Свиридова О.В.

Аналитические и численные методы решения одномерных задач оптимизации

Волжский

2018

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ВОЛЖСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ (ФИЛИАЛ)
ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО
УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОЛГОГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

О.В. Свиридова

Аналитические и численные методы решения одномерных задач оптимизации

Электронное учебно-методическое пособие

2018

УДК 004(07)
ББК 32.81я73
С 247

Рецензенты:
заведующий кафедрой математических методов в экономике, информационных и сервисных технологий Волгоградского института бизнеса, канд. физ.-мат. наук, доцент
Алпатов А.В.,
доцент кафедры математики и информатики Волгоградского государственного медицинского университета, канд. социологических наук
Плешакова Е.О.

Издается по решению редакционно-издательского совета Волгоградского государственного технического университета

Свиридова, О.В
Аналитические и численные методы решения одномерных задач оптимизации [Электронный ресурс] : учебно-методическое пособие / О.В. Свиридова ; ВПИ (филиал) ВолгГТУ. - Электрон. текстовые дан. (1 файл: 540 КБ). – Волжский, 2018. – Режим доступа: http://lib.volpi.ru. – Загл. с титул. экрана.
ISBN 978-5-9948-2818-2

В учебно-методическом пособии приведены теоретические сведения об аналитических и численных методах решения одномерных задач оптимизации, а также примеры решения указанных задач, которые могут быть использованы студентами при подготовке к практическим занятиям и при выполнении контрольной работы.

Предназначено для студентов, обучающихся по направлению 09.03.01 «Информатика и вычислительная техника» в рамках курса «Методы оптимизации».

Ил. 26, табл. 4, библиограф.: 5 назв.

ISBN 978-5-9948-2818-2

Содержание

ВВЕДЕНИЕ

Задача оптимизации, в которой целевая функция задана функцией одной переменной, относится к наиболее простому типу оптимизационных задач. Это связано не только с тем, что именно такие задачи обычно решаются в инженерной практике, но и с тем, что одномерные методы оптимизации часто используются для анализа подзадач, которые возникают при реализации итеративных процедур, ориентированных на решение многомерных задач оптимизации.

Для решения задачи минимизации функции $f(\mathbf{x})$ на отрезке $[a; b]$ на практике, как правило, чаще применяют приближенные (численные) методы, чем аналитические. Они позволяют найти решение этой задачи с необходимой точностью в результате определения конечного числа значений функции $f(x)$ и ее производных в некоторых точках отрезка $[a; b]$.

В первой части учебно-методического пособия приведены аналитические методы решения задачи оптимизации с использованием классического подхода. Во второй части пособия описаны численные методы решения с разделением их на прямые методы и методы, использующие производные функции.

1. АНАЛИТИЧЕСКИЕ МЕТОДЫ РЕШЕНИЯ

1.1. Основные понятия

Аналитический подход к задаче определения локальных и глобальных минимумов состоит в использовании методов математического анализа для поиска уравнений, которым должны удовлетворять эти точки, и для решения этих уравнений.

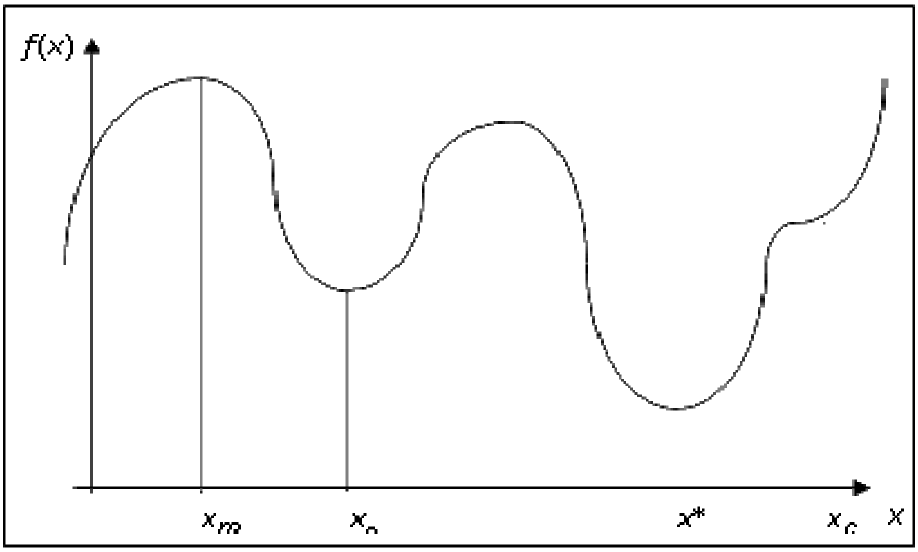

Рис. 1. Стационарные точки непрерывной функции f(x)

Из рисунка 1 видно, что в точках x_m, x_n, x^* и x_0 касательная к графику функции параллельна оси OX, а это означает, что производная функции в этих точках равна нулю. Следовательно, x_m, x_n, x^* и x_0 являются решениями уравнения $f'(x)=0$.

Однако это же справедливо и для точки максимума x_m, и точек минимума x_n, x^* , и для точки перегиба x_0. Таким образом, найденное уравнение является необходимым условием экстремума, но не является достаточным.

В точках x_n и x^* производная $f'(x)$ меняет знак с отрицательного на положительный, в x_m – с положительного на отрицательный, в точке x_0 производная знак не меняет. Следовательно, в точке минимума производная является возрастающей функцией. Степень же возрастания измеряется второй производной, то есть в нашем случае $f''(x_m)<0$, $f''(x_n)>0$, $f''(x^*)>0$. Однако если $f''(x_0)=0$, то ситуация остается неопределенной.

Теорема 1 (необходимое и достаточное условие существования экстремума функций одной переменной). Если функция $f(x)$ и ее производные непрерывны, то точка x^* является точкой экстремума (максимума или минимума) тогда и только тогда, когда порядок ее первой, не обращающейся в ноль в точке x^*, производной есть четное число. При этом, если $f^{(k)}(x^*)<0$, то x^* - точка максимума, если $f^{(k)}(x^*)>0$, то x^* - точка минимума.

Таким образом, при классическом подходе для поиска минимума функции одной переменной необходимо решить уравнение $f'(x)=0$ и установить знак в полученных точках. Аналитическое решение такого уравнения в общем случае невозможно, поэтому используются методы приближенного решения, известные из математического анализа (методы Ньютона, бисекций, и т.д.).

В большинстве случаев задачу оптимизации $f(x)\rightarrow min$ не удается решить, опираясь на необходимые и достаточные условия оптимальности или на геометрическую интерпретацию задачи, и приходится ее решать численно с применением вычислительной техники. Причем наиболее эффективными оказываются методы, разработанные специально для решения конкретного класса задач оптимизации, так как они позволяют полнее учесть ее специфику.

Любой численный метод имеет два этапа:

первый этап любого численного метода (алгоритма) решения задачи оптимизации основан на точном или приближенном вычислении ее характеристик (значений целевой функции; значений функций, задающих допустимое множество, а также их производных);

во втором этапе, на основании полученной информации строится приближение к решению задачи – искомой точке минимума или, если такая точка не единственна, к множеству точек минимума:

$$f^* = \min_{x \in X} f(x)$$

Иногда, если только это требуется, строится и приближение к минимальному значению целевой функции. Для каждой конкретной задачи вопрос о том, какие характеристики следует выбрать для вычисления, решается в зависимости от свойств минимизируемой функции, ограничений и имеющихся возможностей по хранению и обработки информации.

В зависимости от того, какие характеристики, в частности, целевой функции берутся, алгоритмы делятся на алгоритмы:

нулевого порядка – в них используется информация только о значениях минимизируемой функции;

первого порядка – использующие информацию также и о значениях первых производных;

второго порядка – использующие, кроме того, информацию о вторых производных;

и так далее.

Когда решен вопрос о том, какие именно характеристики решаемой задачи следует вычислять, то для задания алгоритма достаточно указать способ выбора точек вычисления.

В зависимости от способа выбора точек вычисления алгоритмы делятся на пассивные и активные (последовательные).

В пассивных алгоритмах все точки выбираются одновременно до начала вычислений.

В активных (последовательных) алгоритмах точки вычисления выбираются поочередно, то есть точка выбирается, когда уже выбраны точки предыдущих вычислений, и в каждой из них произведены предусмотренные алгоритмом вычисления, результаты которых будем обозначать соответственно через $\tilde{x}^{i+1}$.

Таким образом, последовательный алгоритм определяется точкой $\tilde{x}^{i+1}$

и набором отображений вида

$$\tilde{x}^{i-1} : \{x^1, \ldots, x^i; y^1, \ldots, y^i\} \to X,$$

при этом

$$x^{i-1} = \tilde{x}^{i-1}\{x^1, \ldots, x^i; y^1, \ldots, y^i\}.$$

На практике обычно используются наиболее простые виды зависимости, например:

$$\tilde{x}^{i-1}\{x^1, \ldots, x^i; y^1, \ldots, y^i\} = \tilde{x}^{i-1}\{x^i; y^i\},$$

то есть выбор точки очередного вычисления зависит лишь от точки предыдущего вычисления и полученного результата или

$$\tilde{x}^{i-1}\{x^1, \ldots, x^i; y^1, \ldots, y^i\} = \tilde{x}^{i-1} : \{x^i; \lambda_1 * y^1 + \ldots + \lambda_i * y^i\},$$

то есть выбор зависит от линейной комбинации всех полученных результатов (например, в методе сопряженных градиентов).

При этом конкретный алгоритм определяется:

заданием начальной точки;

правилами выбора векторов и чисел на основе полученной в результате вычислений информации;

условием остановки.

Вектор определяет направление *i*-го шага метода минимизации, а коэффициент λ_i – длину этого шага.

Обычно название метода минимизации определяется способом выбора точки $\tilde{x}^{i+1}$.

Наряду с термином <u>шаг метода</u> будем пользоваться также термином <u>итерация метода</u>.

Среди методов минимизации можно условно выделить:

конечношаговые методы;

бесконечношаговые методы.

Конечношаговыми (или конечными) называются методы, гарантирующие отыскание решения задачи за <u>конечное число шагов</u>.

Для бесконечношаговых методов достижение решения гарантируется лишь в пределе.

Сходимость методов оптимизации.

Важной характеристикой бесконечношаговых методов является <u>сходимость</u>.

Метод <u>сходится</u> если $x^k \to x^*$ при $k \to \infty$, где x^* – решение задачи .

Если $f(x^k) \to f(x^*)$ при $k \to \infty$, то иногда также говорят, что метод сходится (по функции), при этом последовательность $\{x^k\}$ называют <u>минимизирующей</u>.

1.2. Постановка задачи одномерной оптимизации

Пусть функция $f(x)$ определена на некотором множестве U вещественной оси $E^1=\{x:\ -\infty<x<\infty\}$. Поскольку максимизация целевой функции ($f(x)\to max$) эквивалентна минимизации противоположной величины ($-f(x)\to min$), будем рассматривать только задачу минимизации функции $f(x)$ на множестве U. Для этого напомним некоторые основные понятия.

Определение 1. Число x^* называется ***точкой глобального (абсолютного) минимума*** или просто ***точкой минимума*** функции $f(x)$ на множестве U, если $f(x^*)\le f(x)$ для всех $x\in U$. Значение $f^*=f(x^*)=min\{f(x)\mid x\in U\}$ называется ***глобальным*** (***абсолютным***) ***минимумом*** или просто ***минимумом функции*** $f(x)$ на U. Множество всех точек минимума функции $f(x)$ на U будем обозначать через U^*.

Определение 2. Число $\tilde{x}\in U$ называется ***точкой локального минимума*** функции $f(x)$ на U, если $f(\tilde{x})\le f(x)$ для всех $x\in U$, достаточно близких к $\tilde{x}$, т.е. если существует число $\varepsilon>0$ такое, что это неравенство выполняется для любого $x \in U \cap \{x: |x-\tilde{x}| < \varepsilon\} = U_\varepsilon(\tilde{x})$.

В зависимости от свойств множества U и функции $f(x)$ множество U^* может содержать одну, несколько или даже бесконечно много точек. Также возможны случаи, когда U^* пусто.

Пример 1. Пусть $f(x) = \sin^2\left(\frac{\pi}{x}\right)$ при $x \neq 0$ и $f(0)=0$. На множестве $U=\{x: 1 \leq x \leq 2\}$ минимальное значение $f(x)$ равно нулю, а множество U^* состоит из единственной точки $x^*=1$. Если $U=\{x: 1/3 \leq x \leq 1\}$, то U^* содержит три точки $\frac{1}{3}, \frac{1}{2}, 1$. В случае $U=\{x: 2 \leq x < +\infty\}$ функция $f(x)$ не имеет наименьшего значения на U. В самом деле, какую бы точку $x \in U$ мы не взяли, найдется точка из U такая, что значение в ней будет меньше $f(x)$. Это значит, что U^* пусто.

Пример 2. Пусть $f(x) = x^2$, $U=\{x: -1 \leq x \leq 1\}$. Минимальное значение $f(x)$ на U равно нулю, множество U^* состоит из единственной точки.

Пример 3. Пусть $f(x) = \ln x$, $U=\{x: 0 < x \leq 1\}$. Здесь U^* пусто, т.к. во всех точках из U функция принимает конечные значения, а для последовательности $x_k = \frac{1}{k}, k = 1, 2, \ldots$ имеем $\lim\limits_{k\to\infty} f(x_k) = -\infty$.

В примерах 1–2 функции ограничены снизу на рассматриваемых множествах, а в примере 3 функция не ограничена снизу.

Определение 3. Функция $f(x)$ называется ***ограниченной снизу*** на множестве U, если существует число M такое, что $f(x) \geq M$ для всех $x \in U$.

В тех случаях, когда $U^* = \varnothing$, естественным обобщением понятия наименьшего значения функции является понятие нижней грани функции.

Определение 4. Пусть функция $f(x)$ ограничена снизу на множестве U. Тогда число f_* называется ***точной нижней гранью*** функции $f(x)$ на множестве U, если 1) $f(x) \geq f_*$ при всех $x \in U$; 2) для любого сколь угодно малого числа $\varepsilon > 0$ найдется точка $x_\varepsilon \in U$ такая, что $f(x_\varepsilon) < f_* + \varepsilon$.

Если функция $f(x)$ неограниченна снизу на U, то в качестве нижней грани *f(x)* на *U* принимается $f_* = -\infty$. Точную нижнюю грань *f(x)* на *U* обозначают через $\inf\limits_{x\in U} f(x) = f_*$.

В примерах 1–2 нижняя грань $f_* = 0$, а в примере 3 $f_* = -\infty$.

Если $U^* \neq \varnothing$, то, очевидно, нижняя грань *f(x)* на *U* совпадает с наименьшим значением f^* этой функции на *U*, т.е. $\inf\limits_{x\in U} f(x) = \min\limits_{x\in U} f(x)$. В этом случае говорят, что функция *f(x)* на *U* достигает своей нижней грани. Отметим, что $\inf\limits_{x\in U} f(x) = f_*$ всегда существует, а $\min\limits_{x\in U} f(x)$, как видно из примеров 1–3, не всегда имеет смысл.

Пример 4. Пусть $f(x) = \dfrac{1}{x}$, $U = \{x:\ 0 \le x < +\infty\}$. Покажем, что функция *f(x)* на *U* не имеет точек минимума, а точная нижняя грань существует.

Предположим, $U^* \neq \varnothing$, т.е. существует хотя бы одна точка $x^* \in [1;\ +\infty)$ такая, что $f(x^*) \le f(x)$ для всех $x \in [1;\ +\infty)$. Выберем произвольное число $x_0 > x^*$. Очевидно, $x_0 \in U$, причем $f(x^*) = \dfrac{1}{x^*} > \dfrac{1}{x_0} = f(x_0)$, что противоречит предыдущему неравенству. Поэтому исходное предположение неверно и $U^* = \varnothing$.

Убедимся в том, что число *0* является точной нижней гранью данной функции *f(x)* на *U*. В самом деле, для всех $x \in [1;\ +\infty)$ имеем $f(x) = \dfrac{1}{x} > 0$, т.е. число *0* удовлетворяет первому из неравенств в определении 4. Далее пусть $\varepsilon > 0$. Возьмем произвольное $x_\varepsilon > \max\left(\dfrac{1}{\varepsilon}, 1\right)$. Тогда, очевидно, $x_\varepsilon \in U$ и $f(x_\varepsilon) = \dfrac{1}{x_\varepsilon} < \dfrac{1}{1/\varepsilon} = 0 + \varepsilon$, т.е. для числа *0* выполняется и второе неравенство из определения точной нижней грани. Поэтому $f_* = \inf\limits_{x\in U} f(x) = 0$.

Если множество точек минимума функции *f(x)* на *U* пусто, то задача минимизации *f(x)* теряет смысл. В этом случае можно ограничиться поиском точки $\tilde{x}^* \in U$, в которой значение *f(x)* с заданной погрешностью ε приближает точную нижнюю грань функции *f(x)* на множестве *U*, т.е. $f(\tilde{x}^*) - f_* < \varepsilon$.

Определение 5. Последовательность точек x_k из *U* называется ***минимизирующей последовательностью*** для функции *f(x)* на множестве *U*, если $\lim\limits_{k\to\infty} f(x_k) = \inf\limits_{x\in U} f(x) = f_*$.

Из определения и существования точной нижней грани следует, что минимизирующая последовательность всегда существует.

Теперь можем перейти к формулировке постановки задачи одномерной оптимизации как задачи минимизации функции *f(x)* на множестве *U*.

Условимся, что запись

$\min\{f(x) \mid :$	(1)

или ей эквивалентная $\min\limits_{x\in U} f(x)$ будет означать, что ставится задача определения величины $f_* = \inf\limits_{x\in U} f(x)$. Причем в задаче (1) неважно, будет ли множество *U** точек минимума *f(x)* на *U* непустым или оно пусто. В случае, когда множество *U** не пусто, требуется наряду с f_* найти точку $x^* \in U^*$.

Заметим, что получить точное решение поставленной задачи (1) удается лишь в редких случаях. Поэтому на практике при решении задачи (1) обычно строят минимизирующую последовательность $\{ x_k \}$ для функции *f(x)* на *U* и затем в качестве приближения для f_* берут величину $f(x_k)$ при достаточно большом *k*. В случае непустого множества *U** достаточно построить минимизирующую последовательность $\{ x_k \}$, которая сходится к множеству *U**. Один достаточно широкий класс функций, для которых $U^* \neq \varnothing$, определяет известная теорема Вейерштрасса, согласно которой функция, непрерывная на замкнутом ограниченном множестве, достигает на этом множестве своих максимального и минимального значений. Таким образом, задача (1) с непрерывной целевой функцией *f(x)* на *U* всегда имеет решение.

Если функция $f(x)$ на множестве U имеет, кроме глобального, локальные минимумы, отличные от него, то минимизация $f(x)$, как правило, сильно затрудняется. Многие методы поиска точки минимума $f(x)$ приспособлены только для функций, у которых каждый локальный минимум является одновременно и глобальным. Этим свойством обладают унимодальные функции.

1.3. Классический подход

Напомним две важные для данного рассмотрения теоремы из классического анализа.

<u>Теорема Вейерштрасса</u>. Если функция $f(x)$ непрерывна на множестве $U=[a, b]$, то существует $\min\limits_{x\in U} f(x)$.

<u>Теорема Ферма.</u> Пусть функция $f(x)$ дифференцируема в точке $x_0\in U$. Если x_0 доставляет локальный минимум $f(x_0)=\min\limits_{x\in U} f(x)$, то $f'(x_0)=0$.

Определение 6. Точка x_0 называется точкой ***локального минимума***, если для всех x из некоторой окрестности x_0 выполняется неравенство $f(x)\geq f(x_0)$.

Пусть $f(x)$ – кусочно-непрерывная и кусочно-гладкая на U функция. Это означает, что на U может существовать лишь конечное число точек, где $f(x)$ терпит разрыв I-го рода, либо непрерывна, но не имеет производной.

Тогда точками минимума могут быть такие точки, в которых:

- $f(x)$ терпит разрыв;
- $f(x)$ непрерывна, но $f'(x)$ не существует;
- $f'(x)=0$;
- либо $x=a$, либо $x=b$.

Рисунки ниже иллюстрируют эти 4 случая.

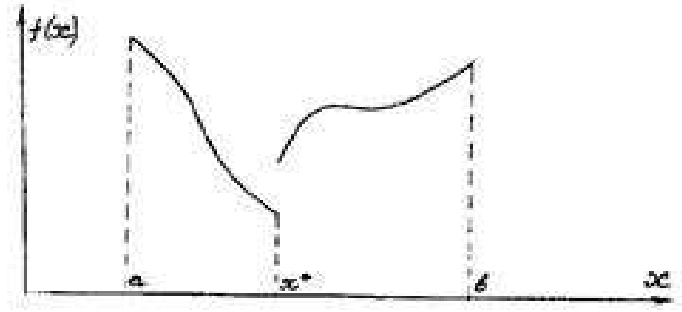

Рис. 2. Функция *f(x)* терпит разрыв в точке *x.**

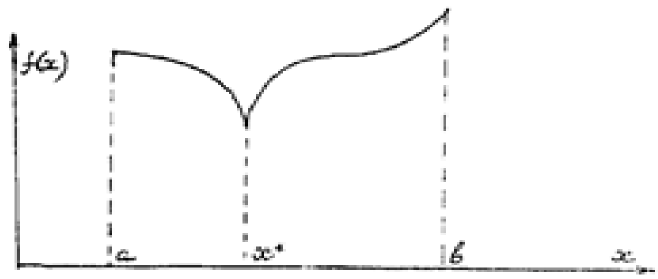

Рис. 3. Функция *f(x)* непрерывна, но производной в точке *x не существует.**

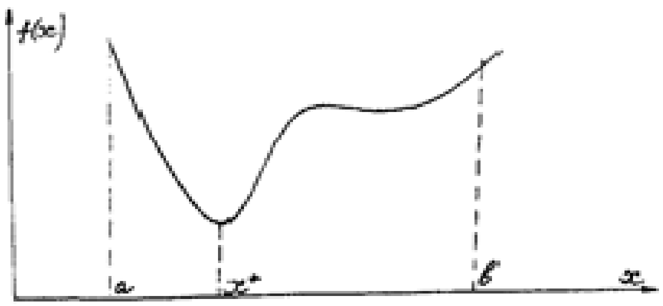

Рис. 4. $f'(x^*) = 0$

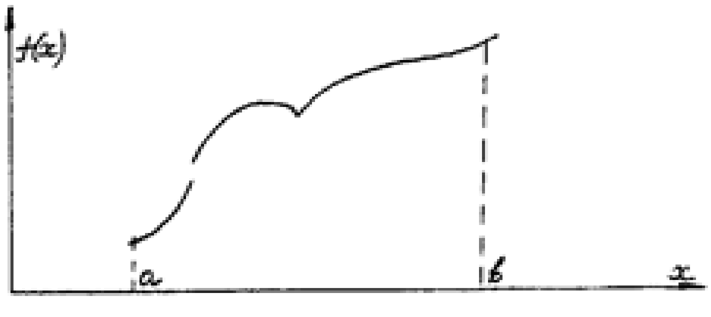

Рис. 5. *x*=a*

Определение 7. Функция $f(x)$ унимодальна на $\boxed{\quad}$, если существуют $\boxed{\qquad\qquad}$ такие, что

1) $f(x)$ строго монотонно убывает на $[a,\alpha]$,

2) $f(x)$ строго монотонно возрастает на $[\beta,b]$,

3) для $x \in [\alpha,\beta]$ $f(x) = \min\limits_{x\in[a,b]} f(x)$.

Если $\boxed{\quad}$, то $f(x)$ строго унимодальна.

Унимодальная функция не обязательно должна быть непрерывной и дифференцируемой. Ниже представлены примеры унимодальных функций.

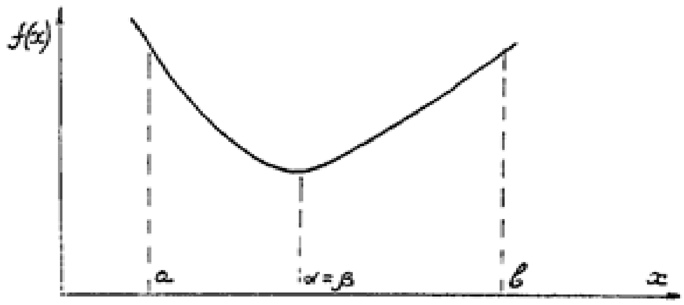

Рис. 6. Строго унимодальная, непрерывная, дифференцируемая функция

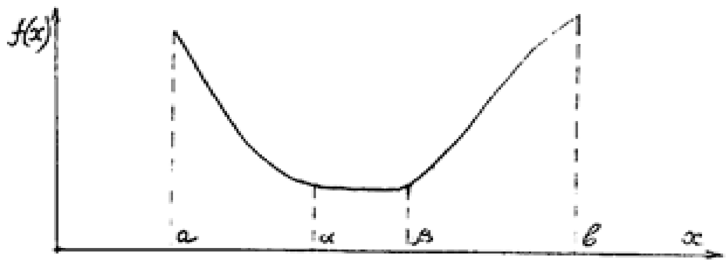

Рис. 7. Нестрого унимодальная, непрерывная, дифференцируемая функция

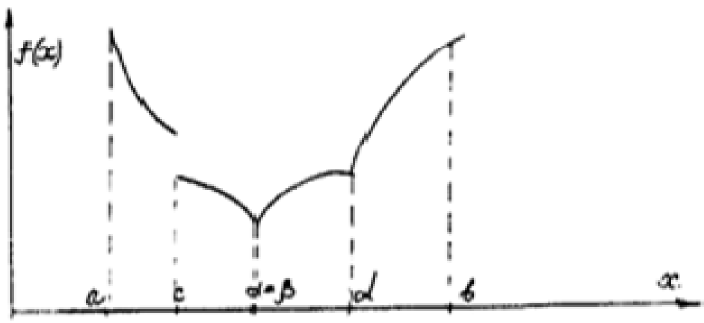

Рис. 8. При $x = c$ $f(x)$ имеет разрыв I - го рода,

при $x = \alpha$, $x = d$ производной у $\boxed{\quad}$ не существует.

2. ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ

Общие сведения о численных методах оптимизации, их классификация

Определение 1. Под ***численным методом*** одномерной минимизации понимается процедура получения числовой последовательности $\{x_k\}$ приближений к точному решению задачи минимизации.

Определение 2. Под ***численным методом*** одномерной минимизации понимается процедура получения вложенных отрезков, покрывающих точное решение:

$[a_0, b_0],\ [a_1, b_1],\ \ldots,\ [a_k, b_k]\ldots$

$[a_k, b_k] \subset [a_{k-1}, b_{k-1}]$.

Порядок метода

Метод имеет порядок k, если он использует информацию о производных $f(x)$ до k-го порядка включительно. Обычно применяются методы 0-го, 1-го и 2-го порядков.

Сходимость метода

Численный метод сходится, если последовательность $\{x_k\}$ сходится к точному решению, то есть $\lim\limits_{k\to\infty} x_k = x^*$ (скорость сходимости характеризуется $|x_k - x^*|$) или метод сходится, если $\lim\limits_{k\to\infty}(b_k - a_k) = 0$; $\lim\limits_{k\to\infty} a_k = \lim\limits_{k\to\infty} b_k = x^*$ (скорость сходимости характеризуется разностью $(b_k - a_k)$).

Критерии останова

Процесс вычислений желательно прервать, если

- достигнута требуемая точность вычислений;
- хорошее приближение не найдено, но скорость продвижения к оптимуму так упала, что нет смысла продолжать дальше;
- метод начал расходится или зациклился.

Часто на практике критерием прерывания по 2-й или 3-й причине является выполнение предельно допустимого числа получений приближенных решений.

Рекомендуется всегда этот критерий вводить в программу, даже если есть большая уверенность в благополучном завершении вычислений.

Если необходимо решить задачу оптимизации с точностью ε, то в качестве критерия окончания вычислений может служить неравенство

$$|x^{k+1} - x^k| \leq \varepsilon$$

Однако, для ряда задач, особенно негладких, этот критерий может привести к ложному решению.

Поэтому наряду с этим критерием обычно применяют один из двух следующих или даже сразу два:

$$|f(x^{k+1}) - f(x^k)| \leq \varepsilon_1$$

$$|f'(x^{k+1})| \leq \varepsilon_2$$

где $\varepsilon_1 . \varepsilon_2$ – малые константы.

2.1. Прямые методы

Методы, использующие только значения функции и не требующие вычисления ее производных, называются **прямыми методами** минимизации.

Большим достоинством прямых методов является то, что от целевой функции не требуется дифференцируемости и, более того, она может быть не задана в аналитическом виде. Единственное, на чем основаны алгоритмы прямых методов оптимизации, это возможность определения значений $f(\mathbf{x})$ в заданных точках.

Рассмотрим наиболее распространенные на практике прямые методы поиска точки минимума. Самым слабым требованием на функцию $f(\mathbf{x})$, позволяющим использовать эти методы, является ее унимодальность. Поэтому далее будем считать функцию $f(\mathbf{x})$ унимодальной на отрезке $[a; b]$.

2.1.1. Метод перебора

Метод перебора (пассивная стратегия поиска) является простейшим из прямых методов оптимизации.

Пусть

$$f(x) \in Q[a; b],$$

где $Q[a; b]$ – множество унимодальных функций на отрезке $[a; b]$. Требуется найти какую–либо из точек минимума x^* функции $f(\mathbf{x})$ на отрезке $[a; b]$ с абсолютной погрешностью $\varepsilon > 0$. Разобьем $[a; b]$ на n равных частей точками деления

$$x_i = a + i\,\frac{b-a}{n},\ \ i = 0, 1, 2,\ldots, n,$$

где $n \geq \dfrac{b-a}{\varepsilon}$.

Вычислив значения $f(\mathbf{x})$ в этих точках, путем сравнения найдем точку x_m, для которой

$$f(x_m) = \min f(x_i),\ \ 0 \leq i \leq n.$$

Далее полагаем $x^* \approx x_m$, $f^* \approx f(x_m)$. При этом максимальная погрешность ε_n определения точки x^* равна $\varepsilon_n = (b\text{-}a)/n$.

Блок-схема метода перебора приведена на рисунке 9.

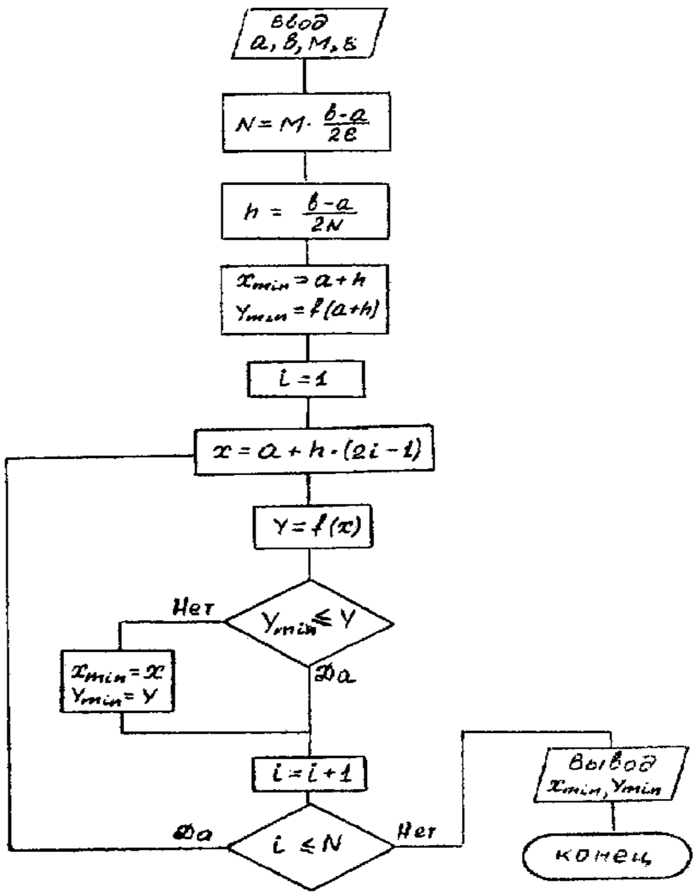

Рис. 9. Блок-схема метода перебора

Пример 1. Найти минимальное значение f^* и точку минимума x^* функции $f(x) = x^4 + 8x^3 - 6x^2 - 72x$ на отрезке [1,5; 2,0]. Точку x^* найти с погрешностью $\varepsilon = 0{,}05$.

$f(x) \in Q[1{,}5;\ 2{,}0]$, так как $f''(x) = 12x^2 + 48x - 12 > 0$ при $\mathbf{x} \in [1{,}5;\ 2]$.

Выбрав $n = \frac{2-1{,}5}{0{,}05} = 10$, вычислим значения $f(x_i)$, $x_i = 1{,}5 + i\,0{,}05$, $i = 0,1,\ldots,10$, поместив их в таблицу 1.

Таблица 1. Значения функции $f(x_i)$

x_i	1,50	1,55	1,60	1,65	1,70	1,75	1,80	1,85	1,90	1,95	2,00
$f(x_i)$	-89,4	-90,2	-91,2	-91,8	-92,08	-92,12	-91,9	-91,4	-90,5	-89,4	-88,0

Из таблицы 1 находим $x^* \approx 1{,}75$, $f^* \approx -92{,}12$.

2.1.2. Методы исключения интервалов

В методе перебора, рассмотренном выше, точки x_i, в которых определяются значения $f(x)$, выбираются заранее. Если же для выбора очередной точки вычисления $f(\boldsymbol{x})$ использовать информацию, содержащуюся в уже найденных значениях $f(x)$, то поиск точки минимума можно сделать более эффективным, т.е. сократить число определяемых для этого значений $f(x)$.

Фактически все одномерные методы поиска основаны на предположении, что исследуемая функция обладает свойством унимодальности, по крайней мере, в допустимой области. Это позволяет определить, в каком из задуманных двумя точками подынтервалов точка оптимума отсутствует.

Методы поиска, которые позволяют определить оптимум функции одной переменной путем последовательного исключения подынтервалов и, следовательно, путем уменьшения интервала поиска, носят название **методов исключения интервалов**.

Теорема 1. Пусть функция f унимодальна на замкнутом интервале $a \le x \le b$, а ее минимум достигается в точке x^*. Рассмотрим точки x_1 и x_2, расположенные в интервале таким образом, что $a < x_1 < x_2 < b$. Сравнивая значения функции в точках x_1 и x_2, можно сделать следующие выводы.

1. Если $f(x_1) > f(x_2)$, то точка минимума $f(\mathbf{x})$ не лежит в интервале (a, x_1), т.е. $x^* \in (x_1, b)$ (рис. 10,а)

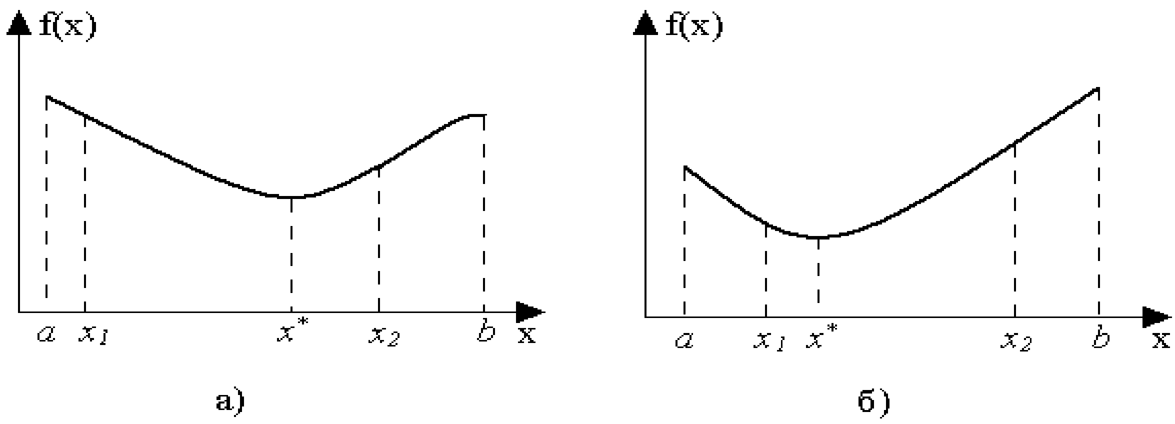

Рис. 10. Графические иллюстрации к теореме 1

2. Если $f(x_1) < f(x_2)$, то точка минимума не лежит в интервале (x_2, b), т.е. $x^* \in (a, x_2)$, (см. рис. 10,б).

Примечание. Если $f(x_1) = f(x_2)$, то можно исключить оба крайних интервала (a, x_1) и (x_2, b); при этом точка минимума должна располагаться в интервале (x_1, x_2).

Согласно теореме 1, которую иногда называют **правилом исключения интервалов**, можно реализовать процедуру поиска, позволяющую найти точку оптимума путем последовательного исключения частей исходного ограниченного интервала. Поиск завершается, когда оставшийся подынтервал уменьшается до достаточно малых размеров. Несомненным достоинством поисковых методов такого рода является то, что они основаны лишь на вычислении значений функции. При этом не требуется, чтобы исследуемые функции были дифференцируемы; более того, допустимы случаи, когда функцию нельзя даже записать в аналитическом виде.

В процессе применения рассматриваемых методов поиска можно выделить два этапа:

1. **этап установления границ интервала**, на котором реализуется процедура поиска границ достаточно широкого интервала, содержащего точку оптимума;

2. **этап уменьшения интервала**, на котором реализуется конечная последовательность преобразований исходного интервала с тем, чтобы уменьшить его длину до заранее установленной величины.

Этап установления границ интервала

На этом этапе сначала выбирается исходная точка, а затем на основе правила исключения строится относительно широкий интервал, содержащий точку оптимума. Обычно поиск граничных точек такого интервала проводится с помощью эвристических методов поиска. В этом случае (k+1)-я пробная точка определяется по рекуррентной формуле (метод Swann W.H.)

$$x_{k+1} = x_k + 2^k \Delta, \ \ k = 0, 1, 2, \ldots,$$

где x_0 – произвольно выбранная начальная точка; Δ − подбираемая некоторым способом величина шага. Знак Δ определяется путем сравнения значений $f(x_0)$, $f(x_0 + |\Delta|)$ и $f(x_0 - |\Delta|)$. Если

$$f(x_0 - |\Delta|) \geq f(x_0) \geq f(x_0 + |\Delta|),$$

то, согласно предположению об унимодальности, точка минимума должна располагаться правее точки x_0 и величина Δ выбирается положительной. Если изменить знаки неравенств на противоположные, то Δ следует выбирать отрицательной. Если же

$$f(x_0 - |\Delta|) \geq f(x_0) \leq f(x_0 + |\Delta|),$$

то точка минимума лежит между $x_0 - |\Delta|$ и $x_0 + |\Delta|$ и поиск граничных точек завершен. Случай, когда

$$f(x_0 - |\Delta|) \leq f(x_0) \geq f(x_0 + |\Delta|),$$

противоречит предположению об унимодальности.

Пример 2. Рассмотрим задачу минимизации функции $f(x) = (100-x)^2$ при заданной начальной точке $x_0 = 30$ и величине шага $|\Delta| = 5$.

Знак Δ определяется на основе сравнения значений

$$f(x_0) \qquad = f(30) = 4900,$$

$$f(x_0 + |\Delta|) = f(35) = 4225,$$

$$f(x_0 - |\Delta|) = f(25) = 5625.$$

Так как

$$f(x_0 - |\Delta|) \geq f(x_0) \geq f(x_0 + |\Delta|),$$

то величина Δ должна быть положительной, а координата точки минимума x^* должна быть больше 30. Имеем $x_1 = x_0 + \Delta = 35$. Далее

$$x_2 = x_1 + 2\Delta = 45,\ f(45) = 3025 < f(x_1),$$

откуда $x^* > 35$.

$$x_3 = x_2 + 2^2\Delta = 65,\ f(65) = 1225 < f(x_2),$$

откуда $x^* > 45$.

$$x_4 = x_3 + 2^3\Delta = 105,\ f(105) = 25 < f(x_3),$$

откуда $x^* > 65$.

$$x_5 = x_4 + 2^4\Delta = 185,\ f(185) = 7225 > f(x_4),$$

следовательно, $x^* < 185$.

Таким образом, шесть шагов вычислений x^* позволили выявить интервал $65 \leq x^* \leq 185$, в котором расположена точка x^*. Эффективность поиска граничных точек зависит от величины шага Δ. Если Δ велико – получаем грубые оценки координат граничных точек. Если Δ мало, то может потребоваться большой объем вычислений.

2.1.3. Метод деления интервала пополам (дихотомии)

После того как установлены границы интервала, содержащего точку оптимума, можно применить более сложную процедуру уменьшения интервала с целью получения уточненных оценок координат оптимума. Поскольку местонахождение точки оптимума априори неизвестно, целесообразно предположить, что размещение пробных точек должно обеспечивать уменьшение интервала в одном и том же отношении. Кроме того, в целях повышения эффективности алгоритма необходимо потребовать, чтобы указанное отношение было максимальным. Подобную стратегию иногда называют **минимаксной стратегией поиска**. Рассмотрим метод деления интервала пополам.

Метод деления интервала пополам является простейшим **последовательным методом минимизации** (методы минимизации, в которых точки x_i определяются в процессе минимума с помощью найденных ранее значений функции $f(x)$ называются последовательными методами). Он позволяет для любой функции $f(x) \in Q[a;b]$ построить последовательность вложенных отрезков $[a;b] \supset [a_1;b_1] \supset \ldots \supset [a_{n-1};b_{n-1}] \supset [a_n;b_n]$, каждый из которых содержит хотя бы одну из точек минимума x^* функции $f(\mathbf{x})$. Пусть $\varepsilon > 0$ – требуемая точность определения точки x^*. Выбрав $\delta \in [0;2\varepsilon]$ (δ может характеризовать погрешность измерений величины $\mathbf{x}$ и ограничена снизу возможностями измерительного прибора), построим последовательности $\{a_n\}$, $\{b_n\}$, $\{x_1^{(n)}\}$ и $\{x_2^{(n)}\}$, $n = 0,1,\ldots,$ используя рекуррентные формулы:

$$a_0 = a, \quad b_0 = b;$$

$$x_1^{(n-1)} = \frac{a_{n-1} + b_{n-1} - \delta}{2}; \quad x_2^{(n-1)} = \frac{a_{n-1} + b_{n-1} + \delta}{2}; \qquad (2)$$

$$a_n = a_{n-1},\ b_n = x_2^{(n-1)},\ \text{если } f(x_1^{(n-1)}) < f(x_2^{(n-1)}),$$

$$a_n = x_1^{(n-1)},\ b_n = b_{n-1},\ \text{если } f(x_1^{(n-1)}) > f(x_2^{(n-1)}).$$

Переход от отрезка $[a_{n-1};\ b_{n-1}]$ к отрезку $[a_n;\ b_n]$ методом деления отрезка пополам иллюстрируется на рис. 11а, если $f(x_1^{(n-1)}) < f(x_2^{(n-1)})$, и на рис. 11б, если $f(x_1^{(n-1)}) > f(x_2^{(n-1)})$. Полагая $x^* \approx (a_n + b_n)/2$, находим x^* с абсолютной погрешностью, не превосходящей величины

$$\varepsilon_n = (b_n - a_n)/2 = (b - a - \delta)/2^{n+1} + \delta/2. \qquad (3)$$

Используя условие $\varepsilon_n \le \varepsilon$ из последнего выражения можно найти необходимое число шагов n для обеспечения требуемой точности ε. Однако на практике часто поступают иначе: определив границы отрезка $[a_n; b_n]$, вычисляют ε_n по формуле (3) и сравнивают с заданной точностью ε.

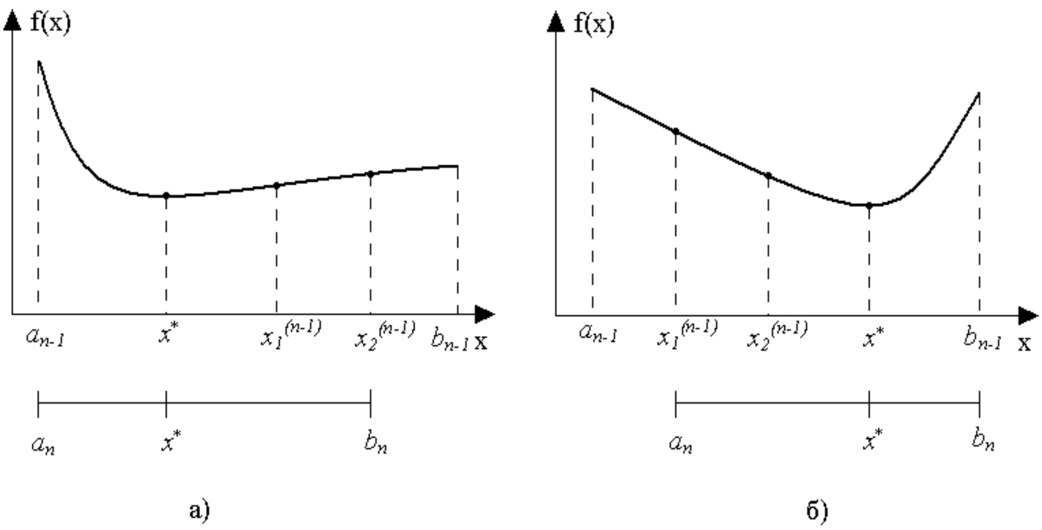

Рис. 11. Уменьшение интервала поиска точки минимума методом дихотомии

На рисунке 12 представлена блок-схема метода дихотомии.

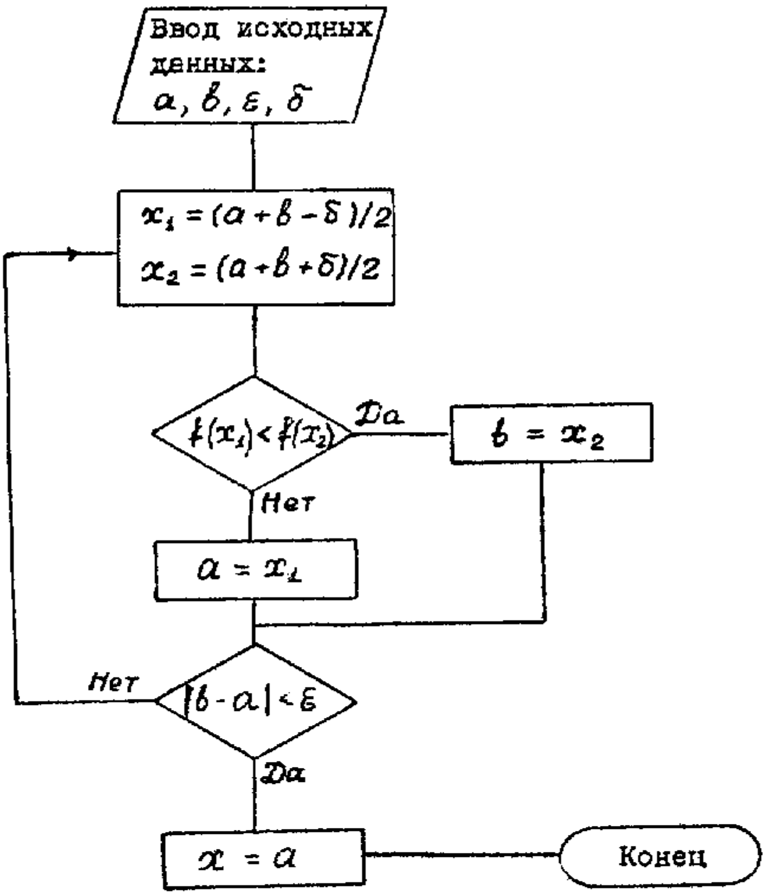

Рисунок 12. Блок-схема метода дихотомии

Пример 3. Найти минимальное значение f^* и точку минимума x^* функции $f(x) = x^4 + 8x^3 - 6x^2 - 72x$ на отрезке [1,5; 2,0]. Точку x^* найти с погрешностью $\varepsilon = 0{,}05$.

Положим $\delta = 0{,}02 < 2\varepsilon = 0{,}1$. Построим последовательность вложенных отрезков $[a_n; b_n]$ по формуле (2), записывая результаты вычислений в табл. 2.

Таблица 2 Значения вложенных отрезков и функций $f(\mathbf{x})$

n	a_n	b_n	$\varepsilon_n=(b_n-a_n)/2$	$x_1^{(n)}$	$x_2^{(n)}$	$f(x_1^{(n)})$	$f(x_2^{(n)})$	Примечание
0	1,5	2,0	0,25	1,74	1,76	-92,135	-92,096	$f(x_1^{(0)})<f(x_2^{(0)})$, $b_1=x_2^{(0)}$
1	1,5	1,76	0,13	1,62	1,64	-91,486	-91,696	$f(x_1^{(1)})>f(x_2^{(1)})$, $a_2=x_1^{(1)}$
2	1,62	1,76	0,07	1,68	1,70	-91,995	-92,084	$f(x_1^{(2)})>f(x_2^{(2)})$, $a_3=x_1^{(2)}$
3	1,68	1,76	0,04					$\varepsilon_n \le \varepsilon$, точность достигнута

Следовательно, $x^* \approx (a_3+b_3)/2 \approx 1{,}72$ и $f^* \approx f(1{,}72) = -92{,}13$.

Для увеличения скорости сходимости метода величину $\delta\in(0;\,2\varepsilon)$ целесообразно выбирать как можно меньшей, однако этот выбор ограничен снизу используемым количеством верных десятичных знаков при задании аргумента **x**. В любом случае δ должно быть больше машинного нуля применяемой ЭВМ.

Пример 4. Найти минимум $f(x)=x^2$ на отрезке $[-1,1]$ с $\varepsilon=0{,}05$

Решение.

Предварительно оценим, сколько шагов для этого потребуется. Выберем $\delta=0{,}02$.

$$k > \log_2((2-0{,}02)/0{,}03) = 6{,}05 .$$

Поскольку k - целое, то потребуется 7 шагов. Осуществим их.

1-й шаг.

$$x_1=-\frac{1-1}{2}-0{,}01=-0{,}01$$

$$x_2=-\frac{1-1}{2}+0{,}01=0{,}01$$

$$f(x_1)=f(x_2) \;\Rightarrow\; a^1=x_1=-0{,}01;\quad b^1=1 .$$

2-й шаг.

$x_1 = 0{,}485$; $x_2 = 0{,}505$

$f(x_1) < f(x_2) \Rightarrow a^2 = -0{,}01; \quad b^2 = 0{,}505$.

3-й шаг.

$x_1 = 0{,}2375$; $x_2 = 0{,}2575$

$f(x_1) < f(x_2) \Rightarrow a^3 = -0{,}01; \quad b^3 = 0{,}2575$.

4-й шаг.

$x_1 = 0{,}10375$; $x_2 = 0{,}12375$

$f(x_1) < f(x_2) \Rightarrow a^4 = -0{,}01; \quad b^4 = 0{,}12375$.

5-й шаг.

$x_1 = 0{,}051375$; $x_2 = 0{,}071375$

$f(x_1) < f(x_2) \Rightarrow a^5 = -0{,}01; \quad b^5 = 0{,}071375$.

6-й шаг.

$x_1 = 0{,}0206875$; $x_2 = 0{,}0406875$.

$f(x_1) < f(x_2) \Rightarrow a^6 = -0{,}01; \quad b^6 = 0{,}0406875$

7-й шаг.

$x_1 = 0{,}005344$; $x_2 = 0{,}025344$

$f(x_1) < f(x_2) \Rightarrow a^7 = -0{,}01; \quad b^7 = 0{,}025344$.

$b^7 - a^7 = 0{,}035344 < 0{,}05$

$b^6 - a^6 = 0{,}0506875 > 0{,}05$

Следовательно, действительно только 7 шагов приводят к решению с заданной точностью. В качестве x^* принимаем -0,01.

На следующем рисунке проиллюстрировано уменьшение отрезка неопределенности по шагам.

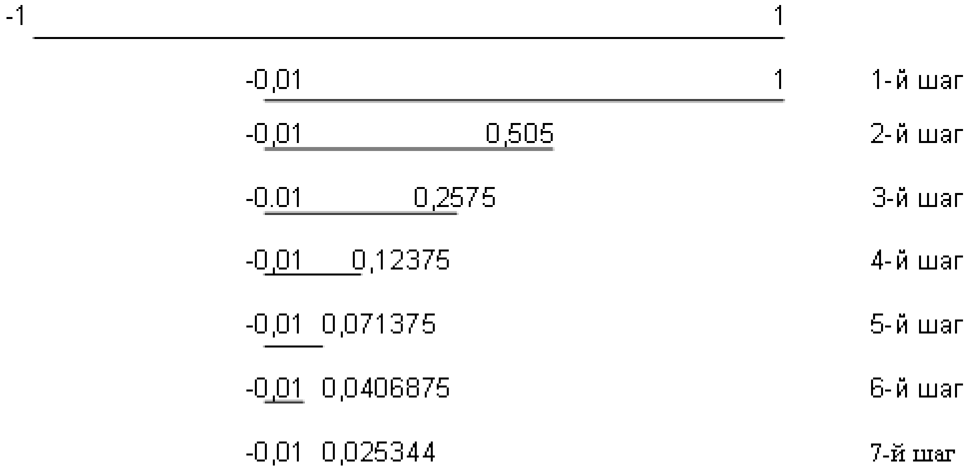

Рис. 13. Пошаговое уменьшение отрезка неопределенности.

2.1.4. Метод золотого сечения

Метод золотого сечения также является последовательным методом минимизации. Этот метод использует найденные значения $f(x)$ более рационально, чем метод деления интервала пополам, что позволяет переходить к очередному интервалу, содержащему x^* после вычисления одного, а не двух значений $f(\mathbf{x})$.

Рассмотрим на исходном отрезке $[a; b]$ точку x_1 и вычислим $f(x_1)$. Зная значение целевой функции в одной точке, невозможно сузить область поиска точки x^*. Поэтому выберем вторую точку x_2 так, чтобы a < $x_1 < x_2 <$ b, и вычислим $f(x_2)$. Возможен один из следующих двух случаев: $f(x_1) \le f(x_2)$ или $f(x_1) \ge f(x_2)$.

Согласно свойству унимодальной функции: если $f(x_1) \le f(x_2)$, то $x^* < x_2$; если же $f(x_1) \ge f(x_2)$, то $x^* > x_1$; в первом случае искомая точка x^* не может быть на отрезке $[x_2; b]$, а во втором − на отрезке $[a; x_1]$ (эти отрезки на рис. 14 отмечены штриховкой). Следовательно, теперь область поиска сужается и следующую точку x_3 следует брать в одном из укороченных отрезков $[a; x_2]$ или $[x_1; b]$.

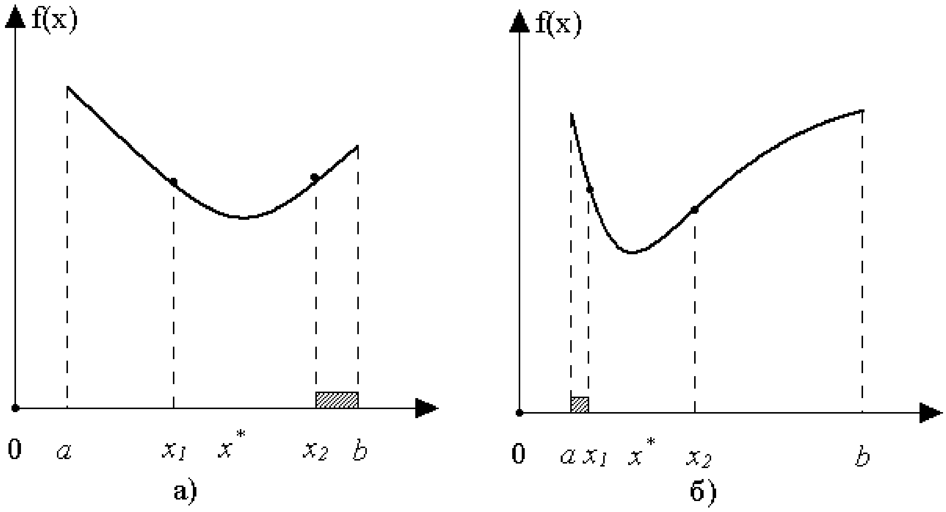

Рис. 14. Уменьшение интервала поиска точки минимума методом золотого сечения

Установим, где на исходном отрезке лучше всего выбрать точки x_1 и x_2. Так как первоначально ничего неизвестно о положении точки x^*, то оба указанных выше случая равновозможны, т.е. “лишним” может оказаться любой из отрезков [x_2; b] и [a; x_1]. Отсюда ясно, что точки x_1 и x_2 должны быть расположены симметрично относительно середины отрезка [a; b]. Чтобы максимально сузить область поиска, эти точки должны быть “поближе” к середине исходного отрезка. Если их взять рядом с серединой исходного отрезка, то на втором этапе сужение области поиска будет незначительным (рис. 15). На втором этапе сужения области поиска потребуется вычислить лишь одно значение $f(x_3)$, которое будем сравнивать с уже имеющимся значением $f(x_1)$ или $f(x_2)$ в зависимости от того, какой из двух случаев реализовался.

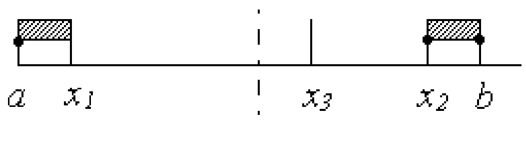

Рис. 15. К выбору пробных точек x_1 и x_2

Поэтому, с одной стороны, точки x_1 и x_2 следует выбирать рядом с серединой отрезка, а с другой – слишком близкими их брать нельзя. Для того чтобы найти “золотую середину” используется метод “золотого сечения”.

Поиск с помощью метода золотого сечения основан на разбиении отрезка прямой на две части, известном как “золотое сечение” – отношение длины всего отрезка к большей части равно отношению большей части к меньшей.

Рассмотрим симметричное расположение двух пробных точек на исходном интервале единичной длины (рис. 16). Пробные точки x_1, x_2 отстоят от граничных точек интервала на расстоянии *F*. При таком симметричном расположении длина остающегося после исключения интервала всегда равна *F* независимо от того, какое из значений функции в пробных точках оказывается меньшим.

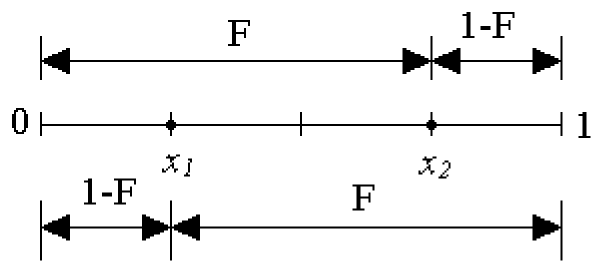

Рис. 16. Поиск пробных точек с помощью метода золотого сечения

Предположим, что исключается правый подынтервал. На рис. 17 показано, что оставшийся подынтервал длины *F* содержит одну пробную точку x_1, расположенную на расстоянии $(1 - F)$ от левой граничной точки. Чтобы точки $x_1 = 1 - F$ и $x_2 = F$ де лили отрезки $[0; F]$ и $[0; 1]$ в одном и том же отношении должно выполняться равенство $\frac{1}{F}=\frac{F}{1-F}$ или $F^2=1-F$, находим положительное значение $F=\frac{\sqrt{5}-1}{2}\approx 0{,}62$.

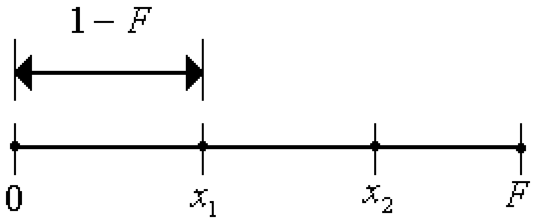

Рис. 17. Интервалы, полученные методом золотого сечения

Таким образом, $x_1 = 1-F = \frac{3-\sqrt{5}}{2}\approx 0{,}38$, $x_2 = F = \frac{\sqrt{5}-1}{2}$.

Для того чтобы симметрия поискового образа сохранялась, расстояние (1 - F) должно составлять F- ю часть длины интервала (которая равна F). При таком выборе F следующая пробная точка x_3 размещается на расстоянии, равном F- й части длины интервала, от правой граничной точки интервала (рис. 18). Отсюда следует, что при выборе F в соответствии с условием $1-F=F^2$ симметрия поискового образца (рис. 16) сохраняется при переходе к уменьшенному интервалу (рис. 18). Схема поиска, при которой пробные точки делят интервал в этом отношении, известна под названием **поиска с помощью метода золотого сечения**.

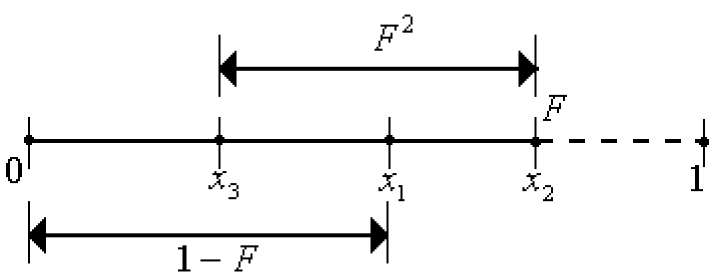

Рис. 18. Симметрия золотого сечения интервала

Для произвольного отрезка $[a;\ b]$ выражения для пробных точек примут вид

$$x_1 = a + F_1(b - a), \quad x_2 = a + F_2(b - a) \tag{4}$$

Зная одну из точек золотого сечения отрезка $[a;\ b]$, другую можно найти по одной из формул

$$x_1 = a + b - x_2, \quad x_2 = a + b - x_1 \tag{5}$$

Пусть $f(\mathbf{x}) \in Q[a;\ b]$ и требуется найти точку минимума x^* функции $f(\mathbf{x})$ на $[a;\ b]$. Построим последовательности $\{a_n\}$, $\{b_n\}$ и $\{\bar{x}_n\}$, $n = 1,2,\ldots$, следующим образом:

$$a_n = a_{n-1},\ b_n = x_2^{(n-1)},\ \bar{x}_n = x_1^{(n-1)}, \text{ если } f(x_1^{(n-1)}) \le f(x_2^{(n-1)}); \tag{6}$$

$$a_n = x_1^{(n-1)},\ b_n = b_{n-1},\ \bar{x}_n = x_2^{(n-1)}, \text{ если } f(x_1^{(n-1)}) > f(x_2^{(n-1)}),\ n = 2,3,\ldots,$$

где $a_1 = a$, $b_1 = b$, $x_1^{(n-1)}$ и $x_2^{(n-1)}$ − первая и вторая точки золотого сечения (4) отрезка $[a_{n-1};\ b_{n-1}]$.

Для определения чисел a_n, b_n, $\bar{x}_n$ по найденным a_{n-1}, b_{n-1}, $\bar{x}_{n-1}$ необходимо выполнить следующие операции:

1. найти одну из точек золотого сечения отрезка $[a_{n-1};\ b_{n-1}]$ по известной другой точке $\bar{x}_{n-1}$, используя формулы (5). При определении x^* с большой точностью, чтобы избежать накопления ошибок округления, обычно точки золотого сечения отрезка $[a_n;\ b_n]$ находят по формулам (4) и в качестве $x_1^{(n-1)}$ и $x_2^{(n-1)}$ используют $\bar{x}_{n-1}$ и ту из найденных точек, которая больше отличается от $\bar{x}_{n-1}$;

2. вычислить значение $f(\mathbf{x})$ во вновь найденной точке золотого сечения (значение в другой точке $\bar{x}_{n-1}$ уже вычислено на одном из предыдущих шагов);

3. сравнить значения $f(x_1^{(n-1)})$ и $f(x_2^{(n-1)})$ и найти a_n, b_n и $\bar{x}_n$ по формулам (6).

Таким образом, на каждом шаге определения a_n, b_n и $\bar{x}_n$ $n = 2,3,\ldots$, требуется вычисление одного значения $f(\mathbf{x})$. Положив $x^* \approx \bar{x}_n$, найдем точку минимума x^* с точностью ε_n:

$$|x^* - \bar{x}_n| \le \varepsilon_n = [(\sqrt{5} - 1)/2]^n \cdot (b - a),$$

отсюда следует, что число шагов n метода золотого сечения, обеспечивающее заданную точность ε нахождения точки x^*, должно удовлетворять неравенству

$$n \ge \ln\left(\frac{\varepsilon_n}{b-a}\right) / \ln\left(\frac{\sqrt{5}-1}{2}\right) \approx -2{,}1 \ln\left(\frac{\varepsilon_n}{b-a}\right) \qquad (7)$$

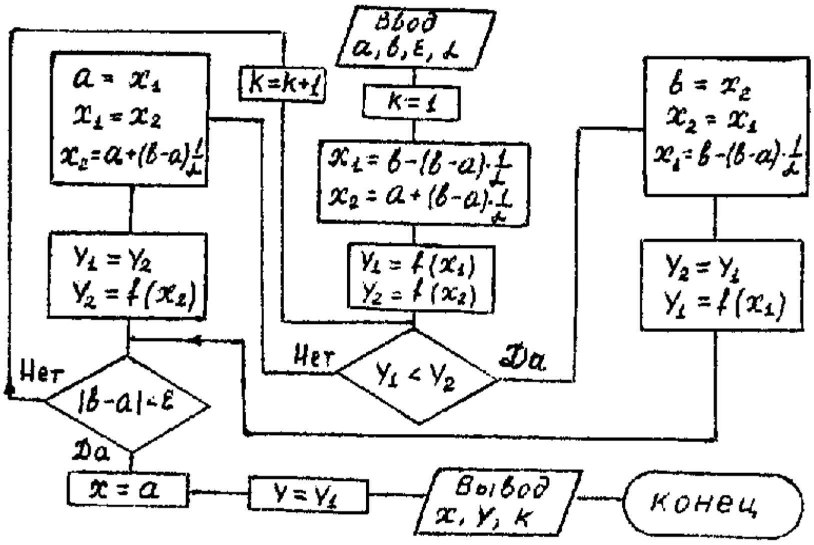

Рисунок 19. Блок-схема метода золотого сечения.

Пример 5. Решить пример 3 методом золотого сечения.

Вычисления проведем по формулам (6), представив результаты в таблице 3, где стрелками отмечены сохраняющиеся при переходе к следующему шагу значения.

Таблица 3. Значения пробных точек и функции f(x)

N	ε_n	a_n	b_n	$x_1^{(n)}$	$x_2^{(n)}$	$f(x_1^{(n)})$	$f(x_2^{(n)})$	Примечание
1	0,309	1,5	2,0	1,691	1,809	-92,049	-91,814	$f(x_1^{(1)}) < f(x_2^{(1)})$, $b_2 = x_2^{(1)}$
2	0,191	1,5	1,809	1,618	1,691	-91,464	-92,049	$f(x_1^{(2)}) > f(x_2^{(2)})$, $a_3 = x_1^{(2)}$
3	0,118	1,618	1,809	1,691	1,736	-92,049	-92,138	$f(x_1^{(3)})>f(x_2^{(3)})$, $a_4 = x_1^{(3)}$
4	0,073	1,691	1,809	1,736	1,764	-92,138	-92,083	$f(x_1^{(4)})< f(x_2^{(4)})$, $b_5 = x_2^{(4)}$
5	0,045				1,736		-92,138	$\varepsilon_n < \varepsilon$, точность достигнута

Из таблицы 3 получаем $x^* \approx \bar{x}_5 = 1{,}736$, $f^* \approx f(\bar{x}_5) = -92{,}138$. Если воспользоваться формулой (7), то n можно определить заранее. В нашем случае $n \ge 4{,}79$, т.е. $n = 5$.

Пример 6. Найти минимум $f(x) = x^2$ на отрезке $[-1,1]$ с $\varepsilon = 0{,}05$

Предварительно определим, сколько потребуется шагов метода золотого сечения.

$$\frac{2}{(1{,}618)^N} < 0{,}05$$

$$N > \frac{\ln(\frac{2}{0{,}05})}{\ln 1{,}618} = 7{,}7$$

Итак, потребуется 8 шагов метода золотого сечения, при этом значения $f(x)$ придется вычислять 9 раз.

<u>1-й шаг.</u>

$$x_1 = 1 - 2 * \frac{1}{1{,}618} = -0{,}2361; \qquad x_2 = -1 + \frac{2}{1{,}618} = 0{,}2361$$

$$f(x_1) = f(x_2) \quad \Rightarrow \quad a^1 = \text{-}0{,}2361; \quad b^1 = 1\,.$$

<u>2-й шаг.</u>

$$x_1 = 0{,}2361; \qquad x_2 = -0{,}2361 + \frac{1{,}2361}{1{,}618} = 0{,}5279$$

$$f(x_1) < f(x_2) \quad \Rightarrow \quad a^2 = \text{-}0{,}2361; \quad b^2 = 0{,}5279\,.$$

<u>3-й шаг.</u>

$$x_1 = 0{,}5279 - \frac{(0{,}5279 + 0{,}2361)}{1{,}618} = 0{,}05573; \qquad x_2 = 0{,}2361$$

$$f(x_1) < f(x_2) \quad \Rightarrow \quad a^3 = \text{-}0{,}2361; \quad b^3 = 0{,}2361\,.$$

<u>4-й шаг.</u>

$$x_1 = 0{,}2361 - \frac{(0{,}2361 + 0{,}2361)}{1{,}618} = -0{,}05573; \qquad x_2 = 0{,}05523$$

$$f(x_1) = f(x_2) \quad \Rightarrow \quad a^4 = \text{-}0{,}05573; \quad b^4 = 0{,}2361\,.$$

<u>5-й шаг.</u>

$$x_1 = 0{,}2361 - \frac{(0{,}2361 + 0{,}05573)}{1{,}618} = 0{,}05573; \qquad x_2 = 0{,}12463$$

$$f(x_1) = f(x_2) \quad \Rightarrow \quad a^5 = \text{-}0{,}05573; \quad b^5 = 0{,}12463\,.$$

<u>6-й шаг.</u>

$$x_1 = 0{,}12463 - \frac{(0{,}12463 + 0{,}05573)}{1{,}618} = -0{,}01316; \qquad x_2 = 0{,}05573$$

$$f(x_1) < f(x_2) \quad \Rightarrow \quad a^6 = \text{-}0{,}05573; \quad b^6 = 0{,}05573\,.$$

<u>7-й шаг.</u>

$$x_1 = 0{,}05573 - \frac{(0{,}05573 + 0{,}05573)}{1{,}618} = -0{,}01316; \qquad x_2 = 0{,}01316$$

$$f(x_1) < f(x_2) \quad \Rightarrow \quad a^7 = \text{-}0{,}05573; \quad b^7 = 0{,}01316\,.$$

<u>8-й шаг.</u>

$$x_1 = 0{,}01316 - \frac{(0{,}01316 + 0{,}05573)}{1{,}618} = -0{,}02942; \qquad x_2 = 0{,}01316$$

$$f(x_1) > f(x_2) \quad \Rightarrow \quad a^8 = \text{-}0{,}02942; \quad b^8 = 0{,}01316\,.$$

$$b^8 - a^8 = 0{,}04258 < \varepsilon \qquad \Rightarrow \quad x^* = \text{-}0{,}02942$$

2.1.5. Метод Фибоначчи

Часто в вычислительных процедурах существенные трудности возникают в связи с вычислениями значений $f(x)$. Например, $f(x)$ вычисляется в процессе эксперимента, либо $f(x)$ задана сложной формулой.

К методам, в которых при ограничениях на количество вычислений значений $f(x)$ достигается в определенном смысле наилучшая точность, относятся методы Фибоначчи и золотого сечения.

Как и в методе дихотомии, процедура будет заключаться в последовательном уменьшении отрезка неопределенности на основании анализа значений функции в двух внутренних точках отрезка с существенным отличием от предыдущего, состоящего в том, что одна из внутренних точек последующего отрезка неопределенности совпадет с одной из двух внутренних точек предыдущего отрезка неопределенности.

Определение. Последовательность чисел $F_0 = F_1 = 1;$ называется последовательностью Фибоначчи.

Зададимся некоторым N и выпишем последовательность чисел Фибоначчи.

Итак, необходимо найти минимум $f(x)$ на отрезке $[a,b]$ с точностью ε. Опишем 1-й шаг метода Фибоначчи.

Как и в предыдущем методе найдем $x_1 < x_2$ на отрезке $[a,b]$:

$$x_1 = a + \frac{F_{N-2}}{F_N} * (b-a); \qquad x_2 = b - \frac{F_{N-2}}{F_N} * (b-a)$$

Из формул видно, что x_1, x_2 симметричны относительно середины отрезка $[a,b]$. Дальнейшая процедура уменьшения отрезка неопределенности совпадает с методом дихотомии. Итак, основное отличие метода Фибоначчи от метода дихотомии состоит в выборе точек x_1, x_2 на каждом шаге.

В силу свойств последовательности Фибоначчи, на каждом шаге, кроме 1-го и предпоследнего, вычисляется одно новое значение функции, другое значение используется из предыдущего шага. Только на 1-м шаге значение $f(x)$ вычисляется дважды, а на предпоследнем, когда x_1 совпадает с x_2, $f(x_1)$ известно из предыдущего шага. Можно показать, что на $(N-1)$ -м шаге x_1 и x_2 совпадут, этим завершится процедура деления отрезка неопределенности. Для получения окончательного результата необходимо вычислить $f(x_1)$ и $f(x_1+\delta)$, где δ - малая величина, параметр метода.

Если $f(x_1)<f(x_1+\delta)$, то полагают, что $x^* \in [a^{N-1}, x_1+\delta]$, в противном случае $x^* \in [x_1, b^{N-1}]$.

Посмотрим, как уменьшается отрезок неопределенности:

$$b^1-a^1=\frac{F_{N-1}}{F_N}(b-a);$$

$$b^2-a^2=\frac{F_{N-1}}{F_{N-1}}(b^1-a^1)=\frac{F_{N-2}}{F_{N-1}}*\frac{F_{N-1}}{F_N}(b-a)$$

....

Таким образом, $(N-1)$-й шаг метода Фибоначчи обеспечивает уменьшение длины отрезка неопределенности в F_N раз.

Необходимо заметить, что для решения задачи минимизации с заданной точностью ε необходимо решить неравенство $\frac{b-a}{F_N}+\delta<\varepsilon$ относительно F_N, получить последовательность чисел Фибоначчи $F_0, F_1, \ldots, F_{N-1}, F_N$ и использовать ее с конца.

Замечание 1.

Нетрудно заметить, что теоретически достаточно найти первую точку метода x_1, остальные точки можно получать, используя свойство их симметрии относительно центра отрезка, однако в этом случае быстро накапливается погрешность. Чтобы избежать накопления погрешности, следует пересчитывать точки x_1 и x_2 по соответствующим формулам.

Замечание 2.

Поскольку N определяется сначала как функция от ε, алгоритм не позволяет получить более точный результат путем продолжения счета. Для обеспечения другой точности необходимо реализовать новую вычислительную процедуру.

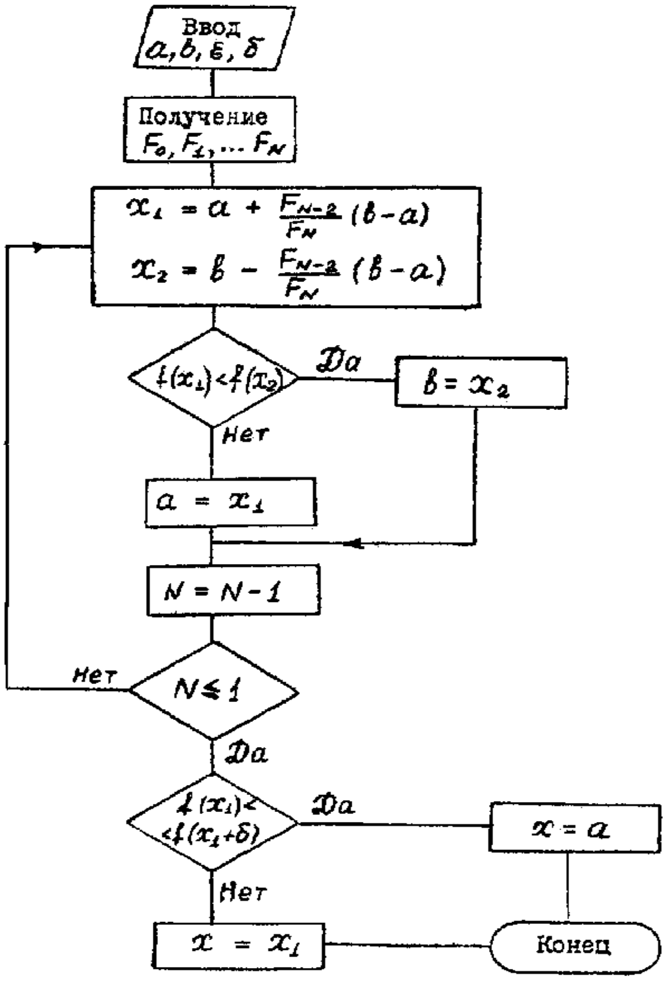

Рисунок 19. Блок-схема метода Фибоначчи.

Пример 7

Найти минимум $f(x)=x^2$ на отрезке $[-1,1]$ с $\varepsilon=0,05$

Начнем с определения F_N:

$$(b-a)/F_N < 0,05 \qquad \frac{2}{F_N} < 0,05 \quad \Rightarrow F_N > 40$$

Таблица 4. Значения чисел Фибоначчи

F_0	1
F_1	1
F_2	2
F_3	3
F_4	5
F_5	8
F_6	13
F_7	21
F_8	34
F_9	55

Для решения поставленной задачи потребуется 9 шагов по методу Фибоначчи, при этом понадобится 9 раз вычислять $f(x)$. Заметим, что для решения этой же задачи методом дихотомии мы проделали 7 шагов, то есть $f(x)$ вычисляли 14 раз.

<u>1-й шаг.</u>

$$x_1 = -1 + 2 * \frac{21}{55} = -0,2364; \qquad x_2 = 0,2364$$

$$f(x_1) = f(x_2) \Rightarrow a^1 = -0,2364; \quad b^1 = 1.$$

<u>2-й шаг.</u>

$$x_1 = -0,2364 + 1,2364 * \frac{13}{34} = 0,2364$$

$$x_2 = 1 - 1,2364 * \frac{13}{34} = 0,5273$$

$$f(x_1) < f(x_2) \Rightarrow a^2 = -0,2364; \quad b^2 = 0,5273.$$

<u>3-й шаг.</u>

$$x_1 = 0,05453; \qquad x_2 = 0,2364$$

$f(x_1) < f(x_2) \Rightarrow a^3 = -02364; \quad b^3 = 0,2364 .$

4-й шаг.

$x_1 = -0,05453$; $\quad x_2 = 0,05454$

$f(x_1) < f(x_2) \Rightarrow a^4 = -0,2364; \quad b^4 = 0,05453 .$

5-й шаг.

$x_1 = -0,1273$; $\quad x_2 = -0,05457$

$f(x_1) < f(x_2) \Rightarrow a^5 = -0,1273; \quad b^5 = 0,05453 .$

6-й шаг.

$x_1 = -0,05457$; $\quad x_2 = -0,018202$

$f(x_1) > f(x_2) \Rightarrow a^6 = -0,05457; \quad b^6 = 0,05453 .$

7-й шаг.

$x_1 = -0,05457 + 0,036367 = -0,018203$

$x_2 = -0,018163$

$f(x_1) > f(x_2) \Rightarrow a^7 = -0,018203; \quad b^7 = 0,05453 .$

8-й шаг.

$x_1 = -0,018164$; $\quad x_2 = -0,018164$

$$x_1 = x_2 .$$

9-й шаг.

$x_1 = -0,018163$; $\quad x_1 + \delta = x_2 = 0,028163$

$f(x_1) < f(x_2) \Rightarrow a^9 = -0,018203; \quad b^9 = 0,028163 .$

$x^* = -0,018203$

Замечание.

Вычисления проводились с 5 знаками после запятой, поэтому точки последующего и предыдущего шага совпадают не полностью.

2.1.6. Метод квадратичной интерполяции (парабол)

Этот метод применяется, когда сделано несколько шагов более грубыми методами, то есть отрезок локализации точки минимума достаточно мал.

Итак, найти $\min\limits_{x\in[a,b]} f(x)$ с точностью ε.

Выбираем $x_1 : a < x_1 < b$.

Через точки $[a, f(a)],\ [x_1, f(x_1)],\ [b, f(b)]$ проводим параболу, то есть строим интерполяционный многочлен 2-й степени:

$$L_2(x) = f(a)*\frac{(x-x_1)(x-b)}{(a-x_1)(a-b)} + f(x_1)*\frac{(x-a)(x-b)}{(x_1-a)(x_1-b)} + f(b)*\frac{(x-a)(x-x_1)}{(b-a)(b-x_1)}$$

Находим вершину параболы x_2 из условия $L_2'(x)=0$,

$$x_2 = x_1 + \frac{1}{2}*\frac{(b-x_1)^2*(f(a)-f(x_1))-(x-a)^2*(f(b)-f(x_1))}{(b-x_1)*(f(a)-f(x_1))+(x_1-a)*(f(b)-f(x_1))}.$$

Если $x_2 > x_1$, то необходимо выяснить $f(x_1) \geq f(x_2)$ или нет.

Если $f(x_1) \geq f(x_2)$, то по трем точкам $(x_1, f(x_1)),\ (x_2, f(x_2)),\ (b, f(b))$ строится следующая парабола.

Если $f(x_1) < f(x_2)$, то следующая парабола строится по точкам:

$(a, f(a)),\ (x_1, f(x_1)),\ (x_2, f(x_2))$

В случае если $x_1 > x_2$, соответствующим образом подбирается очередная тройка чисел, через которые проводится парабола. Каждая новая вершина параболы принимается за очередное приближенное решение.

На рисунке представлена блок-схема метода парабол, в которой итерационный процесс заканчивается, как только абсциссы вершины последующей и предыдущей параболы разнятся менее чем на ε.

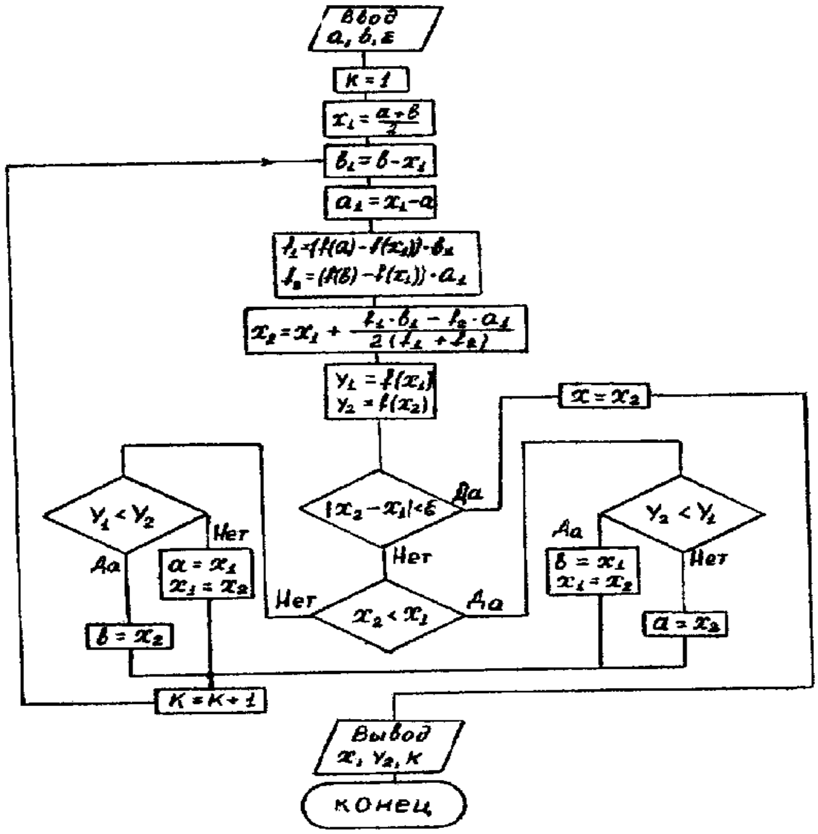

Рисунок 20. Блок-схема метода квадратичной интерполяции.

Пример 8

Найти минимум $f(x) = x^2$ на отрезке $[-1,1]$ с точностью $\varepsilon = 0,05$

Пусть $x_1 = 0,5$

<u>1-й шаг</u>:

по трем точкам (-1, 1), (0.5, 0.25), (1, 1) строим параболу и находим ее вершину

$$\times_2 = 0,5 + \frac{1}{2} * \frac{(1-0,5)^2 * (1-0,25) - (1,5)^2 * (1-0,25)}{0,5 * (1-0,25) + 1,5 * (1-0,25)} = 0\ ,$$

$$|0,5 - 0| = 0,5 > 0,01$$

Требование по точности не выполнено, поэтому необходимо продолжить вычисления

$$f(x_2) = 0\ ,$$

$f(x_2) < f(x_1)$, следовательно, следующую параболу строим по трем точкам:

$(0;0)$, $(0,5;0,25)$, $(1,1)$ и находим

$$x_2 = 0,5 + \frac{1}{2} * \frac{(1-0,5)^2 * (0-0,25) - (0,5-0)^2 * (1-0,25)}{(1-0,5) * (0-0,25) + (0,5-0) * (1-0,25)} = 0,$$

$$|0-0| = 0 < 0,01$$

Вычисления завершены, в качестве приближенного решения получили $x = 0$.

Нетрудно убедится в том, что $x = 0$ является точным решением задачи. Это совпадение объясняется тем, что $f(x)$ есть многочлен 2-й степени, построенные интерполяционные многочлены, естественно, полностью совпали с $f(x)$.

Замечание.

В описанном алгоритме предполагалось, что $f(x)$ унимодальна.

В случае, если этот метод применять для неунимодальной $f(x)$, алгоритм может зациклиться.

Чтобы избежать возможного зацикливания, целесообразно перед построением очередной параболы проводить проверку выпуклости трех выбранных точек. Значение функции в средней точке должно быть меньше значений на концах рассматриваемого отрезка.

2.2. Методы, использующие производные функции

Прямые методы используются при минимальных требованиях к целевой функции $f(\mathbf{x})$ – она считается унимодальной, и вычислению подлежат значения только самой функции, но не ее производных.

Если усилить эти требования, предположив, что $f(\mathbf{x})$ является дифференцируемой или дважды дифференцируемой выпуклой функцией, и считать, что возможно вычисление производных $f(\mathbf{x})$ в произвольно выбранных точках, то эффективность процедур поиска точки минимума можно существенно повысить.

Рассмотрим методы минимизации, в которых используются значения производных целевой функции. Напомним, что для выпуклой дифференцируемой функции равенство $f'(\mathbf{x}) = 0$ является не только необходимым, но и достаточным условием глобального минимума. Поэтому, если известно, что x^* является внутренней точкой отрезка $[a; b]$, то приближенное равенство $f'(\mathbf{x}) \approx 0$ или $|f'(\mathbf{x})| \le \varepsilon$ может служить условием остановки вычислений в рассматриваемых ниже методах.

2.2.1. Условие Липшица

Для определения глобального экстремума можно использовать равномерную сетку с достаточно мелким шагом (метод перебора). Однако такой подход требует слишком большого времени счета. Чтобы избежать этого, можно использовать некоторые априорные сведения о характере критерия оптимальности, позволяющие изменять шаг в процессе поиска в зависимости от ранее вычисленных значений. С этой целью можно использовать условие Липшица.

Определение. Функция $f(\mathbf{x})$ удовлетворяет условию Липшица в области G, если существует такая постоянная величина $\boldsymbol{L > 0}$ (**константа Липшица**), что для любых двух векторов $\mathbf{x}_1, \mathbf{x}_2 \in G$ выполняется неравенство

$$|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \le L|\mathbf{x}_1 - \mathbf{x}_2| \tag{9}$$

Это условие означает, что $f(\mathbf{x})$ убывает и возрастает не быстрее линейной функции с заданным коэффициентом L.

З а м е ч а н и я :

1. Если неравенство (9) выполняется с константой L, то оно справедливо и при всех $L' > L$. Поэтому для функции, удовлетворяющей условию Липшица, существует бесконечное множество констант L из (9).

При использовании алгоритмов минимизации, включающих L как параметр, наилучшие результаты достигаются, как правило, если в качестве L берется минимальная из констант Липшица.

2. Из условия (9) непосредственно следует непрерывность $f(\mathbf{x})$ на отрезке $[a; b]$. Поэтому, согласно теореме Вейерштрасса, функция $f(\mathbf{x})$, удовлетворяющая на отрезке $[a; b]$ условию Липшица, имеет на нем хотя бы одну точку минимума.

3. Условие (9) означает, что модуль углового коэффициента любой хорды графика $f(\mathbf{x})$ не превосходит L.

Переходя в (9) к пределу при $|x_1 - x_2| \to 0$, убеждаемся, что если в некоторой точке существует касательная к графику функции $f(\mathbf{x})$, то модуль ее углового коэффициента также не может превышать L. Так, функция $f(x) = \sqrt{x}$ на отрезке [0; 1] условию Липшица не удовлетворяет, потому что при $x \to +0$ угловой коэффициент касательной к ее графику k неограниченно возрастает (рис. 21).

4. Если функция $f(x)$ имеет на отрезке $[a; b]$ непрерывную производную, то она удовлетворяет на этом отрезке условию Липшица с константой $L = \max|f'(x)|$.

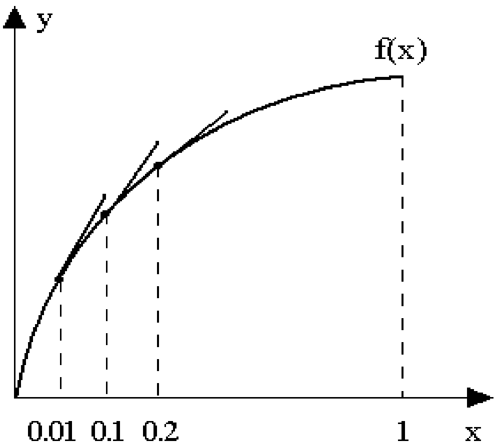

Рис. 21. График функции f(x) = $\sqrt{x}$, x$\in$[0; 1], не удовлетворяющей условию Липшица

По формуле конечных приращений для произвольных точек $x_1, x_2 \in [a; b]$ имеем:

$$f(x_1) - f(x_2) = f'(\xi)(x_1 - x_2),$$

где ξ – некоторая точка, лежащая между x_1 и x_2. Отсюда с учетом условия $|f'(\xi)| \leq \max|f'(\mathbf{x})| = L$ получаем неравенство (10) для $f(\mathbf{x})$.

5. Если $a = x_0 < x_1 < \ldots < x_n = b$, а функция $f(x)$ непрерывна на $[a; b]$ и удовлетворяет условию (5.9) на каждом из отрезков $[x_i; x_{i+1}]$, $i = 0,1,\ldots, n$-1, с константой L_i, то она удовлетворяет условию Липшица и на всем отрезке $[a; b]$ с константой $L = \max\limits_{0 \leq i \leq n-1} L_i$.

Использование условия Липшица позволяет ускорить процедуру детерминированного перебора значений проектируемых параметров при определении критерия оптимальности. На рис. 22 представлена функция одной переменной, удовлетворяющая условию Липшица на отрезке $0 \leq x \leq$ а с константой L_0. После того как найден локальный минимум $x_1^{\text{лок}}$ функции $f(x)$, удается ускорить поиск остальных локальных минимумов по сравнению с полным перебором на равномерной сетке.

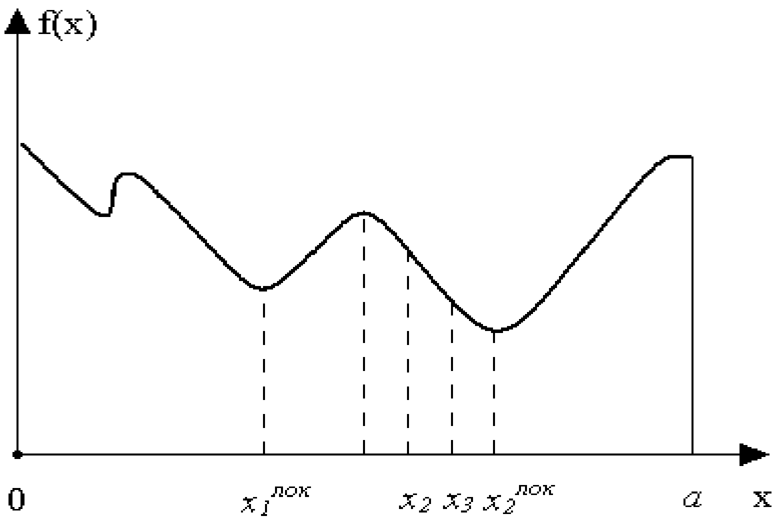

Рис. 22. Функция, удовлетворяющая условию Липшица

Пусть h_0 – некоторый шаг, осуществленный из точки локального минимума $x_1^{лок}$. Если в точке $x_1^{лок} + h_0$ значение функции f не меньше, чем в точке $x_1^{лок}$, т.е. $f(x_1^{лок} + h_0) \geq f(x_1^{лок})$, то, исходя из условия (9), можно определить ближайшую точку, в которой функция $f(x)$ может достигнуть значения $f(x_1^{лок})$ (подозрительного на глобальный минимум).

Действительно, согласно (9) функция $f(x)$ убывает не быстрее, чем линейная функция:

$$\varphi(\mathbf{x}) = f(x_1^{лок}. + h_0) + L_0[\mathbf{x} - (x_1^{лок} + h_0)].$$

Отсюда следует, что точка x_2, в которой необходимо произвести новое пробное вычисление функции $f(x)$, может быть получена из условия

$$\varphi(x) = f(x_1^{лок}),$$

т.е.

$$x_2 = (x_1^{лок} + h_0) + \frac{f(x_1^{лок} + h_0) - f(x_1^{лок})}{L_0} \quad (10)$$

Поскольку $f(x_2) \geq f(x_1^{\text{лок}})$ (рис. 22), к ней также можно применить сформулированный выше алгоритм для определения следующей пробной точки x_3.

В точке x_3 выполняется равенство $f(x_3) = f(x_1^{\text{лок}})$, поэтому из нее производится поиск локального минимума одним из известных методов. В точке $x_2^{\text{лок}}$ достигается значение меньшее, чем $f(x_1^{\text{лок}})$, поэтому все дальнейшие шаги до конца отрезка определяются исходя из $f(x_2^{\text{лок}})$. Ввиду того, что не оказывается новых точек, подозрительных на глобальный оптимум, поиск заканчивается.

2.2.2 Метод Ньютона – Рафсона

Метод Ньютона часто используется на завершающем этапе минимизации, когда x^* – точка минимума грубо найдена другим, менее трудоемким методом и требуется найти x^* с большой точностью. Кроме того, если функция $f(\mathbf{x})$ содержит члены, включающие $\mathbf{x}$ в третьей и более высоких степенях, то непосредственное получение аналитического решения уравнения $f'(\mathbf{x}) = 0$ может оказаться затруднительным. В таких случаях используются приближенные методы последовательного поиска стационарной точки функции $f(\mathbf{x})$. Ньютон разработал схему, ориентированную на нахождение корня нелинейного уравнения, которая позднее была уточнена Рафсоном.

В рамках схемы Ньютона – Рафсона предполагается, что $f(x)$ – дважды дифференцируемая функция, причем $f''(x) > 0$ (это гарантирует выпуклость $f(x)$). Работа алгоритма начинается в точке x_0, которая представляет начальное приближение координаты стационарной точки, или корня уравнения $f'(x) = 0$. В очередной точке x_k $(k = 0,1,\ldots)$ строится линейная аппроксимация функции $f^{'}(x)$, и точка, в которой аппроксимирующая линейная функция обращается в нуль, принимается в качестве следующего приближения x_{k+1} (рис. 23). Если точка x_k принята в качестве текущего приближения к стационарной точке, то уравнение касательной к графику точке $x = x_k$ имеет вид

$$y = f(x_k) + f^{'}(x_k)(x - x_k).$$

Точка $x = x_{k+1}$, найденная из условия $y = 0$, определяется формулой

$$x_{k+1} = x_k - f(x_k)/f^{'}(x_k).$$

Полагая $f(x) = f^{'}(x)$, тогда для решения уравнения $f^{'}(x) = 0$ необходимо построить последовательность

$$x_{k+1} = x_k - \frac{f^{'}(x_k)}{f^{''}(x_k)},\ k = 0,1,\ldots, \tag{11}$$

где x_0 – точка, выбранная в качестве начального приближения.

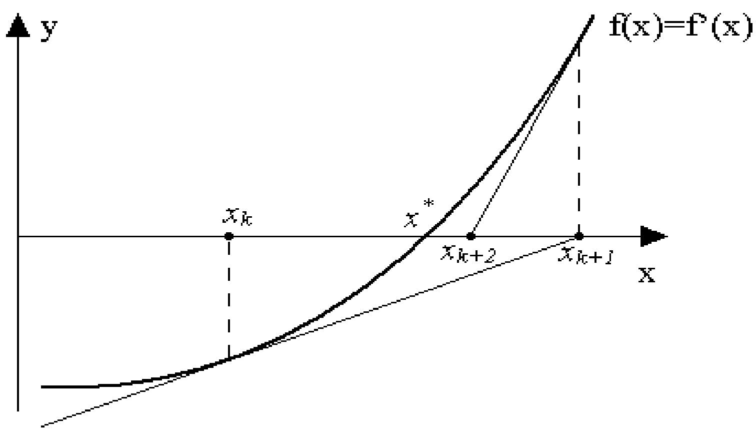

Рис. 23. Метод Ньютона – Рафсона (сходимость)

В зависимости от выбора начальной точки и вида функции алгоритм может как сходиться к истинной стационарной точке, так и расходиться, что отражено на рис. 24. Если начальная точка расположена правее x_0, то получаемые в результате последовательных приближений точки удаляются от стационарной точки x^*.

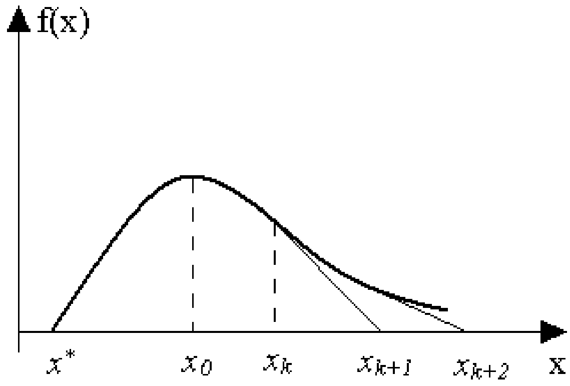

Рис. 24. Метод Ньютона – Рафсона (отсутствие сходимости)

Оценка скорости сходимости может быть сформулирована следующим образом. Пусть $f(x)$ – дважды дифференцируемая на E^n функция, причем $f''(x) \geq \mu > 0$ при всех $x \in E$ и $f''(x)$ удовлетворяет условию Липшица на X с константой L. Тогда, если начальное приближение x_0 удовлетворяет условию

$$q = \frac{.L}{2\mu^2}\left|f'(x_0)\right| < 1,$$

то последовательность (11) сходится к единственной точке минимума x^* функции $f(x)$ на X, причем

$$\left|x^* - x_k\right| \leq \frac{2\mu}{L} q^{2^k}, \ k = 0,1,\ldots$$

Вычисления по формуле (10) производят до тех пор, пока не выполнится неравенство $|f'(x_k)| \le \varepsilon$, после чего полагают $x^* \approx x_k$, $f^* \approx f(x_k)$.

На следующем рисунке приведена блок-схема алгоритма.

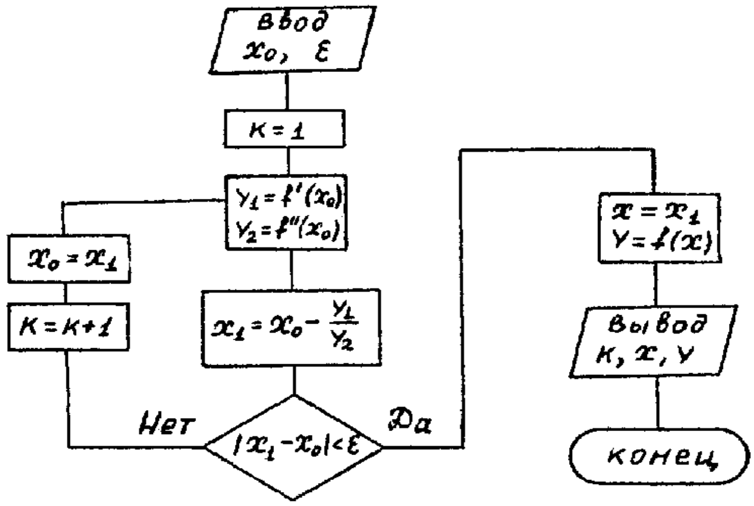

Рисунок 25. Блок - схема метода Ньютона – Рафсона

Пример 9. Рассмотрим следующую задачу: минимизировать

$$f(x) = 2x^2 + (16/x),\ \varepsilon = 0{,}05.$$

Для того чтобы определить стационарную точку функции $f(x)$, воспользуемся методом Ньютона – Рафсона, положив $x_1 = 1$:

$$f'(x) = 4x - (16/x^2),\ \ f''(x) = 4 + (32/x^3).$$

Итерация 1. $x_1 = 1$, $f'(x_1) = -12$, $f''(x_1) = 36$, $x_2 = 1-(-12/36) = 1{,}33$.

Итерация 2. $x_2 = 1{,}33$, $f'(x_2) = -3{,}73$, $f''(x_2) = 17{,}6$, $x_3 = 1{,}33 - \frac{-3{,}73}{17{,}6}$ = 1,54.

Итерация 3. $x_3 = 1{,}54$, $f'(x_3) = -0{,}59$,

$f''(x_3) = 12{,}77$, $x_4 = 1{,}54 - (-\frac{0{,}59}{12{,}77}) = 1{,}58$.

Итерация 4. $x_4 = 1{,}58$, $f'(x_4) = -0{,}09$,

$f''(x_4) = 12{,}12$, $x_5 = 1{,}58 - (-\frac{0{,}09}{12{,}12}) = 1{,}587$.

Итерация 5. $x_6 = 1{,}587 - \frac{0{,}005}{12{,}006} = 1{,}587$.

Так как $f'(x_5) < \varepsilon$, то $x^* \approx x_6 \approx 1{,}587$, $f^* \approx f(x_6) \approx 15{,}12$.

2.2.3 Метод секущих

Метод секущих является комбинацией метода Ньютона и общей схемы исключения интервалов. Как уже отмечалось, равенство $f'(x) = 0$ является необходимым и достаточным условием глобального минимума выпуклой дифференцируемой функции $f(x)$. Поэтому если на концах отрезка $[c; d]$ производная $f'(x)$ имеет разные знаки, т.е. $f'(c)\,f'(d) < 0$, то на интервале $(a; b)$ найдется точка, в которой $f'(x)$ обращается в нуль, и поиск точки минимума $f(\mathbf{x})$ на $[c; d]$ эквивалентен решению уравнения

$$f'(x) = 0,\ x \in (a; b) \tag{12}$$

Отсюда следует, что при $f'(c)\,f'(d) < 0$ любой приближенный метод решения уравнения (12) можно рассматривать как метод минимизации выпуклой дифференцируемой функции $f(x)$ на отрезке $[c; d]$.

Предположим, что в процессе поиска стационарной точки функции $f(x)$ в интервале $(a; b)$ обнаружены две точки c и d, в которых знаки производной различны. В этом случае алгоритм метода секущих позволяет аппроксимировать функцию $f'(x)$ “секущей прямой” и найти точку, в которой секущая графика $f'(x)$ пересекает ось абцисс (рис. 26). Таким образом, следующее приближение к стационарной точке x^* определяется по формуле:

$$z = d - \frac{f'(d)}{[f'(d) - f'(c)]/(d - c)}.$$

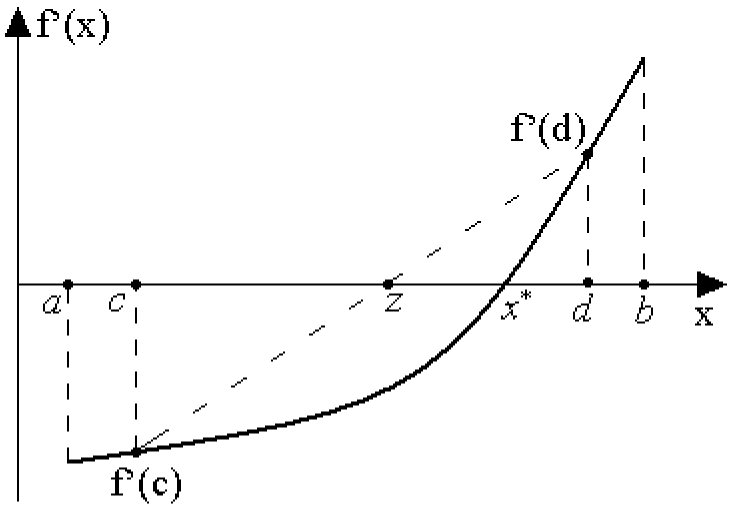

Рис. 26. Метод секущих

Если $|f'(z)| \le \varepsilon$, поиск следует закончить. В противном случае необходимо выбрать одну из точек c или d таким образом, чтобы знаки производной в этой точке и точке z были различны, а затем повторить основной шаг алгоритма. Например, в ситуации, изображенной на рис. 26, в качестве двух следующих точек должны быть выбраны точки z и d.

Легко видеть, что в отличие от метода средней точки метод секущих основан на исследовании не только знака, но и значений производной в пробных точках и поэтому в ряде случаев позволяет исключить более половины интервала поиска (см. рис. 26).

Пример 10. Опять рассмотрим следующую задачу: минимизировать

$$f(x) = 2x^2 + (16/x)$$

в интервале $1 \le \mathbf{x} \le 5$.

$$f'(x) = \frac{df(x)}{dx} = 4x - \frac{16}{x^2}.$$

И т е р а ц и я 1.

Шаг 1. $d = 5$, $c = 1$, $f'(d) = 19{,}36$, $f'(c) = -12$.

Шаг 2.

$$z = 5 - \frac{19{,}36}{(19{,}36 + 12)/4} = 2{,}53.$$

Шаг 3. $f'(z) = 7{,}62 > 0$; положить $d = 2{,}53$.

И т е р а ц и я 2.

Шаг 2.

$$z = 2{,}53 - \frac{7{,}62}{(7{,}62 + 12)/1{,}53} = 1{,}94.$$

Шаг 3. $f'(z) = 3{,}51 > 0$; положить $d = 1{,}94$. Итерации продолжаются до тех пор, пока не будет выполняться неравенство $|f'(z)| \le \varepsilon$.

Библиографический список

1. Корнеенко В.П. Методы оптимизации — М.: Высшая школа, 2007.

2. Пантелеев А.В. Методы оптимизации в примерах и задачах. Учебное пособие: 3-е изд. — М.: Высшая школа, 2008.

3. Гончаров В.А. Методы оптимизации: Учебное пособие – М.: Юрайт, 2010.

4. Струченков В.И. Методы оптимизации. Основы теории, задачи, обучающиекомпьютерные программы: 2-е изд., перераб. Учебное пособие. — М: Экзамен, 2007.

5. Симаков Е. Е., Ким Е. Решение транспортных задач с применением программирования в системе MathCAD // Молодой ученый. — 2014. — №5. — С. 8-13.

Электронное учебное издание

Ольга Викторовна **Свиридова**

Аналитические и численные методы решения одномерных задач оптимизации

Учебно-методическое пособие

Электронное издание сетевого распространения

Редактор Матвеева Н.И.

Темплан 2018 г. Поз. № 11.
Подписано к использованию 28.03.2018. Формат 60x84 1/16.
Гарнитура Times. Усл. печ. л. 3,56.

Волгоградский государственный технический университет.
400005, г. Волгоград, пр. Ленина, 28, корп. 1.

ВПИ (филиал) ВолгГТУ.
404121, г. Волжский, ул. Энгельса, 42а.