad (1)$$

Якщо на змінні x_1, x_2, …, x_n не накладається ніяких обмежень, то така задача називається ***класичною задачею оптимізації*** *(без обмежень).*

Будемо вважати, що цільова функція є гладкою і унімодальною. У шуканій точці мінімуму усі часткові похідні дорівнюють нулю:

$$\frac{\partial F}{\partial x_i}(x_1, x_2, \dots, x_n) = 0 \quad , \quad i = \overline{1, m}. \qquad (2)$$

Таким чином, задача зводиться до задачі розв'язування системи нелінійних рівнянь.

Метод Ньютона:

$$\vec{x}_{k+1} = \vec{x}_k + \vec{\Delta}_k ,$$

$$R(\vec{x}_k) \cdot \vec{\Delta}_k = -\nabla F(\vec{x}_k) ,$$

де k *- номер ітерації,* $\vec{\Delta}_k$ *- поправка, яка знаходиться з СЛР.*

$$R(\vec{x}_k) = \begin{pmatrix} \frac{\partial^2 F(\vec{x}_k)}{\partial x_1 \partial x_1} & \frac{\partial^2 F(\vec{x}_k)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F(\vec{x}_k)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F(\vec{x}_k)}{\partial x_2 \partial x_1} & \frac{\partial^2 F(\vec{x}_k)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 F(\vec{x}_k)}{\partial x_2 \partial x_n} \\ \cdots & & & \\ \frac{\partial^2 F(\vec{x}_k)}{\partial x_n \partial x_1} & \frac{\partial^2 F(\vec{x}_k)}{\partial x_n \partial x_2} & & \frac{\partial^2 F(\vec{x}_k)}{\partial x_n \partial x_n} \end{pmatrix},$$

Критерій закінчення обчислень:

$$|\vec{\Delta}_k| < \varepsilon .$$

Дякую за увагу!