

# Лекция 9

1. Основные понятия и определения задачи оптимизации
2. Аналитические и численные методы решения задачи безусловной одномерной оптимизации
3. Методы сканирования (прямого перебора)

# Задача оптимизации. Проектные параметры

**Оптимизация** – это процесс выбора наилучшего варианта из всех возможных. В инженерных расчетах методы оптимизации позволяют выбрать наилучший вариант конструкции, распределения ресурсов и т.п. В экономических расчетах – получить максимальную прибыль, добиться минимальных затрат и т.п.

В процессе решения задачи оптимизации необходимо найти оптимальные значения **проектных параметров** (параметров плана), определяющих данную задачу:  $x_1, x_2, \dots, x_n$ . Содержательный смысл этих параметров зависит от конкретной задачи. Это могут быть линейные размеры, масса, температура и тому подобные параметры оптимизируемого объекта.

# Задача оптимизации. Целевая функция

Выбор оптимального решения или сравнение альтернатив производится с помощью **целевой функции**  $f(x_1, x_2, \dots, x_n)$ , зависящей от проектных параметров. **Решение задачи оптимизации заключается в отыскании таких значений проектных параметров, при которых достигается минимум или максимум целевой функции.** Примерами целевых функций могут служить мощность установки, прочность конструкции, объем выпуска продукции, стоимость перевозки грузов, прибыль и т.п.

Геометрически целевая функция представляет собой поверхность в  $(n+1)$ -мерном пространстве. В частности, при  $n=1$  это кривая на плоскости  $y=f(x)$ ; при  $n=2$  – поверхность в 3-мерном пространстве  $y = f(x_1, x_2)$ .

# Безусловная и условная оптимизация

Существует **два типа задач оптимизации**: безусловные и условные.

**Безусловная оптимизация** – это отыскание минимума (максимума) функции и определение соответствующих значений аргументов в  $n$ -мерном пространстве без каких-либо ограничений на значения аргументов. Обычно рассматриваются задачи минимизации, так как задача отыскания максимума легко сводится к задаче минимизации путем замены знака целевой функции на противоположный.

При **условной оптимизации** задача имеет **ограничения**, определяющие множество  $S$  в  $n$ -мерном пространстве, в пределах которого ищется оптимальное решение. Эти ограничения задаются совокупностью некоторых функций  $g_i(x_1, x_2, \dots, x_n)$ ;  $i=1, 2, \dots, m$ , удовлетворяющих уравнениям или неравенствам:

$$g_i(x_1, x_2, \dots, x_n) \geq 0; i=1, 2, \dots, m$$

Теория и методы решения задач условной оптимизации – предмет исследования важного раздела прикладной математики – **математического программирования**.

# Пример постановки задачи оптимизации

Рассмотрим пример постановки задачи оптимизации. Пусть контейнер в форме прямоугольного параллелепипеда объемом  $V=1\text{м}^3$  должен иметь оптимальную поверхность, чтобы на его изготовление уходило как можно меньше материала. При постоянной толщине стенок это означает, что площадь полной поверхности параллелепипеда должна быть минимальной.

Обозначим длины ребер контейнера  $x_1$ ,  $x_2$ ,  $x_3$ . Тогда площадь контейнера

$$S = 2(x_1x_2 + x_2x_3 + x_1x_3)$$

Так как  $V = x_1x_2x_3 = 1$ , то  $x_3 = 1/(x_1x_2)$  и

$$S = 2(x_1x_2 + 1/x_1 + 1/x_2)$$

Таким образом, задача оптимизации свелась к задаче безусловной минимизации функции двух переменных, то есть отыскания значений  $x_1$  и  $x_2$ , при которых достигается

$$\min S = 2(x_1x_2 + 1/x_1 + 1/x_2)$$

# Пример постановки задачи оптимизации

Пусть требуется, чтобы контейнер при этом имел длину не менее **2 м**. Тогда задача становится задачей условной минимизации: отыскать такие значения  $x_1 > 0$  и  $x_2 > 0$ , при которых достигается

$$\min S = 2(x_1x_2 + 1/x_1 + 1/x_2)$$

с учетом ограничения  $x_1 \geq 2$ .

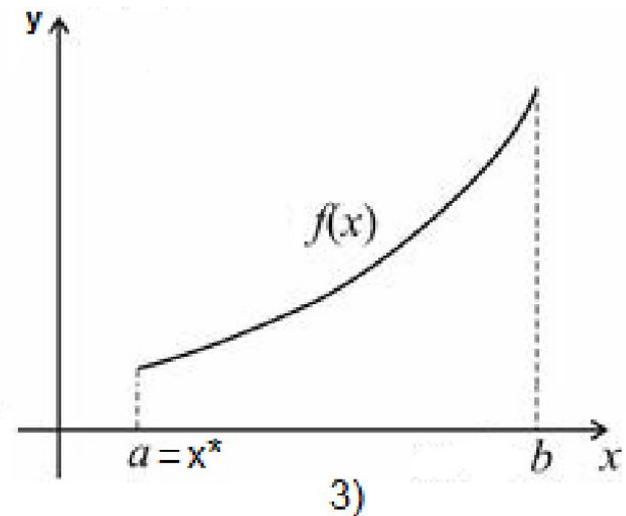
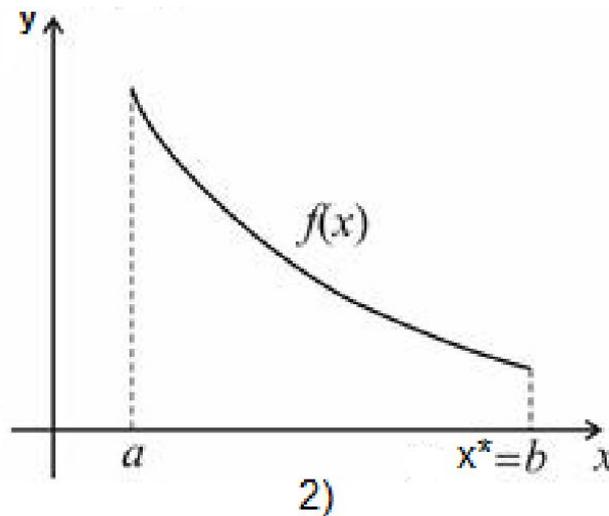
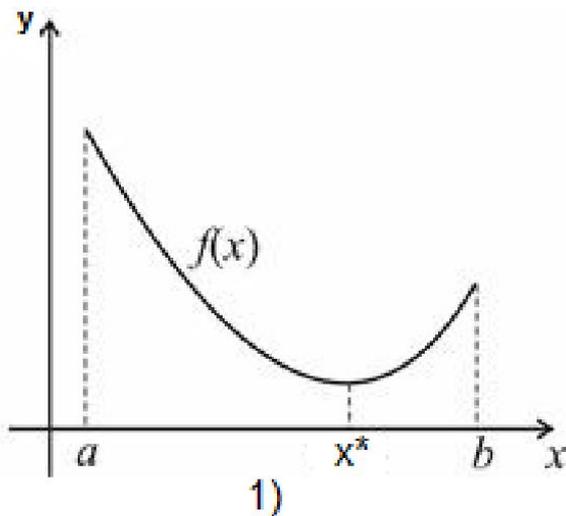
# Локальные и глобальный минимумы



Задачей одномерной оптимизации является нахождение точек локального минимума и соответствующих им значений функции. В случаях, когда требуется найти глобальный минимум, отыскиваются все локальные минимумы и выбирается наименьший из них.

# Унимодальные функции

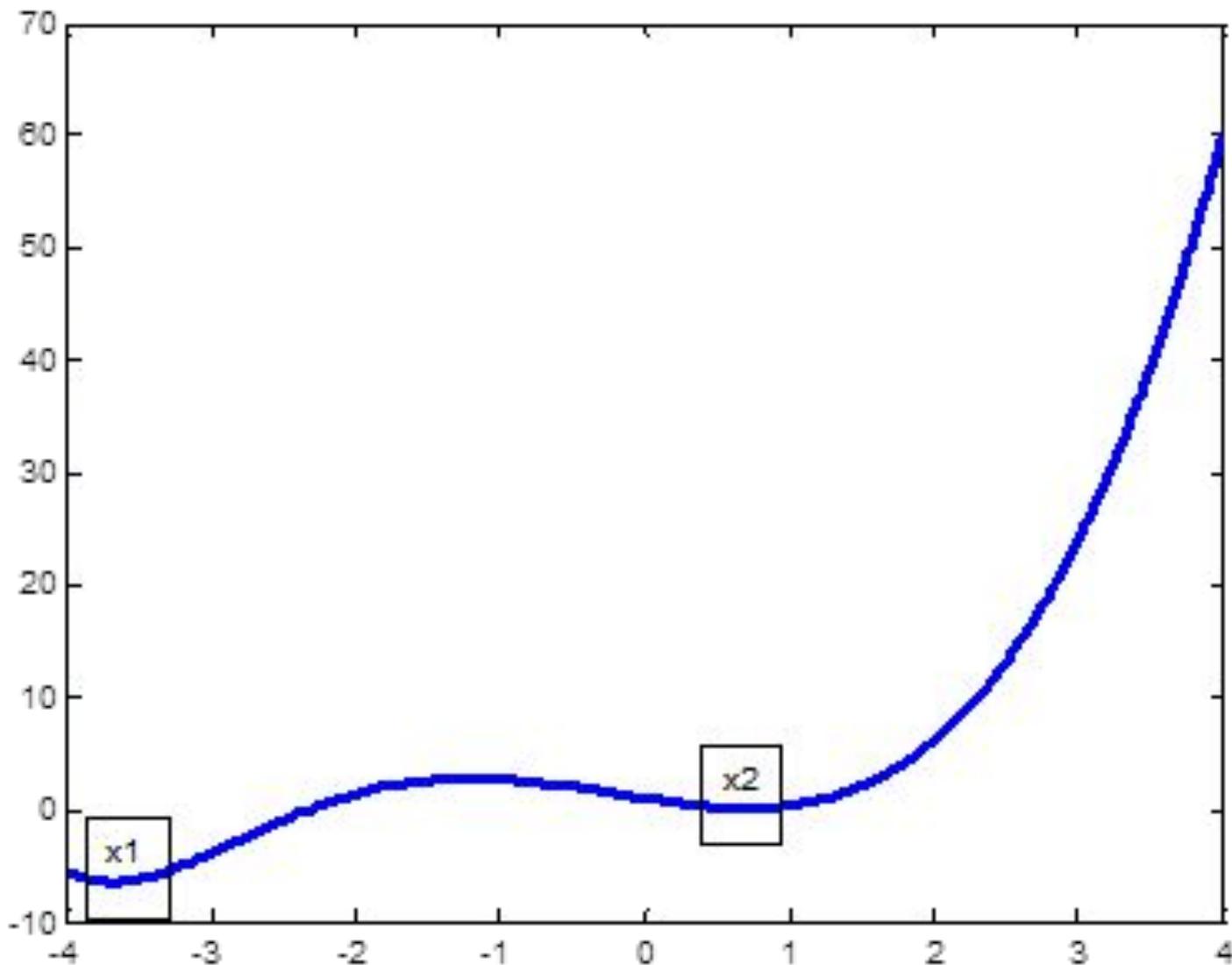
Отрезок, на котором локализован единственный минимум, называется **отрезком неопределенности**. Все численные методы одномерной оптимизации применяются только для функций, **унимодальных** на отрезке неопределенности. Функция  $f(x)$  называется унимодальной на отрезке  $[a;b]$ , если она имеет на этом отрезке единственный минимум в точке  $x^*$ , и  $f(x)$  строго возрастает при  $x < x^*$  и  $x > x^*$ . При этом возможны три случая.



# Условия унимодальности функции

Обычно при решении задачи одномерной оптимизации речь идет о поиске единственного экстремума функции. В этом случае **необходимым условием унимодальности** функции и существования экстремума является **неубывание первой производной  $f'(x)$**  на отрезке неопределенности  **$[a;b]$** . **Достаточным условием** является положительный знак второй производной  **$f''(x) > 0$**  на отрезке

# График функции $f(x) = x^3 - x + e^{-x}$



# Пример проверки условий унимодалности

Пусть, например, требуется отыскать локальные минимумы функции  $f(x) = x^3 - x + e^{-x}$ . График функции приведен на слайде 10. Из него видно, что  $f(x)$  имеет два локальных минимума:  $x_1$  и  $x_2$ , причем  $x_1$  – точка глобального минимума. Исходя из графика, можно выделить следующие отрезки неопределенности: для точки  $x_1$  –  $[-4; -3]$ , а для точки  $x_2$  –  $[0; 1]$ .

Проверим условия унимодалности функции на выбранных отрезках. Первая производная  $f'(x) = 3x^2 - 1 - e^{-x}$ . Вторая производная  $f''(x) = 6x + e^{-x}$ . На отрезке  $[-4; -3]$  первая производная  $f'(x)$  возрастает от  $-7.6$  до  $+5.9$ , обращаясь в  $0$  в точке минимума внутри отрезка. Вторая производная  $f''(x) > 0$  на всем отрезке, изменяясь от  $+1$  до  $+2.1$ .

На отрезке  $[0; 1]$  первая производная  $f'(x)$  возрастает от  $-2$  до  $+1.6$ , обращаясь в  $0$  в точке минимума внутри отрезка. Вторая производная  $f''(x) > 0$  на всем отрезке, изменяясь от  $+1$  до  $+6.4$ .

Таким образом, условия унимодалности и существования минимума функции выполняются на обоих отрезках неопределенности.

# Аналитический метод отыскания локального минимума

Аналитический метод отыскания локального минимума функции  $f(x)$  требует решения уравнения  $f'(x) = 0$ . Для этого необходимо, чтобы:

- 1)  $f(x)$  была задана аналитически;
- 2)  $f(x)$  была дифференцируема на отрезке  $[a;b]$ ;
- 3) уравнение  $f'(x) = 0$  имело аналитическое решение.

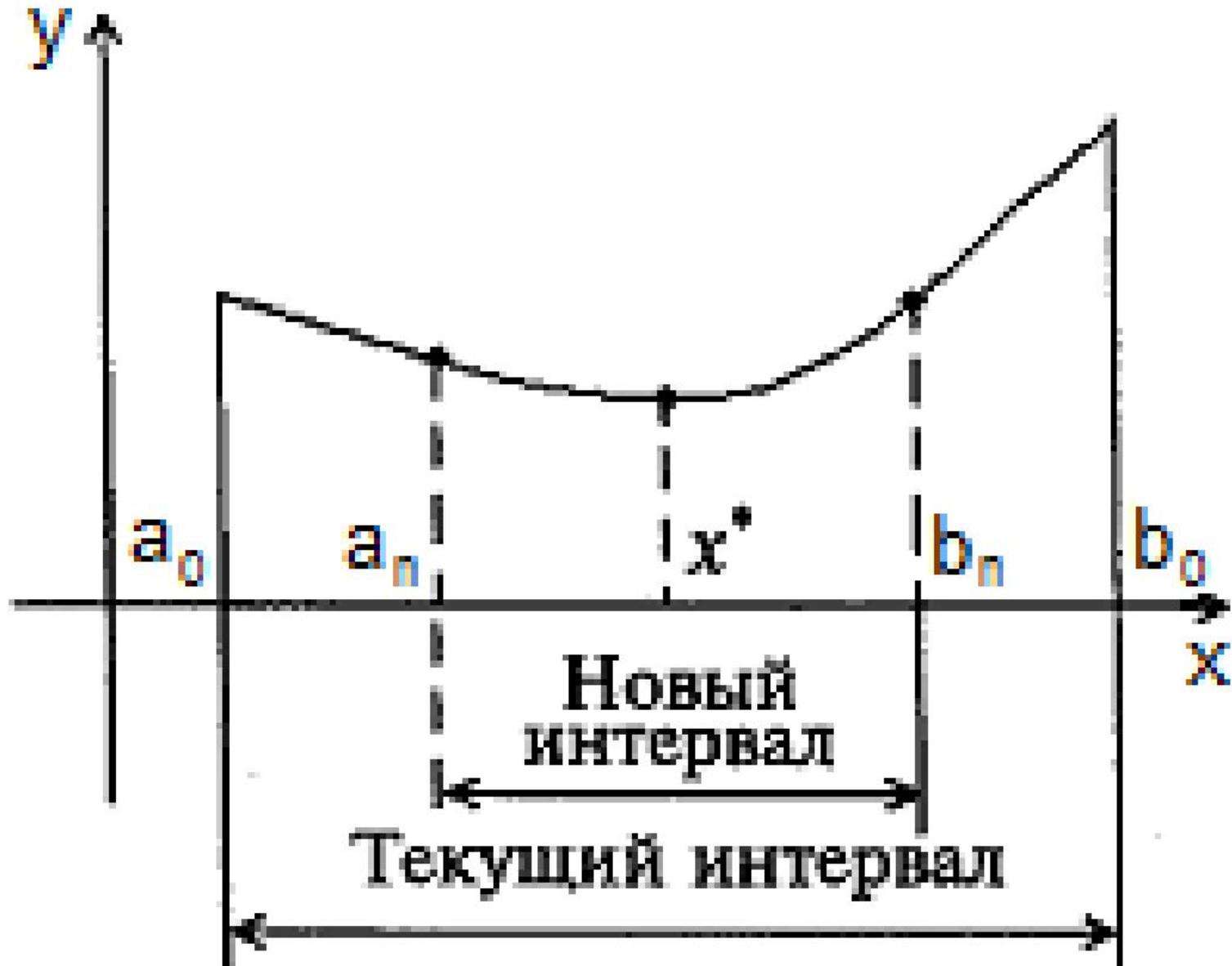
Даже если выполняются первые два условия, то, в большинстве случаев, уравнение  $f'(x) = 0$  является нелинейным и может быть решено только численными методами. Поэтому имеет смысл использовать численные методы непосредственно для решения задачи минимизации.

# Методы поиска

Для численного решения задачи безусловной одномерной оптимизации используются различные **методы поиска**. Их сущность состоит в последовательном сужении отрезка неопределенности. Вначале длина отрезка равна  $b_0 - a_0$ , а в итоге она должна стать меньше заданной допустимой погрешности решения  $\varepsilon$ , так что  $x^* \in [a_n; b_n]$ , причем  $b_n - a_n < \varepsilon$ .

За приближенное значение точки минимума  $\ddot{x}$  может быть принята любая точка отрезка  $[a_n; b_n]$ , при этом  $|\ddot{x} - x^*| \leq b_n - a_n < \varepsilon$ . Обычно принимают  $\ddot{x} = (a_n + b_n)/2$ , то есть середину отрезка.

# Методы поиска



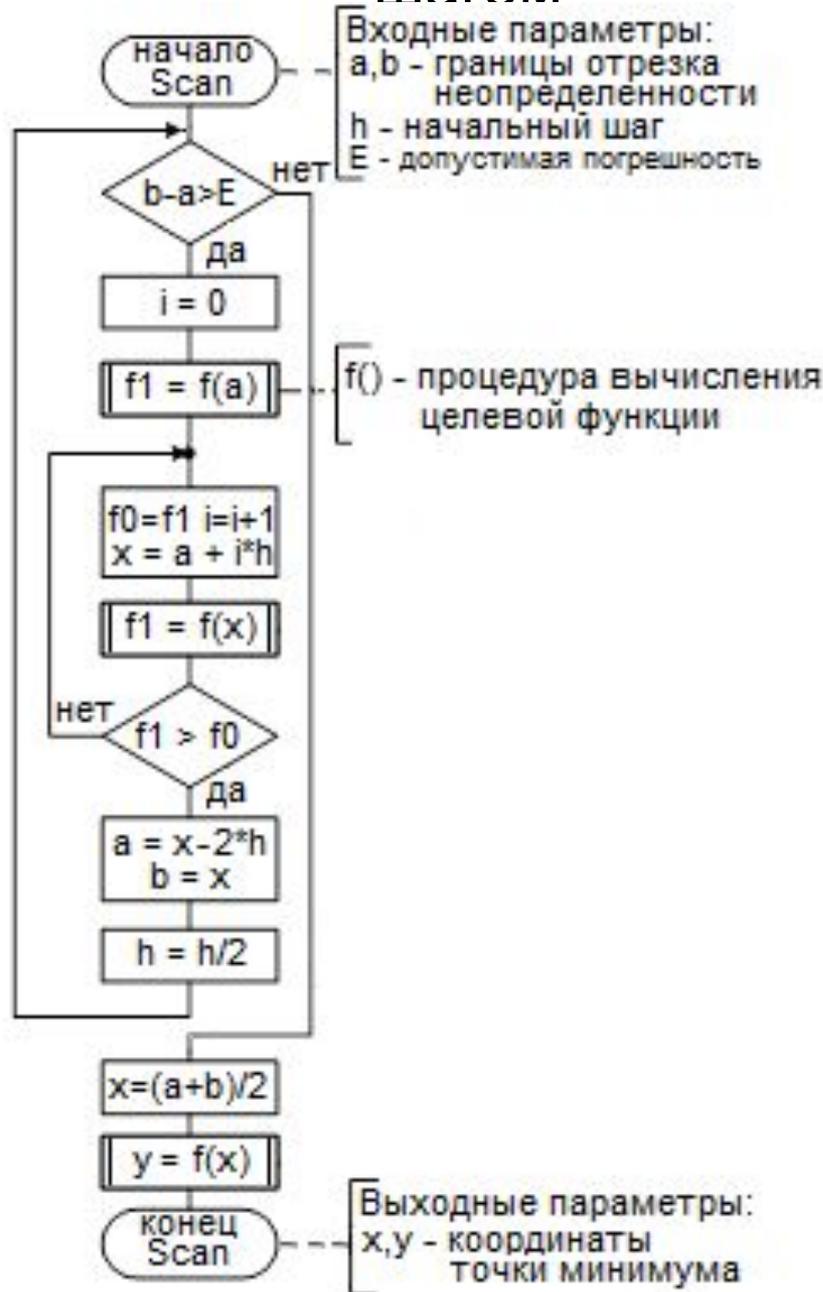
# Методы сканирования (прямого перебора)

Простейшим методом поиска является метод *сканирования* или *прямого перебора*. Отрезок неопределенности  $[a;b]$  разбивают на  $n$  частей, где  $n = (b-a)/h$ , и  $h \leq \varepsilon$ . Последовательно вычисляя  $f(x=a+i \cdot h)$ ,  $i=0, 1, \dots, n$ , находят точку наименьшего значения функции.

Метод сканирования весьма трудоемок. На практике чаще применяют одну из основных модификаций метода – *метод прямого перебора с переменным шагом*. Суть его заключается в следующем. От начальной точки отрезка неопределенности  $a$  двигаются с некоторым начальным шагом  $h_0$  до тех пор, пока функция в точках разбиения убывает. Если функция в очередной точке стала возрастать, то происходит сужение интервала неопределенности путем возврата от этой точки, которая станет правой границей нового отрезка, на два шага назад. Полученная таким образом точка станет левой границей нового отрезка. Новый отрезок вновь сканируют таким же образом, но уже с уменьшенным в два раза шагом. Процесс повторяется, пока длина очередного отрезка неопределенности  $b_n - a_n$  не станет меньше заданной допустимой погрешности  $\varepsilon$ .

# Схема алгоритма метода прямого перебора с переменным

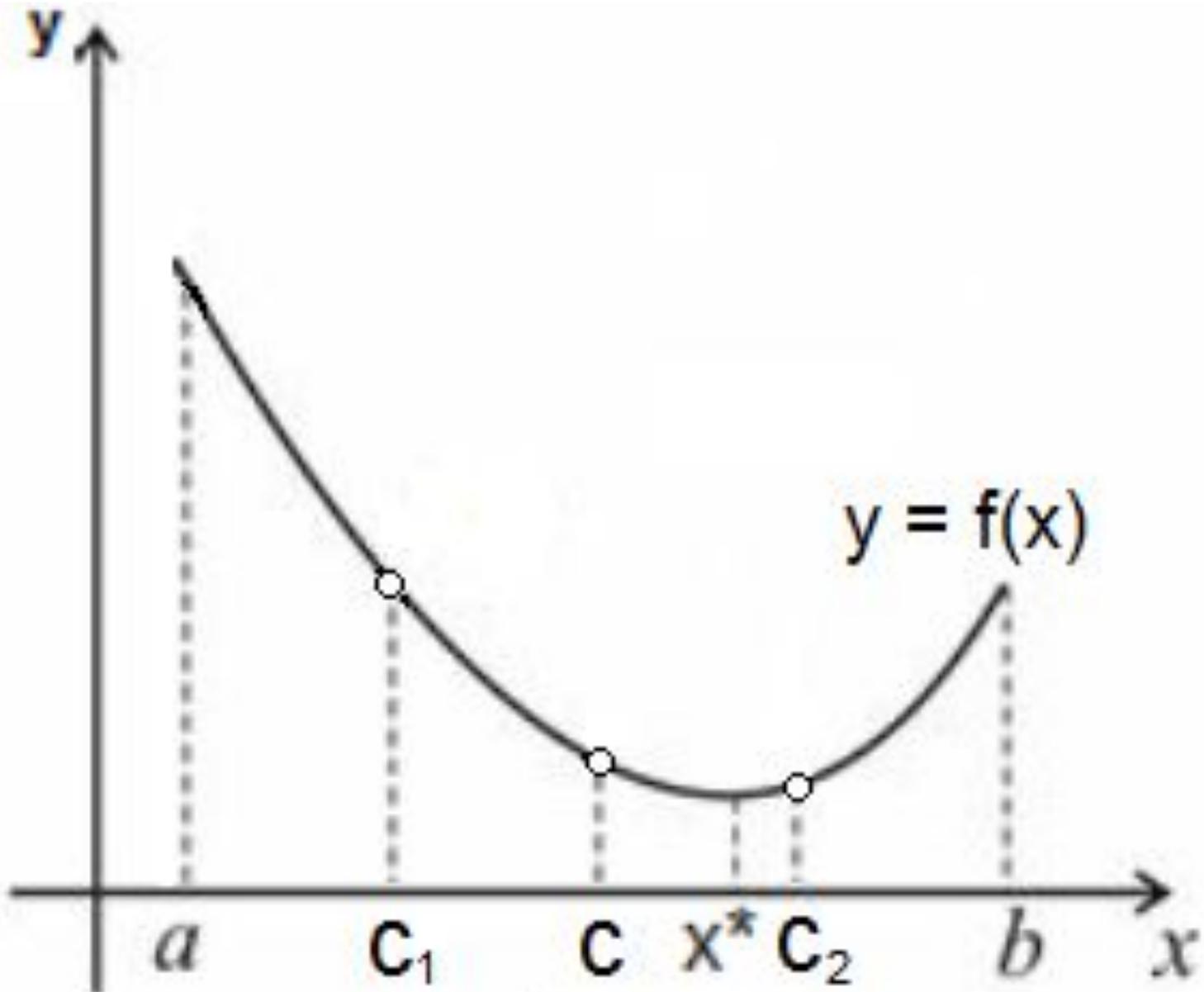
## ШАГОМ



# Методы последовательного поиска

Несмотря на меньшую трудоемкость вычислений методом прямого перебора с переменным шагом по сравнению с обычным методом сканирования, он все равно остается очень трудоемким. Для сокращения количества вычислений целевой функции применяют методы *последовательного поиска*, в которых строится последовательность отрезков  $[a_0; b_0]$ ,  $[a_1; b_1]$ , ..., стягивающихся к точке минимума  $x^*$ . При этом интервал неопределенности постепенно сужается и в конце концов становится меньше допустимой погрешности.

# Метод деления отрезка пополам



# Сущность метода деления отрезка пополам

Сначала находится середина начального отрезка  $[a;b]$  – точка  $c=(a+b)/2$ , и вычисляется значение функции в этой точке  $f(c)$ . Затем отрезок  $[a;c]$  делится пополам точкой  $c_1=(a+c)/2$ , и вычисляется значение функции  $f(c_1)$ .

Если  $f(c_1) < f(c)$ , то точка минимума  $x^* \in [a;c]$ , и правая граница нового отрезка переносится в точку  $c$ :  $b = c$ .

В противном случае пополам делится отрезок  $[c;b]$ , находится точка  $c_2=(c+b)/2$  и значение функции в этой точке  $f(c_2)$ .

Если  $f(c_2) < f(c)$ , то точка минимума  $x^* \in [c;b]$ , и уже левая граница нового отрезка переносится в точку  $c$ :  $a = c$ .

В противном случае (как на слайде 18) точка минимума  $x^* \in [c_1; c_2]$ , и переносятся обе границы нового отрезка:  $a = c_1$ ,  $b = c_2$ . Затем такая же процедура выполняется на новом отрезке, и так далее, пока длина отрезка неопределенности не станет меньше допустимой погрешности приближения.

При любом варианте сужения на каждой итерации интервал неопределенности сужается вдвое. Количество вычислений функции на каждой итерации, кроме первой, равно одному или двум.

# Свойства метода деления отрезка пополам

При любом варианте сужения на каждой итерации интервал неопределенности сужается вдвое. Количество вычислений функции на каждой итерации, кроме первой, равно одному или двум.

Нетрудно определить количество итераций, необходимое для сужения отрезка до необходимой величины. Длину отрезка неопределенности после  $n$  итераций  $\Delta_n$  можно вычислить по формуле:

$$\Delta_n = \frac{b-a}{2^n}$$

Чтобы выполнялось условие  $\Delta_n < \varepsilon$ , необходимо, чтобы

$$n \geq \log_2 \frac{b-a}{\varepsilon}$$

Например, если начальная длина отрезка равна 1, а допустимая погрешность  $\varepsilon = 10^{-6}$ , то  $n$  оказывается равным 20.

# Схема алгоритма метода деления отрезка

## ПОПОПАМ

