Н. Ю. Прокопенко

МЕТОДЫ ОПТИМИЗАЦИИ

Учебное пособие

Нижний Новгород
2018

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Нижегородский государственный архитектурно-строительный университет»

Н. Ю. Прокопенко

Методы оптимизации

Утверждено редакционно-издательским советом университета
в качестве учебного пособия

Нижний Новгород
ННГАСУ
2018

ББК 32.813я73
П78
УДК 004.89(075.8)

Рецензенты:

Тюсова М. К. – к. соц. н., доцент кафедры математики и системного анализа Нижегородского института управления Российской академии народного хозяйства и государ ственной службы при Президенте РФ (НИУ РАНХиГС).

Елесин А. В. – к. ф.-м. н., ведущий сотрудник НИИМ ННГУ.

Прокопенко Н. Ю. Методы оптимизации [Текст]: учеб. пособие /Н. Ю. Прокопенко; Нижегор. гос. архитектур. - строит. ун-т. – Н. Новгород: ННГАСУ, 2018. – 118 с. ISBN 978-5-528-00287-3

Учебное пособие предназначено для студентов, обучающихся по направлению подготовки 09.03.03 Прикладная информатика, профиль Прикладная информатика в экономике, дисциплине «Методы оптимизации».

ISBN 978-5-528-00287-3

Содержание

1. Оптимизационные задачи нелинейного программирования. Их классификация

Методы оптимизации – методы поиска экстремума функции при наличии ограничений или без ограничений очень широко используются на практике. Это, прежде всего, оптимальное проектирование (выбор наилучших номинальных технологических режимов, элементов конструкций, структуры технологических цепочек, условий экономической деятельности, повышение доходности и т. д.), оптимальное управление построением нематематических моделей объектов управления (минимизации невязок различной структуры модели и реального объекта) и многие другие аспекты решения экономических и социальных проблем (например, управление запасами, трудовыми ресурсами, транспортными потоками и т. д.).

Постановка задачи оптимизации

Заданы множество X и функция $f(x)$, определенная на X. Требуется найти точки минимума или максимума.

$f(x) \to \min \quad x \in X,$ (1)

где $f(x)$ – целевая функция;

X – допустимое множество;

$\forall x \in X$ – допустимая точка задачи.

В основном мы будем иметь дело с конечномерными задачами оптимизации, то есть с задачами, в которых допустимое множество X лежит в евклидовом пространстве R^n ($x \in R^n$).

Точка $x^* \in X$, являющаяся решением задачи, может быть точкой глобального или локального минимума.

Определение. Точка $x^* \in X$ называется

1) точкой глобального минимума функции $f(x)$ на множестве X или глобальным решением задачи оптимизации, если $f(x^*) \le f(x)$ при $\forall x \in X$. (2)

2) точкой локального минимума функции $f(x)$ на множестве X или локальным решением задачи оптимизации, если

$\exists \varepsilon > 0$, такое, что для $\forall x \in X \bigcap U_{\varepsilon}(x^{*})$ $f(x^{*}) \leq f(x)$ (3),

где

$U_{\varepsilon}(x^{*}) = \{x \in R^{n} \mid \|x - x^{*}\| \leq \varepsilon\}$ – шар радиуса $\varepsilon > 0$ с центром в x^{*}.

Если неравенство в (2) или в (3) выполняется как строгое при $x \neq x^{*}$, то x^{*} – точка строгого минимума (строгое решение) в глобальном или локальном смысле.

Ясно, что глобальное решение является и локальным; обратное неверно.

Для отражения того факта, что точка $x^{*} \in X$ является точкой глобального минимума функции $f(x)$ на множестве *X*, обычно используется запись $f(x^{*}) = \min\limits_{x \in X} f(x)$ или эквивалентная ей запись $x^{*} = \arg\min\limits_{x \in X} f(x)$.

Множество всех точек глобального минимума $f(x)$ на *X*, обозначают через $Arg\min\limits_{x \in X} f(x) = \{x^{*} \in X \mid f(x^{*}) = f^{*}\}$, где f^{*} – минимальное значение функции $f(x)$ на множестве *X*. В этом случае $\arg\min f(x)$ – это просто произвольная точка из множества $Arg\min\limits_{x \in X} f(x)$.

По аналогии с (1) записывают задачу максимизации функции $f(x)$ на множестве *X*, в виде $f(x) \to \max$, $x \in X$ (4).

Заменяя в данных выше определениях слово «минимум» на «максимум» и заменяя знак неравенств в (2) (3), на противоположный, получаем соответствующие понятия для задачи (4).

Решения задач (1) и (4), то есть точки минимума и максимума функции $f(x)$ на множестве *X* называются также ***точками экстремума***, а сами задачи (1), (4) – ***экстремальными задачами***.

Ясно, что задача (4) эквивалентна задаче $-f(x) \to \min$, $x \in X$, в том смысле, что множества глобальных и локальных, строгих и нестрогих решений этих задач соответственно совпадают. Это позволяет без труда переносить результаты, полученные для задачи минимизации, на задачи максимизации, и наоборот.

Пример. Рассмотрим график некоторой функции *Y=F(X)* (рис. 1). Тогда имеем:

X1 – точка локального максимума;

X2 – точка локального минимума;

X3 – точка глобального максимума;

X4 – точка глобального минимума.

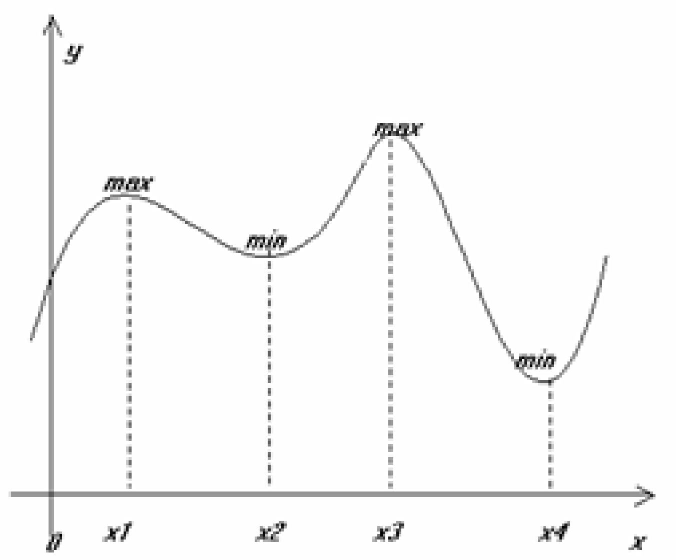


Рис. 1. Глобальные и локальные экстремумы

Замечание

1) Всякая точка глобального минимума является и точкой локального минимума этой функции. Обратное, вообще говоря, неверно.

2) Аналогичные определения глобального максимума и локального максимума можно получить путем замены знака неравенства на противоположный.

3) Если функция обладает свойством унимодальности, то локальный минимум автоматически является глобальным минимумом.

4) Если функция не является унимодальной, то возможно наличие нескольких локальных оптимумов, при этом глобальный минимум можно определить путем нахождения всех локальных оптимумов и выбора наименьшего из них.

При изучении задач оптимизации в первую очередь возникает вопрос о существовании решения. Ответ на этот вопрос дает теорема Вейерштрасса.

Теорема Вейерштрасса.

Пусть X – замкнутое ограниченное множество в R^n, $f(x)$ – непрерывная функция на множестве X. Тогда точка глобального минимума функции $f(x)$ на X существует (*глобальное решение задачи существует*).

Классификация задач оптимизации

Классификацию задач оптимизации можно проводить по нескольким признакам в зависимости от вида функции $f(x)$ и множества X:

1) детерминированные, стохастические, задачи оптимизации с неопределенностями; статические, динамические (например, задачи управления);
2) безусловной и условной оптимизации;
3) с непрерывными и дискретными переменными (частично-целочисленные, целочисленные, с булевыми переменными);
4) однокритериальные и многокритериальные;
5) линейные и нелинейные;
6) одномерные и многомерные, причем многомерные задачи могут быть малой и большой размерности;
7) с выпуклыми и невыпуклыми целевыми функциями;
8) одноэкстремальные и многоэкстремальные.

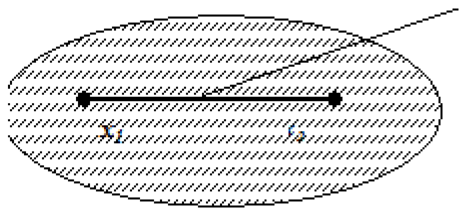
Нелинейные задачи составляют широкий класс настолько сложных задач, что до сих пор невозможно разработать общие методы, подобные симплекс-методу в линейном программировании, которые позволяли бы решить любые нелинейные задачи. Все возможные методы решения задач нелинейного программирования можно разделить на два больших класса: точные и приближенные методы решения. Точные методы позволяют определить решение для некоторых более узких задач, прежде всего задач с выпуклыми (вогнутыми) функциями *F(x)* и $g_i(x)$. Приближенные (итерационные) методы позволяют решить практически любую задачу нелинейного программирования, однако имеют свои недостатки: скорость сходимости (число шагов), точность и др.

Выпуклые множества

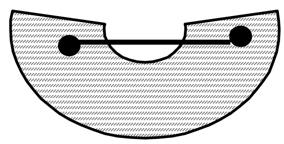
Понятие выпуклости играет важную роль в изучении задач оптимизации.

Определение. Непустое множество S из R^n называется выпуклым, если отрезок прямой, соединяющий любые точки множества S, также принадлежит этому множеству. То есть если $\forall x_1, x_2 \in S, \forall \lambda \in [0,1] \quad \lambda * x_1 + (1-\lambda) * x_2 \in S$.

Точка вида $x_3 = \lambda * x_1 + (1-\lambda) * x_2$ называется выпуклой комбинацией точек x_1 и x_2.



Выпуклое множество

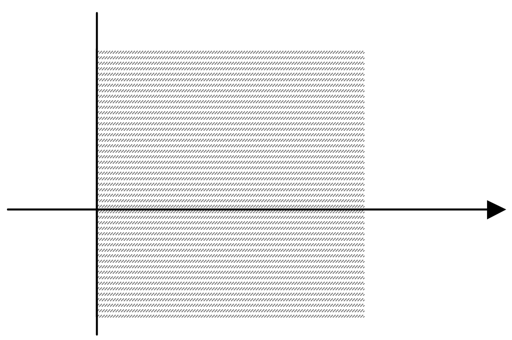


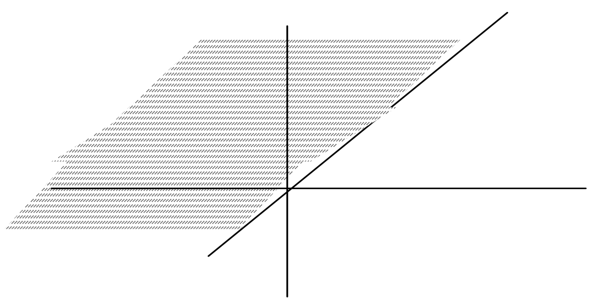
Невыпуклое множество

Примеры выпуклых множеств:

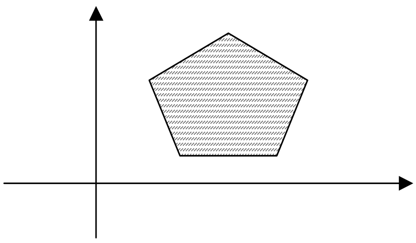
1) $S = \{x \in R^n : p^T * x \le \alpha\}$ – полупространство в R^n.

В пространстве $\boldsymbol{R^2}$.

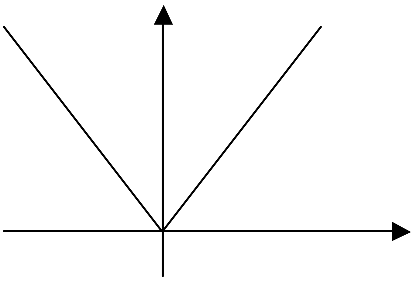




2) Многогранное множество. В R^2:



3) $S = \{x \in R^2 : x_2 \geq |x_1|\}$ – выпуклый конус в R^2.



4) $S = \{x \in R^2 : x_1^2 + x_2^2 \leq 4\}$ – круг с радиусом 2 и центром $(0, 0)$.

Выпуклые функции

Определение. Пусть $f : S \to R^n$, где S – непустое выпуклое множество в R^n. Функция f выпукла на множестве S, если для $\forall x_1, x_2 \in S$ и $\lambda \in (0,1)$:

$$f[\lambda * x_1 + (1-\lambda) * x_2] \leq \lambda * f(x_1) + (1-\lambda) * f(x_2) \quad (5).$$

Функция f строго выпукла на множестве S, если для $\forall x_1, x_2 \in S$, $x_1 \neq x_2$ и $\lambda \in (0,1)$: $f[\lambda * x_1 + (1-\lambda) * x_2] < \lambda * f(x_1) + (1-\lambda) * f(x_2)$.

Функция $f : S \to R^n$ называется вогнутой (строго вогнутой), если функция $-f$ выпукла (строго выпукла) на S.

Что означает соотношение (5): для выпуклой функции значение f в точках отрезка, соединяющего x_1 и x_2, не превосходит средневзвешенного (с тем же λ) значения величин $f(x_1)$ и $f(x_2)$.

Определение (для функции одной переменной). *График функции F(X) называется* ***вогнутым*** *(рис. 2) на интервале (A,B),* ***выпуклым*** *(рис. 3) на интервале (A,B), если все точки графика расположены ниже (выше) любой его касательной на этом интервале.*

Теорема 1. *Если функция F(X) дважды дифференцируема на интервале и* *во всех его точках* $f''(x) > 0 (f''(x) < 0)$, *то ее график* ***выпуклый (вогнутый)*** *на этом интервале.*

Определение. Точкой перегиба *графика функции называется точка этого графика, которая отделяет выпуклую его часть от вогнутой (рис. 4).*

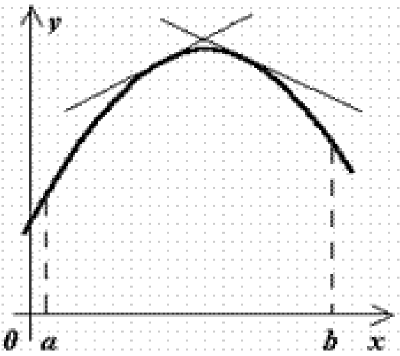


Рис. 2. Вогнутая функция

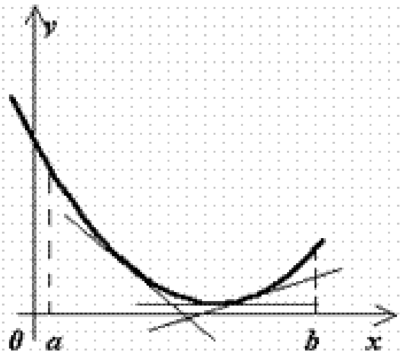


Рис. 3. Выпуклая функция

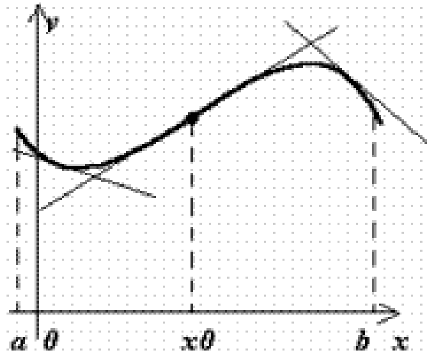


Рис. 4. Функция общего вида

Теорема 2. *Пусть функция F(X) определена в некоторой окрестности точки B и дважды дифференцируема в ней всюду, кроме, быть может, самой точки B. Тогда, если при переходе через точку B вторая производная меняет знак, то точка B есть* ***точка перегиба графика функции.***

Примеры выпуклых функций:

$f_1(x) = 3 * x + 4$;

$f_2(x) = |x|$;

$f_3(x) = x^2 - 2 * x$.

Понятие унимодальной функции.

Определение. $f(x)$ унимодальна на $[a,b]$, если существуют $\alpha, \beta \ : a \le \alpha \le \beta \le b$ 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 $\alpha = \beta$, то $f(x)$ строго унимодальна.

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

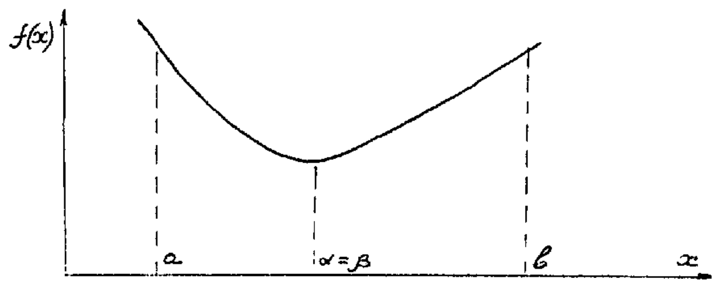


Рис. 5. Строго унимодальная, непрерывная, дифференцируемая функция

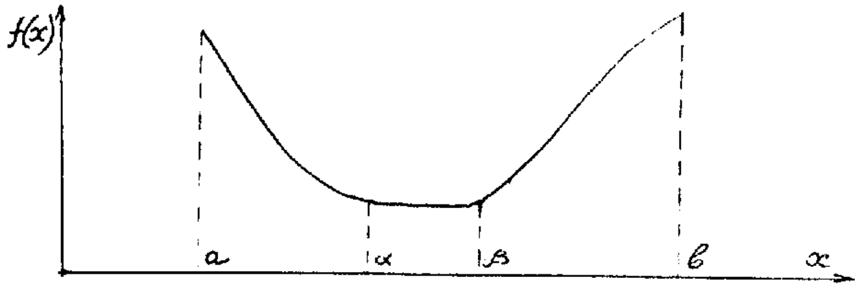


Рис. 6. Нестрого унимодальная, непрерывная, дифференцируемая функция

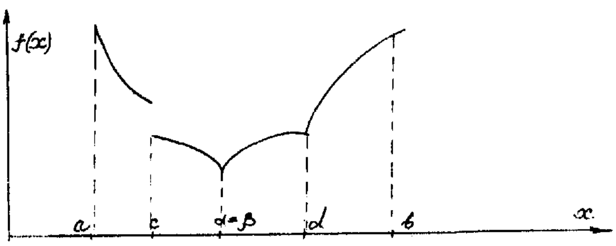


Рис. 7. При $x = c$ имеет разрыв I-го рода, при $x = \alpha$, $x = d$ производной у функции $f(x)$ не существует

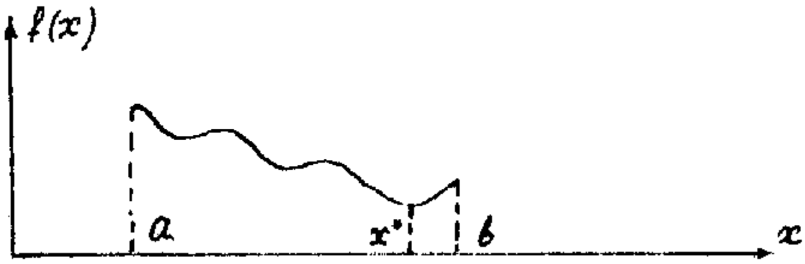


Рис. 8. Пример многоэкстремальной функции, где x^* – точка глобального минимума

Если функция f на отрезке $[a,b]$ многоэкстремальна (то есть существует несколько минимумов функции f), то возникает задача отыскания глобального минимума функции $f(x)$.

Замечание. *Для унимодальной функции локальный оптимум является глобальным.*

Исходя из определения унимодальной функции и направления выпуклости функции можно утверждать, что:

1) если унимодальная функция выпукла, то она имеет глобальный минимум;

2) если унимодальная функция вогнута, то она имеет глобальный максимум.

Алгоритм определения глобального максимума (минимума):

Шаг 1. Приравнять $\frac{df}{dx} = 0$ и найти все стационарные точки.

Шаг 2. Выбрать все стационарные точки, которые расположены в интервале *[A,B]*. Обозначим эти точки через $x_1, x_2, \ldots, x_n$. Проверку наличия локального оптимума следует проводить только на множестве указанных точек, дополненном точками *A* и *B*.

Шаг 3. Найти наибольшее (наименьшее) значение *F(x)* из множества $F(A)$, $F(B)$, $F(x_1)$, …, $F(x_n)$. Это значение соответствует глобальному максимуму (минимуму).

Общие сведения о численных методах оптимизации, их классификация

В большинстве случаев задачу оптимизации: $f(x) \to \min$, $x \in X$ 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:

1) 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);

2) на втором этапе, на основании полученной информации строится приближение к решению задачи – искомой точке минимума x^*, или, если такая точка не единственна, к множеству точек минимума.

Иногда, если только это требуется, строится и приближение к минимальному значению целевой функции $f^* = \min_{x \in X} f(x)$.

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:

1) нулевого порядка, в них используется информация только о значениях минимизируемой функции;

2) первого порядка, использующие информацию также и о значениях первых производных;

3) 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 x^{i+1} выбирается, когда уже выбраны точки предыдущих вычислений $x^1, x^2, .., x^i$ и в каждой из них произведены предусмотренные алгоритмом вычисления, результаты которых будем обозначать соответственно через $y^1, y^2, ..., y^i$.

Таким образом, последовательный алгоритм определяется точкой $x^1 \in X$ и набором отображений вида $\tilde{x}^{i+1} : \{x^1,...,x^i; y^1,...,y^i\} \to X \qquad i \ge 1$, при этом $x^{i+1} = \tilde{x}^{i+1}\{x^1,...,x^i; y^1,...,y^i\}$.

На практике обычно используются наиболее простые виды зависимости, например: $\tilde{x}^{i+1}\{x^1,...,x^i; y^1,...,y^i\} = \tilde{x}^{i+1}\{x^i; y^i\}$, то есть выбор точки очередного вычисления зависит лишь от точки предыдущего вычисления и полученного результата: $\tilde{x}^{i+1}\{x^1,...,x^i; y^1,...,y^i\} = \tilde{x}^{i+1} : \{x^i; \lambda_1 * y^1 + ... + \lambda_i * y^i\}$, то есть выбор x^{i+1} зависит от x^i и линейной комбинации всех полученных результатов (например, в методе сопряженных градиентов).

В дальнейшем для нахождения x^{k+1} будем пользоваться соотношением вида $x^{k+1} = x^k + \alpha_k * h^k, \quad \alpha_k \in \mathrm{R}, \; \mathrm{k} = 0, 1, 2...$

При этом конкретный алгоритм определяется:

1) заданием точки x^0;

2) правилами выбора векторов h^k и чисел α_k на основе полученной в результате вычислений информации;

3) условием остановки.

Правила выбора h^k и чисел α_k могут предусматривать вычисления характеристик не только в точках $x^0, x^1,..., x^k$, но и в других точках, отличных от $x^0, x^1,..., x^k$.

Вектор h^k определяет направление $k+1$-го шага метода минимизации, а коэффициент α_k – длину этого шага.

Обычно название метода минимизации определяется способом выбора h^k, а его различные варианты связываются с разными способами выбора α_k.

Наряду с термином <u>шаг метода</u> пользуется также термин <u>итерация метода</u>.

Среди методов минимизации можно условно выделить конечношаговые методы и бесконечношаговые методы.

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

Для *бесконечно шаговых* методов достижение решения гарантируется лишь в пределе.

Сходимость методов оптимизации

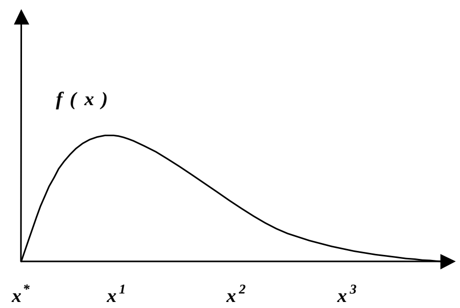
Важной характеристикой бесконечношаговых методов является сходимость.

Метод сходится, если $x^k \to x^*$ при $k \to \infty$, где x^* – решение задачи

$f(x) \to \min \quad x \in X$.

Если $f(x^k) \to f(x^*)$, то иногда также говорят, что метод сходится (по функции), при этом последовательность x^k называют минимизирующей. Минимизирующая последовательность может не сходиться к точке минимума.

Пример:



В данном примере минимизирующая последовательность $x^k = k$ не сходится к точке минимума $x^* = 0$.

В случае, когда точка минимума x^* не единственна, под сходимостью метода понимается сходимость x^k к множеству X^* точек минимума функции $f(x)$.

Пусть $x^k \to x^*$ при $k \to \infty$.

Говорят, что: 1) последовательность x^k сходится к точке x^* линейно (с линейной скоростью, со скоростью геометрической прогрессии), если существуют такие константы $q \in (0,1)$ и k_0, что

$$\left\| x^{k+1} - x^* \right\| \le \left\| x^k - x^* \right\| * q \quad \text{при} \quad k \ge k_0$$

2) последовательность x^k сходится к точке x^* сверхлинейно (со сверхлинейной скоростью), если:

$$\left\| x^{k+1} - x^* \right\| \le q_k * \left\| x^k - x^* \right\|, \ q_k \to 0_+ \ \text{при} \ k \to \infty$$

3) последовательность x^k сходится к точке x^* с квадратичной скоростью сходимости, если существуют такие константы $c \ge 0$ и k_0, что

$$\left\| x^{k+1} - x^* \right\| \le c * \left\| x^k - x^* \right\|^2 \quad \text{при} \quad k \ge k_0.$$

Большинство теорем о сходимости методов оптимизации доказывается в предположении о выпуклости целевой функции.

Для невыпуклых задач численные методы оптимизации позволяют отыскивать лишь локальные решения.

Условия остановки *(критерии окончания счета).*

Условие остановки может определяться имеющимися в наличии вычислительными ресурсами (например, числом вычислений характеристик минимизируемой функции) $f(x)$.

На практике часто используют следующие условия остановки:

$$\left\| x^{k+1} - x^k \right\| \le \varepsilon_1 \qquad (1)$$

$$\left\| f(x^{k+1}) - f(x^k) \right\| \le \varepsilon_2 \qquad (2)$$

$$\left\| f'(x^{k+1}) \right\| \le \varepsilon_3 \qquad (3)$$

Обычно пользуются одним из условий, но иногда используют критерии, состоящие в одновременном выполнении двух или всех трех условий (1) – (3).

Критерий (3) относится лишь к задачам безусловной оптимизации. Его выполнение означает, что в точке x^{k+1} с точностью ε_3 выполняется условие стационарности.

2. Безусловная нелинейная оптимизация функции одной переменной. Классификация методов

Постановка задачи.

$f(x)$ – числовая функция вещественной переменной x. Под задачей одномерной минимизации на отрезке $[a,b]$ понимается поиск $x^* \in [a,b]$ такого, что

$$f(x^*) = \min_{x\in[a,b]} f(x) \qquad (1)$$

$\min\limits_{x\in[a,b]} f(x)$ – наименьшее значение $f(x)$ на $[a,b]$;

x^*, удовлетворяющее (1), называется точкой минимума $f(x)$ на $[a,b]$;

X^* – множество точек минимума на $[a,b]$;

X^* может быть пустым, содержать конечное или бесконечное число точек.

Замечание. Задача поиска максимума $f(x)$ сводится к задаче поиска минимума: $\max f(x) = \min(-f(x))$.

Задача одномерной минимизации состоит из двух частей:

- локализации точек минимума, то есть указания отрезков, содержащих единственную точку минимума;
- поиска точки минимума на заданном отрезке при наличии информации о том, что на этом отрезке заведомо существует единственный минимум.

Задача локализации минимума обычно решается с помощью классического метода, основанного на дифференциальном исчислении.

Кроме того, существуют и некоторые вычислительные процедуры, позволяющие в определенных условиях такую задачу решать.

Классический подход

Напомним важные для данного рассмотрения теоремы из классического анализа.

Теорема Вейерштрасса. Если $f(x)$ непрерывна на $[a,b]$, то $\min\limits_{x\in[a,b]} f(x_0)$ существует.

Теорема Ферма. Пусть $f(x)$ дифференцируема в точке x^*. Если x^* доставляет локальный минимум $f(x)$, то $f'(x^*) = 0$.

Определение. x^* называется точкой локального минимума, если существует $\delta > 0$ такое, что для $x \in [x^* - \delta,\ x^* + \delta]$ выполняется $f(x) > f(x^*)$.

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

Тогда точками минимума могут быть такие точки, в которых:

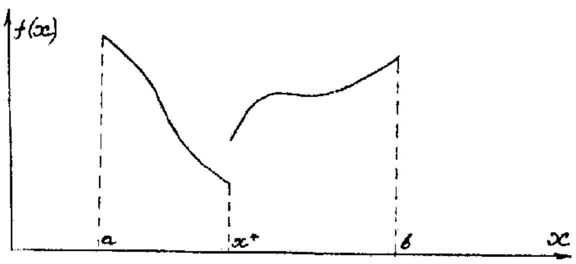
– $f(x)$ терпит разрыв;

– $f(x)$ непрерывна, но $f'(x)$ не существует;

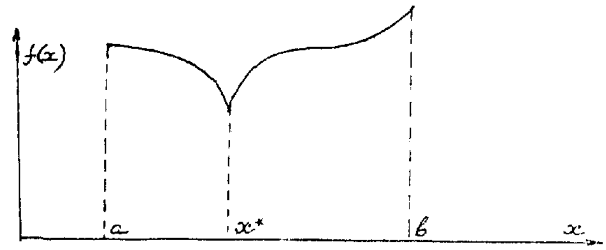
– $f'(x) = 0$;

– либо $x^* = a$, либо $x^* = b$.

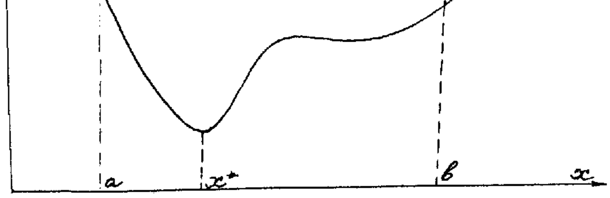
Рисунки ниже иллюстрируют эти 4 случая.



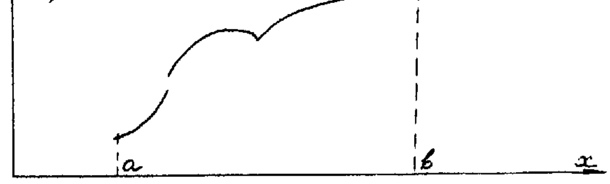
$f(x)$ терпит разрыв в точке x^*



$f(x)$ непрерывна, но производной не существует



$f'(x^*) = 0$



$x^* = a$

Теорема 1. Необходимые условия *того, что x* является точкой* ***Локального минимума (максимума)*** *дважды дифференцируемой функции F на открытом интервале (A,B), выражаются следующими соотношениями:*

1) $\left.\frac{df}{dx}\right|_{x=x^*}=0$,

2) $\left.\frac{d^2 f}{dx^2}\right|_{x=x^*}\geq 0\ (\leq 0)$.

Эти условия необходимы, но не достаточны для того, чтобы точка x^* была точкой локального минимума (максимума).

Определение. Стационарной точкой *называется точка x*, в которой*

$$\left.\frac{df}{dx}\right|_{x=x^*}=0.$$

Если стационарная точка не соответствует локальному оптимуму (минимуму или максимуму), то она является ***точкой перегиба или седловой точкой.***

Теорема 2. *Пусть в точке x* первые (n-1) производные функции обращаются в нуль, а производная порядка n отлична от нуля. Тогда:*

1) *если n – нечетное, то x* – точка перегиба;*

2) *если n – четное, то x* – точка локального оптимума.*

Кроме того,

А) *если эта производная положительная, то x* – точка локального минимума;*

Б) *если эта производная отрицательная, то x* – точка локального максимума.*

Замечание. *Выше предполагалось, что рассматриваемая функция дифференцируема, или что её первая производная существует и непрерывна. Однако если функция не является дифференцируемой во всех точках области определения, то даже необходимое условие наличия оптимума, позволяющее идентифицировать стационарные точки, может не выполняться в точке оптимума.*

Общие сведения о численных методах одномерной 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:

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

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

Порядок метода. Метод имеет порядок $\boldsymbol{k}$, если он использует информацию о производных $f(x)$ до $\boldsymbol{k}$-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 расходиться 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 $|x^{k+1} - x^k| \le \varepsilon$.

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: $|f(x^{k+1}) - f(x^k)| \le \varepsilon_1$ и $|f'(x^{k+1})| \le \varepsilon_2$, где $\varepsilon_1, \varepsilon_2$ – малые константы.

Методы минимизации 0-го порядка

Напоминаем, что эти методы оперируют только со значениями $f\{x\}$.

Метод перебора

Суть метода перебора состоит в следующем:

в N точках отрезка X $x_i = a + (2*i-1)*\dfrac{b-a}{2*N}$, $i = \overline{1,N}$ вычисляются значения функции f и в качестве минимального значения берут: $f(x^*) = \min\limits_{i=1,N} f(x_i)$.

Погрешность метода ε не превосходит величины $M*\dfrac{b-a}{2*N}$. Это легко показать, так как если x^* точка глобального минимума, то очевидно, найдется такая точка x_i, $i = \overline{1,N}$, что $|x_i - x^*| < \dfrac{b-a}{2*N}$.

Тогда с учетом условия Липшица:

$$0 \le \min_{j=1,N} f(x_j) - f(x^*) \le f(x_i) - f(x^*) \le M*|x_i - x^*| < M*\frac{b-a}{2*N}.$$

Таким образом, задаваясь необходимой величиной погрешности $\varepsilon = M * \frac{b-a}{2*N}$, можно определить требуемое количество точек N на отрезке $[a,b]$: $N = M * \frac{b-a}{2*\varepsilon}$.

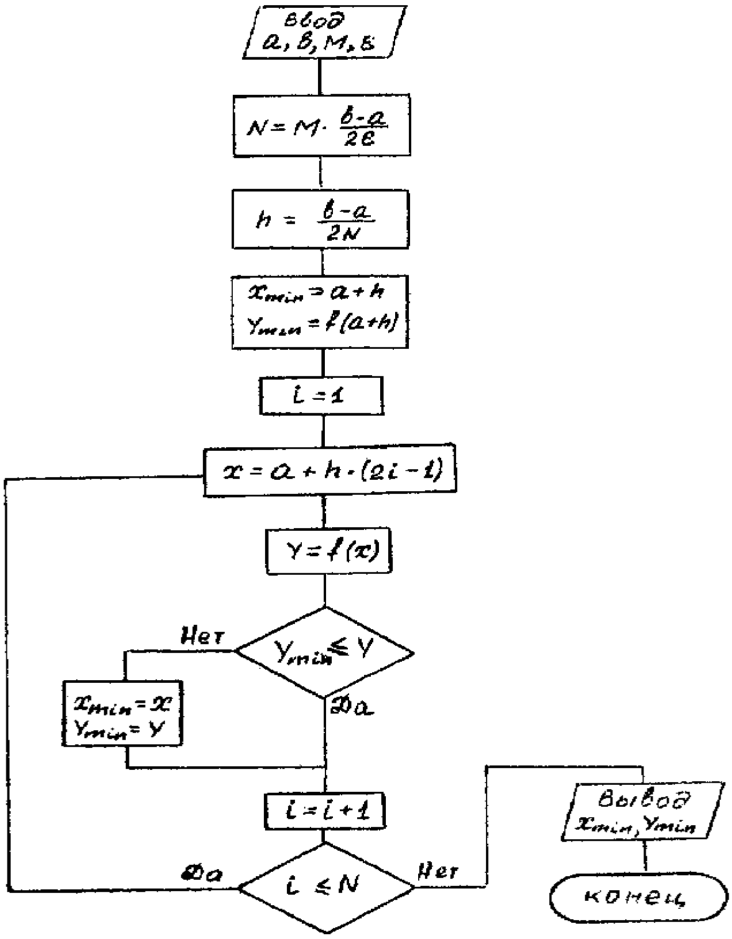


Рис. 9. Блок-схема метода перебора

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

Ранее было дано определение у*нимодальной функции*. Унимодальность функций является исключительно важным свойством. Фактически все одномерные методы поиска, используемые на практике, основаны на предположении, что исследуемая функция в допустимой области, по крайней мере, обладает свойством унимодальности. Полезность этого свойства определяется тем фактом, что для унимодальной функции *F(X)* сравнение значений *F(X)* в двух различных точках интервала поиска позволяет определить, в каком из заданных двумя указанными точками подинтервалов точка оптимума отсутствует.

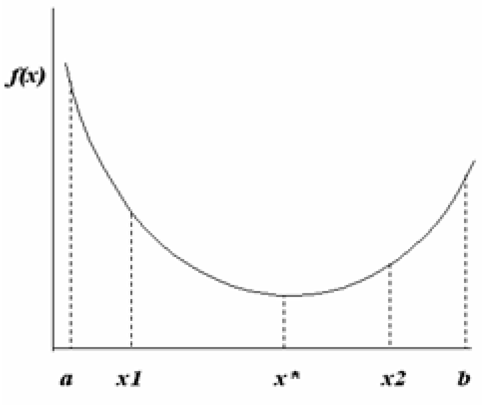


Рис. 10.

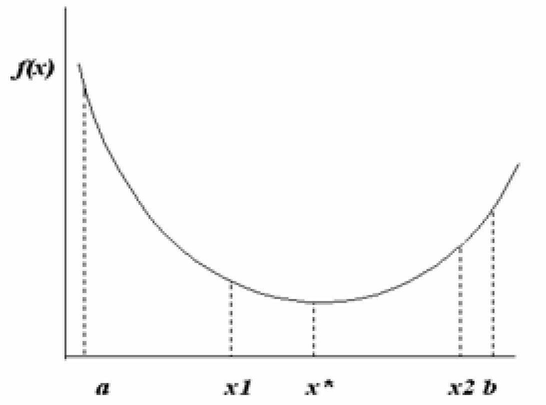


Рис. 11.

Теорема. *Пусть функция F унимодальна на замкнутом интервале* $a \le x \le b$, *а её 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(X) не лежит в интервале* (a, x_1), *т. е.* $x^* \in (x_1, b)$ (см. рис. 10).

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. 11).*

Замечание. *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 находиться в интервале* (x_1, x_2).

Согласно приведенной выше теореме, 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 подинтервал уменьшается до достаточно малых размеров. Заметим, что правило исключения интервалов устраняет необходимость полного перебора всех допустимых точек. Несомненным достоинством поисковых методов такого рода является то, что они основаны лишь на вычислении значений функций. При этом не требуется, чтобы исследуемые функции были дифференцируемы; более того, допустимы случаи, когда функцию нельзя даже записать в аналитическом виде. Единственным требованием является возможность определения значений функции $F(x)$ в заданных точках x с помощью прямых расчетов или имитационных экспериментов.

Вообще в процессе применения рассматриваемых методов поиска можно выделить два этапа:

- ***Этап установления границ интервала***, на котором реализуется процедура поиска границ достаточно широкого интервала, содержащего точку оптимума;

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

В данном разделе рассматриваются методы решения одномерных задач оптимизации вида

$$f(x) \to \max\{x \mid a \le x \le b\},$$

где x – скаляр, a и b – соответственно концы интервала, из которого берутся значения переменной x.

В основном рассматриваются алгоритмы, связанные с построением улучшающей последовательности. Решением задачи называется x^*, при котором $F(x^*) \ge F(x)$ для любого значения $a \le x \le b$. При практическом решении задач не будем различать два значения x_i и x_{i+1}, если $|x_i - x_{i+1}| \le \varepsilon$, где ε – задаваемая погрешность решения.

Метод сканирования

Метод заключается в последовательном переборе всех значений $a \le x \le b$ с шагом ε (погрешность решения) с вычислением критерия оптимальности F в каждой точке. Путем выбора наибольшего из всех вычислений значений F и находится решение задачи x^*.

Достоинство метода в том, что можно найти глобальный максимум критерия, если $F(x)$ – многоэкстремальная функция. К недостаткам данного метода относится значительное число повторных вычислений $F(x)$, что в случае сложной функции $F(x)$ требует существенных затрат времени.

На практике можно реализовать одну из основных модификаций метода – последовательное уточнение решения, или сканирование с переменным шагом (рис. 11).

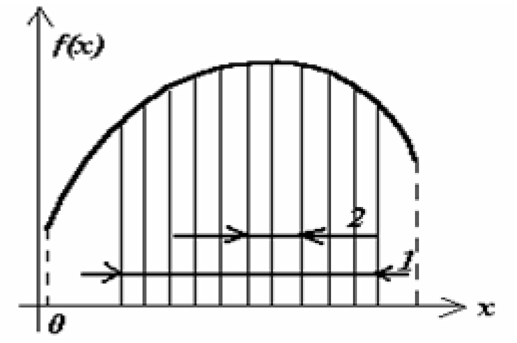


Рис. 11.

Рассмотрим иллюстрацию модифицированного метода сканирования:
1 – интервал, включающий в себя искомый максимум функции после первого этапа сканирования (исходный участок разбит на 5 участков); 2 – то же после второго этапа.

На первом этапе сканирование осуществляют с крупным шагом, затем отрезок, внутри которого получено наибольшее значение $F(x)$, разбивается на более мелкие отрезки, ищется новый отрезок, внутри которого уточненное значение максимума. Новый отрезок опять делится на более мелкие и т. д., тех пор, пока величина отрезка, содержащего максимальное значение $F(x)$, не будет меньше заданной погрешности.

Главный недостаток этого варианта метода – возможность пропуска «острого» глобального максимума $F(x)$.

2.1 Метод дихотомии

Необходимо найти $\min\limits_{x\in[a,b]} f(x)$ с точностью ε в предположении, что $f(x)$ унимодальна на $[a,b]$.

Это значит, что если x^* – точное решение задачи минимизации $f(x)$ на $[a,b]$, а $\tilde{a}$ – приближенное, то $|\tilde{a}-x^*|<\varepsilon$.

Метод основан на делении текущего отрезка $[a,b]$, где содержится искомый экстремум, на две равные части с последующим выбором одной из половин, в которой локализуется минимум в качестве следующего текущего отрезка. Экстремум локализуется путем сравнения двух значений критерия оптимальности в точках, отстоящих от середины отрезка на $\delta=\varepsilon/2$, где ε — погрешность решения задачи оптимизации.

Рассмотрим один шаг метода.

Точкой $c=(a+b)/2$ делим отрезок $[a,b]$ пополам, поскольку $f(x)$ унимодальна, то $f(c)<f(a)$ и $f(c)<f(b)$, однако точка минимума может оказаться как в левой, так и в правой части отрезка $[a,b]$.

На рисунке 12 представлены две унимодальные функции, имеющие точку минимума в разных половинах отрезка $[a,b]$.

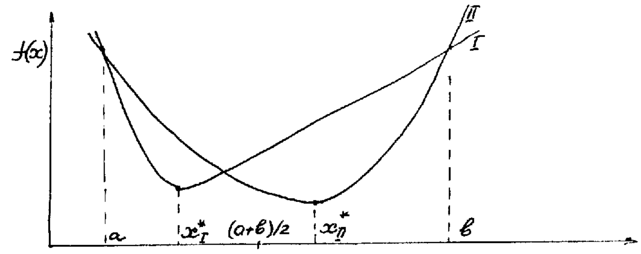


Рис. 12.

Следовательно, по значению функции в средней точке отрезка нельзя сузить отрезок неопределенности.

Внутри отрезка $[a,b]$ выберем две точки: $x_1=(a+b)/2-\delta/2$; $x_2=(a+b)/2+\delta/2$, где δ – параметр метода, $\delta<\varepsilon$.

<u>Замечание:</u> выбор малых констант должен быть согласован с машинной точностью, то есть δ, ε не должны быть меньше машинной точности.

В силу унимодальности $f(x)$ точка минимума x^* попадет либо в отрезок $[a, x_2]$, либо в $[x_1, b]$.

Если $f(x_1) < f(x_2)$, то $x^* \in [a, x_2]$ не может в этом случае попасть в отрезок $[x_2, b]$, так как нарушилась бы унимодальность $f(x)$.

Если $f(x_1) > f(x_2)$, то $x^* \in [x_1, b]$.

Если $f(x_1) = f(x_2)$, то $x^* \in [x_1, x_2]$. Но часто этот случай не выделяют отдельно, а включают знак равенства в один из выше рассмотренных случаев.

Таким образом, после первого шага метода постановка задачи осталась прежней с уменьшенным отрезком неопределенности, в котором находится точка минимума.

Многократное применение описанной выше процедуры приведет к тому, что отрезок неопределенности сузится до желаемого. Поскольку алгоритм работает в условиях леммы о вложенных отрезках, сходимость метода гарантирована.

На рисунке 13 представлена блок-схема метода дихотомии.

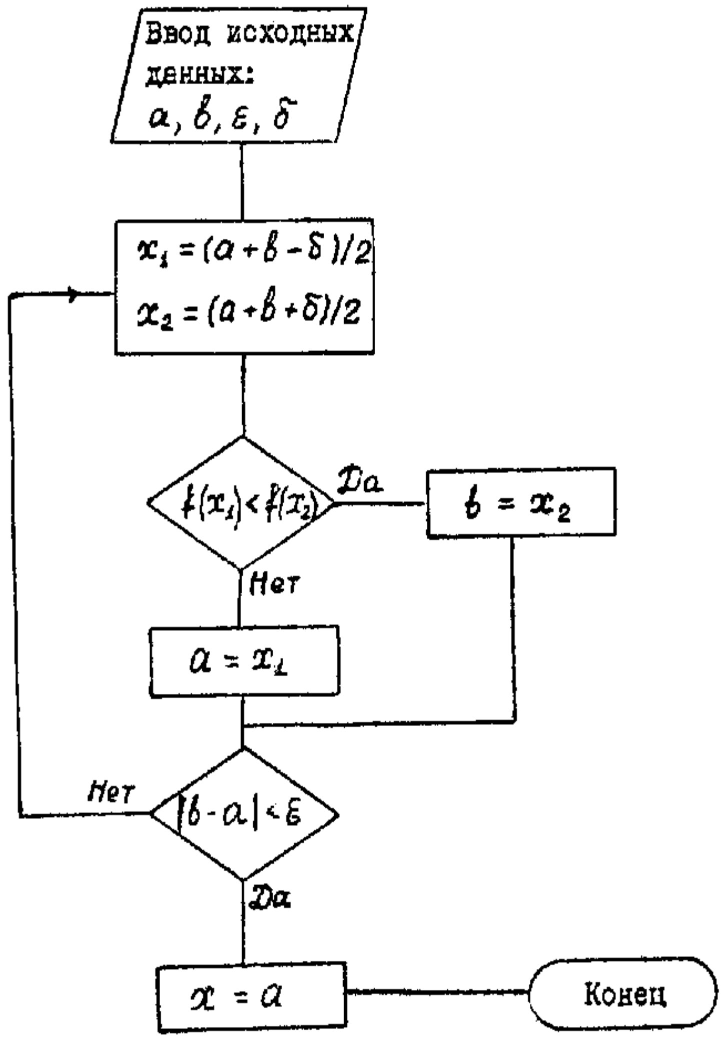


Рис. 13. Блок-схема метода дихотомии

Пометим индексами номер шага.

a^i – левый конец отрезка неопределенности при i- й итерации.

b^i – соответственно правый конец.

Тогда $b^1 - a^1 = \dfrac{b-a}{2} + \dfrac{\delta}{2}$ согласно описанию 1-го шага

$$b^2 - a^2 = \frac{b^1 - a^1}{2} + \frac{\delta}{2} = \frac{(b-a)/2 + \delta/2}{2} + \frac{\delta}{2} = (b-a+\delta)/4 + \delta/2 =$$

$$= (b-a-\delta)/4 + \delta = (b-a-\delta)/2^2 + \delta.$$

По математической индукции предположим, что

$$b^k - a^k = (b-a-\delta)/2^k + \delta.$$

Рассмотрим $b^{k+1} - a^{k+1}$:

$$b^{k+1}-a^{k+1}=(b^k-a^k)/2+\delta/2=\frac{(b-a-\delta)/2^k+\delta}{2}+\frac{\delta}{2}=(b-a-\delta)/2^{k+1}+\delta$$

Таким образом, методом математической индукции доказали, что после выполнения ***k***-го шага метода дихотомии $b^k-a^k=\dfrac{(b-a-\delta)}{2^k}+\delta$.

Поскольку задача решается с точностью ε, то необходимое число шагов метода должно удовлетворять неравенству: $(b-a-\delta)/2^k+\delta<\varepsilon$ или

$$\frac{(b-a-\delta)}{2^k}<\varepsilon-\delta \quad \Rightarrow 2^k>\frac{b-a-\delta}{\varepsilon-\delta} \;\Rightarrow\; k>\log_2\frac{b-a-\delta}{\varepsilon-\delta}.$$

Таким образом, заранее по постановке задачи мы можем узнать число шагов метода.

Замечание 1. Из последней оценки видим, что число шагов метода не зависит от вида $f(x)$.

Замечание 2. Этот метод можно применять для минимизации функций, не являющихся унимодальными, однако, без гарантии, что будет найден обязательно глобальный минимум.

Пример. Методом дихотомии 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.

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

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

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

2-й шаг.

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

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

3-й шаг.

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

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

4-й шаг.

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

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

5-й шаг.

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

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

6-й шаг.

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

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

7-й шаг.

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

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

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

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

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.

-1 1

-0,01 1 1-й шаг

-0,01 0,505 2-й шаг

-0.01 0,2575 3-й шаг

-0,01 0,12375 4-й шаг

-0,01 0,071375 5-й шаг

Рис. 14. Пошаговое уменьшение отрезка неопределенности

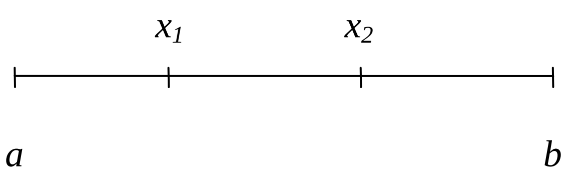
2.2. Метод золотого сечения

Метод золотого сечения также является последовательным методом минимизации. Опираясь на свойства золотого сечения отрезка, этот метод использует найденные значения *F(x)* более рационально, чем метод деления отрезка пополам, что позволяет переходить к очередному отрезку, содержащему точку *X** после вычисления одного, а не двух значений *F(x)*.

Метод основан на делении текущего отрезка [*a; b*], где содержится искомый экстремум, на две неравные части, подчиняющиеся правилу золотого сечения, для определения следующего отрезка, содержащего максимум.

Правило золотого сечения: *отношение всего отрезка к большей его части равно отношению большей части отрезка к меньшей.*

Рассмотрим следующую задачу. Возьмем отрезок $[a,b]$, найдем внутри этого отрезка $x_1,\ x_2\ \ (x_1 < x_2)$, образующие золотое сечение.



Для этого необходимо выполнение следующих условий:

$$\frac{x_2 - a}{b - a} = \frac{b - x_1}{b - a} = \frac{x_1 - a}{x_2 - a} = \frac{b - x_2}{b - x_1} = \frac{1}{\alpha}.$$

Найдем α, при котором возможны равенства

$$\frac{x_2 - a}{b - a} = \frac{1}{\alpha} \quad \Rightarrow \mathrm{x}_2 = \mathrm{a} + (b - a) * \frac{1}{\alpha} \ ,$$

$$\frac{b - x_1}{b - a} = \frac{1}{\alpha} \quad \Rightarrow x_1 = b - (b - a) * \frac{1}{\alpha} \ ,$$

$$\frac{x_1 - a}{x_2 - a} = \frac{1}{\alpha} \quad \Rightarrow \frac{b - (b - a) * \frac{1}{\alpha} - a}{a + (b - a) * \frac{1}{\alpha} - a} = \frac{1}{\alpha} \ ,$$

$$\frac{(b - a) * (1 - \frac{1}{\alpha})}{(b - a) * \frac{1}{\alpha}} = \alpha \quad \Rightarrow \frac{1}{\alpha^2} + \frac{1}{\alpha} - 1 = 0$$

$$\alpha^2 - \alpha - 1 = 0, \quad \alpha_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1 \pm \sqrt{5}}{2}$$

Поскольку $\alpha > 0 \quad \Rightarrow \ \alpha = \dfrac{1 + \sqrt{5}}{2} \approx 0{,}618$.

Метод золотого сечения состоит в том, что начиная с 1-го шага отрезок делится точками $x_1, \ \mathrm{x}_2$ в пропорции золотого сечения:

$$x_1 = a + \frac{3 - \sqrt{5}}{2}(b - a), \ \ x_2 = a + \frac{\sqrt{5} - 1}{2}(b - a).$$

При каждом шаге отрезок неопределенности уменьшается в $\alpha = 0{,}618$ раз.

$$b^1 - a^1 = \frac{b - a}{\alpha}$$

....

$$b^{\mathrm{n}} - a^n = \frac{b - a}{\alpha^n}$$

Если ε – заданная точность, то число шагов метода золотого сечения следует находить как решение неравенства

$$\frac{b - a}{\alpha^n} < \varepsilon \quad \Rightarrow \ \alpha^n > \frac{b - a}{\varepsilon} \ \Rightarrow n > \frac{\ln(\frac{b - a}{\varepsilon})}{\ln \alpha}.$$

Существуют аналитические формулы для расчета новой точки на отрезке, где находится минимальное значение $F(x)$.

При n шагах метода золотого сечения $f(x)$ вычисляется $n+1$ раз, так как на 1-м шаге $f(x)$ вычисляется дважды, а на последующих шагах по одному разу, при этом одна из внутренних точек отрезка неопределенности последующего шага совпадает с одной из точек предыдущего шага.

Метод золотого сечения обеспечивает более быструю сходимость к решению, чем многие другие методы, и применим, очевидно, только для одноэкстремальных функций, т. е. функций, содержащих один экстремум того типа, который ищется в задаче.

Алгоритм:

1 шаг. Определить точки x_1 и x_2, принадлежащие отрезку $[a;b]$ по формулам:

$$x_1 = a + \frac{3-\sqrt{5}}{2}(b-a), \quad x_2 = a + \frac{\sqrt{5}-1}{2}(b-a)$$

2 шаг. Проверяем критерий останова. Вычисляем ε_n:

$$\left| x^* - x_n \right| \le \varepsilon_n = \frac{\sqrt{5}-1}{2}(b-a).$$

Если $\varepsilon_n < \varepsilon$, то СТОП, иначе – следующий шаг.

3 шаг. Вычисляем $f(x_1)$ и $f(x_2)$ и сравниваем. Если $f(x_1) < f(x_2)$, то $b=x_2$, $x_2=x_1$, x_1 вычисляем. Если $f(x_1) > f(x_2)$, то $a=x_1$, $x_1=x_2$, x_2 вычисляем.

4 шаг. Переходим на 2-й шаг.

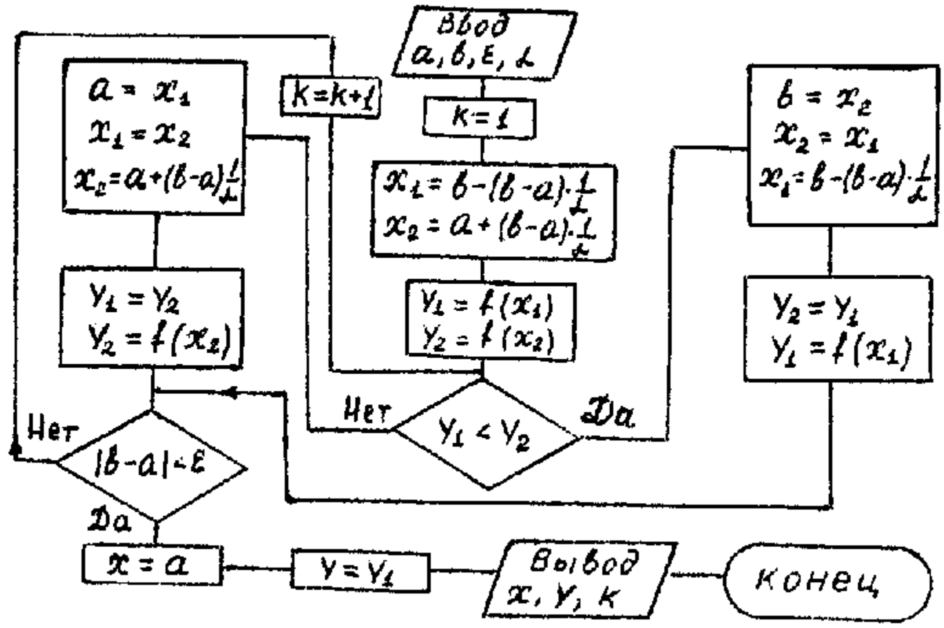


Рис. 15. Блок-схема метода золотого сечения

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

Предварительно определим, сколько потребуется шагов метода золотого сечения.

$$\frac{2}{(1{,}618)^n} < 0{,}05 \qquad \Rightarrow \qquad n > \frac{\ln(\frac{2}{0{,}05})}{\ln 1{,}618} = 7{,}7$$

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

1-й шаг.

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

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

2-й шаг. $x_1 = 0{,}2361$; $x_2 = -0{,}2361 + \dfrac{1{,}2361}{1{,}618} = 0{,}5279$

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

3-й шаг.

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

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

<u>4-й шаг.</u>

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

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

<u>5-й шаг.</u>

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

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

<u>6-й шаг.</u>

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

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

<u>7-й шаг.</u>

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

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

<u>8-й шаг.</u>

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

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

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

Сравнительная характеристика методов исключения интервалов

Выше были рассмотрены примеры решения задач различными методами исключения интервалов. Еще раз подчеркнем, что эти методы пригодны для любых непрерывных одноэкстремальных функций (для метода половинного деления необходимо, чтобы функция не имела горизонтальных участков). Сходимость методов и их эффективность не зависят от свойств функции. Основное достоинство метода сканирования заключается в снижении количества повторов вычисления для решения с заданной погрешностью, но при этом повышается вероятность пропуска «острого» глобального экстремума. Все три метода легко поддаются алгоритмизации. Для повышения точности нахождения решения необходимо просто уменьшить задаваемую погрешность. При сравнительном анализе можно сделать вывод, что метод золотого сечения оказывается более эффективным по сравнению с остальными двумя методами, поскольку он требует наименьшего числа оцениваний значения функции для достижения одной и той же заданной точности.

2.3 Метод Ньютона

Метод Ньютона относится к методам 2-го порядка и рекомендуется к применению в том случае, когда задача минимизации достаточно хорошо локализована.

Обычно это имеет место в том случае, когда на начальном этапе применяется один из методов 0-го порядка, а затем осуществляется переход к методу Ньютона. Для этого необходимым условием является гладкость $f(x)$, существование не равных нулю $f'(x)$, $f''(x)$, $f'''(x)$ для $x \in [a,b]$.

Тогда если x_k – k-е приближение к точке минимума, то расчетной формулой метода Ньютона будет формула $x_{k+1} = x_k - \dfrac{f'(x_k)}{f''(x_k)}$.

Метод Ньютона имеет высокую скорость сходимости:

$|x_{k+1} - x^*| < C * |x_k - x^*|^2$,

где x^* – точка минимума; C – некоторая положительная константа.

Процесс вычисления продолжается до тех пор, пока не будет достигнуто $|x_{k+1} - x_k| < \varepsilon$.

На следующем рисунке приведена блок-схема алгоритма.

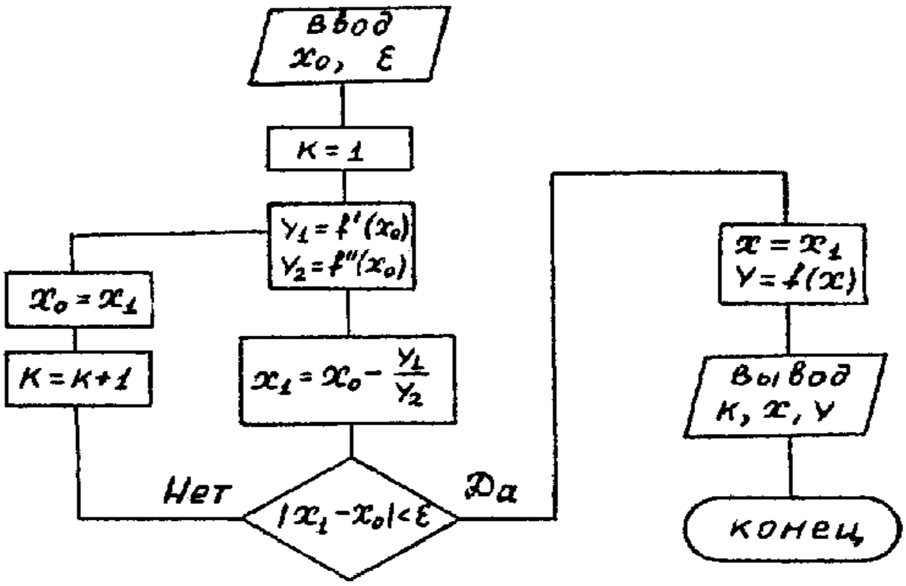


Рис. 16. Блок-схема метода Ньютона

Пример. Найти минимум $f(x) = (1-x) * \sin x$ на отрезке [0,1] с точностью $\varepsilon = 0{,}01$.

Зададим x_0, например, возьмем середину отрезка [0,1]: $x_0 = 0{,}5$.

Для данного примера расчетной формулой метода Ньютона будет

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} = x_k - \frac{(1-x_k)*\cos x_k - \sin x_k}{x_k * \sin x_k - 2*\cos x_k}$$

1-й шаг. $x_1 = 0{,}5 - \dfrac{(1-0{,}5)*\cos 0{,}5 - \sin 0{,}5}{0{,}5*\sin 0{,}5 - 2*\cos 0{,}5} = 0{,}47319$

$|x_1 - x_0| = |0{,}47319 - 0{,}5| = 0{,}0263 > 0{,}01$

2-й шаг. $x_2 = 0{,}47319 - \dfrac{(1-0{,}47319)*\cos 0{,}47319 - \sin 0{,}47319}{0{,}5*\sin 0{,}47319 - 2*\cos 0{,}47319} = 0{,}488169$

$|x_2 - x_1| = |0{,}48169 - 0{,}47139| = 0{,}0085 < 0{,}01$

Ответ: $x^* = 0{,}48$.

Замечание. Число сохраняемых знаков после округления определяется заданной точностью ε.

Сравнение методов одномерного поиска

Наилучшими критериями сравнения методов поиска, описанных выше, являются их эффективность и универсальность. Под эффективностью алгоритма обычно понимают число вычислений функции, необходимое для достижения требуемого сужения интервала неопределенности. Универсальность алгоритма означает, что его можно легко применить для решения самых разнообразных задач. С точки зрения универсальности малоэффективный метод деления отрезка пополам или метод сканирования имеет, по крайней мере, одно преимущество – его можно с успехом применять и для неунимодальных функций, если они достаточно гладкие. Большим достоинством этих методов является то, что при оптимизации функции не обязательно знать отрезок, содержащий оптимум. Методы, использующие квадратичную аппроксимацию, дают более быструю сходимость, чем методы исключения интервалов.

Нередко заранее известно, является ли рассматриваемая целевая функция унимодальной. В таких случаях следует воспользоваться несколькими разными алгоритмами и посмотреть, дают ли они все один и тот же оптимум. Отсюда следует один важный вывод, который следует иметь в виду, решая задачи оптимизации: не существует универсального алгоритма, который позволял бы решать любые задачи. Решая сложные задачи оптимизации, следует пользоваться разными методами, так как это позволяет увеличить долю удачных решений.

3. Безусловная оптимизация функции многих переменных

Под задачей многомерной безусловной оптимизации будем понимать задачу нахождения минимума или максимума функции $f(X)$, X – вектор n-мерного евклидова пространства; $x_1, x_2, ..., x_n$ компоненты вектора X. Обычно эта задача записывается как $f(X) \to \min\ (\max)$, $\quad X \in R^n$.

Определение. Экстремумом функции *двух переменных называется её максимальное или минимальное значение на заданном множестве изменения переменных.*

Экстремумы и методы их нахождения имеют широкое применение в экономических исследованиях, при выборе наилучших вариантов инвестиций, производственных программ, вложения денег в покупки и др.

<u>Определение.</u> *Точка M (x_0, y_0) называется точкой максимума (минимума) функции z = f(x, y), если существует окрестность точки M, такая, что для всех точек (x, y) из этой окрестности выполняется неравенство*

$f(x_0, y_0) \geq f(x, y)$

$(f(x_0, y_0) \leq f(x, y))$.

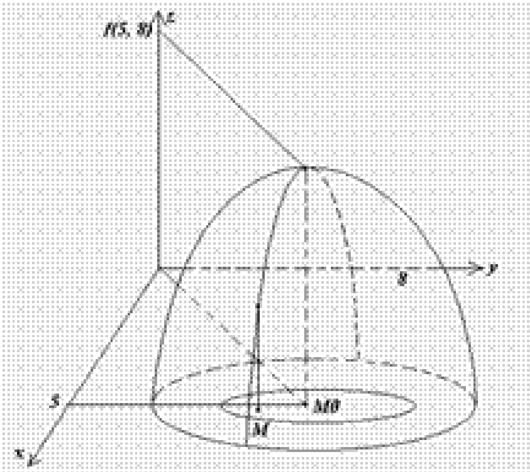


Рис. 17.

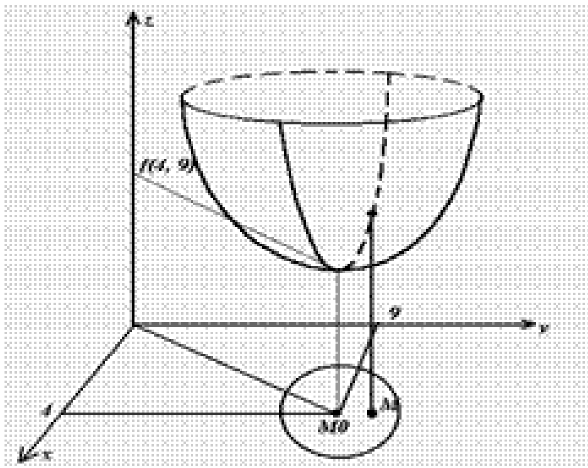


Рис. 18.

Пример. На рис. 17 представлен график функции двух переменных, точка $M_0(5, 8)$, в которой достигается максимум функции, окрестность точки $M_0(5, 8)$, максимальное значение функции $F(X, Y)$, равное $F(5, 8)$; на рис. 18 – график функции, точка $M_0(4, 9)$, в которой достигается минимум функции, окрестность точки $M_0(4, 9)$, минимальное значение функции $F(4, 9)$.

Приведенное определение экстремума является определением локального экстремума, функция может иметь несколько локальных максимумов или минимумов. Ясно, что при нахождении лучшего решения следует ориентироваться на наибольший из локальных максимумов, если ищется наибольшее значение функции, и на наименьший из локальных минимумов, если ищется наименьшее значение функции.

<u>Определение.</u> *Наибольшая величина из локальных максимумов называется **глобальным максимумом**, наименьший из локальных минимумов – **глобальным минимумом.***

Задача нахождения локальных экстремумов, а тем более глобальных, для функции нескольких переменных является достаточно трудной, в общем случае для произвольного числа переменных практически неразрешимой. Для выпуклых функций разработаны специальные методы нахождения экстремумов.

При исследовании функции часто возникает необходимость нахождения таких пар (x, y) из области определения функции, при которых функция принимает одинаковые значения. Рассмотрим множество точек (X,Y) из области определения функции $F(X, Y)$, таких, что $F(X, Y)=Const$, где запись *“Const”* означает, что значение функции зафиксировано и равно некоторому числу из области значений функции.

<u>Определение</u>. ***Линией уровня функции*** *$U=F(X, Y)$ называется линия $F(X, Y)=C$ на плоскости $XOУ$, в точках которой функция сохраняет постоянное значение $U=C$.*

Линии уровня геометрически изображаются на плоскости изменения независимых переменных в виде кривых линий. Получение линий уровня можно представить себе следующим образом. Рассмотрим множество C, которое состоит из точек трехмерного пространства с координатами $(X, Y, F(X, Y)=Const)$, которые, с одной стороны, принадлежат графику функции $Z=F(X, Y)$, с другой – лежат в плоскости, параллельной координатной плоскости $XOУ$, и отстоящей от неё на величину, равную заданной константе. Тогда для построения линии уровня достаточно поверхность графика функции пересечь плоскостью $Z=Const$ и линию пересечения спроектировать на плоскость $XOУ$. Проведенное рассуждение является обоснованием возможности непосредственно строить линии уровня на плоскости $XOУ$.

<u>Определение</u>. *Множество линий уровня называют **картой линий уровня**.*

Хорошо известны примеры линий уровня – уровни одинаковых высот на топографической карте и линии одинакового барометрического давления на карте погоды.

На рис. 19 изображены максимум, минимум и седловая точка в задаче оптимизации функции двух переменных при отсутствии ограничений.

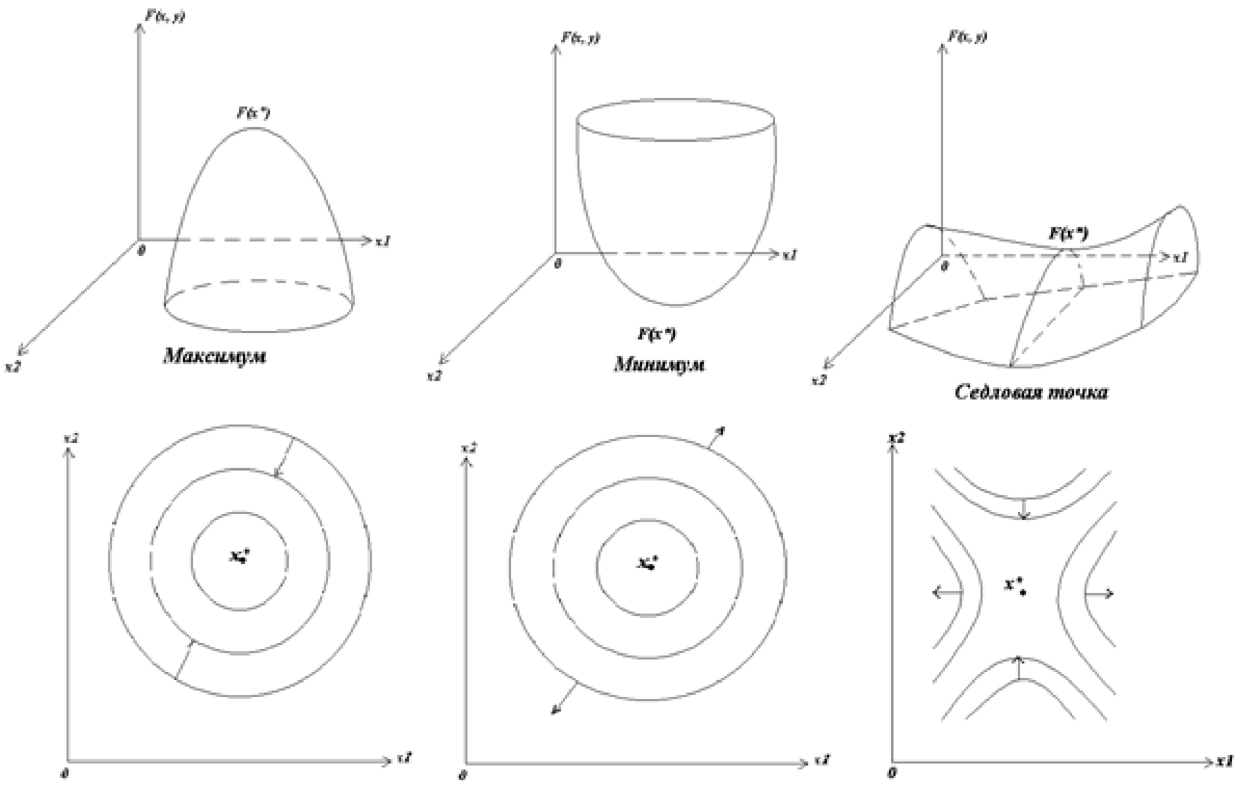


Рис. 19.

Пример. Найти линии уровня функции $U=X^2+Y^2$.

Решение. Уравнение семейства линий уровня имеет вид $X^2+Y^2=C$ $(C>0)$. Придавая C различные действительные значения, получим концентрические окружности с центром в начале координат.

Определение. ***Поверхностью уровня функции U=F(X, Y, Z)*** *называется поверхность $F(X, Y, Z)=C$, в точках которой функция сохраняет постоянное значение $U=C$.*

Пример. Найти поверхности уровня функции $U=X^2+Z^2-Y^2$.

Решение.

Уравнение семейства поверхностей уровня имеет вид $X^2+Z^2-Y^2=C$.

Если $C=0$, то получаем $X^2+Z^2-Y^2=0$ – конус; если $C<0$, то $X^2+Z^2-Y^2=C$ – *с*емейство двуполостных гиперболоидов.

Определение. *Градиентом функции* $u=f(x,y)$ в точке $M_0(x_0,y_0)$ называется вектор, координаты которого равны значениям частных производных функции в точке M_0:

$$\overline{grad}\, f(M_0) =$$

$$\left.\frac{df}{dx}\right|_{x=x^*} = 0,$$

$$\left.\frac{d^2 f}{dx^2}\right|_{x=x^*} \geq 0\ (\leq 0).$$

$$gra\overline{d}f(M_0) = \left\{\frac{du}{dx}\ \Big|_{M_0};\left.\frac{du}{dy}\right|_{M_0}\right\}.$$

Градиент функции направлен по нормали к поверхности уровня, т. е. перпендикулярно к касательной плоскости, проведенной в точке M в сторону наибольшего возрастания функции в данной точке.

Рассмотрим вектор, противоположный градиенту. Он называется ***антиградиентом***. Координаты этого вектора равны:

$$\left(-\frac{\partial f}{\partial x},-\frac{\partial f}{\partial y}\right).$$

Антиградиент функции $f(X, Y)$ в точке $M_0(x_0, y_0)$ указывает направление наиболее быстрого убывания функции в точке M_0. Любое направление, образующее острый угол с антиградиентом, является направлением убывания функции в этой точке.

Определение. *Матрицей Гессе* $H(x)$, дважды непрерывно дифференцируемой в точке x функции $f(x)$, называется матрица частных производных второго порядка, вычисленных в данной точке.

$$f''(x)=\begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Матрица Гессе

<u>Определение</u>. Функция *f(x)* называется *квадратичной формой*, если её можно представить в виде $f(x)=\sum_{i,j=1}^{n} a_{ij}x_i x_j$. Квадратичная форма – это функция на векторном пространстве, задаваемая однородным многочленом второй степени от координат вектора.

В матричном виде квадратичную форму можно представить в следующем виде: $f(x) = (1/2)x^TAx-b^Tx+c$ (3.1)

<u>Определение</u>. Квадратичная форма $x^T\cdot H(x)\cdot x$ и соответствующая ей матрица Гессе H(x) называется <u>положительно-определенной (H(x)>0)</u>, если для любого ненулевого вектора x справедливо следующее: $x^T\cdot H(x)\cdot x > 0$ (3.2).

Квадратичная форма $x^T\cdot H(x)\cdot x$ и соответствующая ей матрица Гессе H(x) называется <u>положительно-полуопределенной (H(x)≥0)</u>, если для любого ненулевого вектора x справедливо следующее: $x^T\cdot H(x)\cdot x \geq 0$.

Квадратичная форма $x^T\cdot H(x)\cdot x$ и соответствующая ей матрица Гессе H(x) называется <u>отрицательно-определенной (H(x)<0)</u>, если для любого ненулевого вектора x справедливо следующее: $x^T\cdot H(x)\cdot x < 0$.

Квадратичная форма $x^T\cdot H(x)\cdot x$ и соответствующая ей матрица Гессе H(x) называется <u>отрицательно-полуопределенной (H(x)≤0)</u>, если для любого ненулевого вектора x справедливо следующее: $x^T\cdot H(x)\cdot x \leq 0$.

Квадратичная форма $x^T\cdot H(x)\cdot x$ и соответствующая ей матрица Гессе H(x) называется <u>неопределенной</u> , если существуют ненулевые вектора x, для которых выполняются неравенства: $x^TAx < 0$ и $x^TAx > 0$.

Квадратичная форма $x^{T} \cdot H(x) \cdot x$ и соответствующая ей матрица Гессе H(x) называется тождественно равной нулю ($H(x) \equiv 0$), если для любого вектора x справедливо следующее: $x^{T} \cdot H(x) \cdot x = 0$.

На рисунке 20 изображено как выглядят квадратичные формы соответственно для положительно-определенной матрицы (a), отрицательно-определенной матрицы (b), положительно-неопределенной матрицы (c), неопределенной матрицы (d).

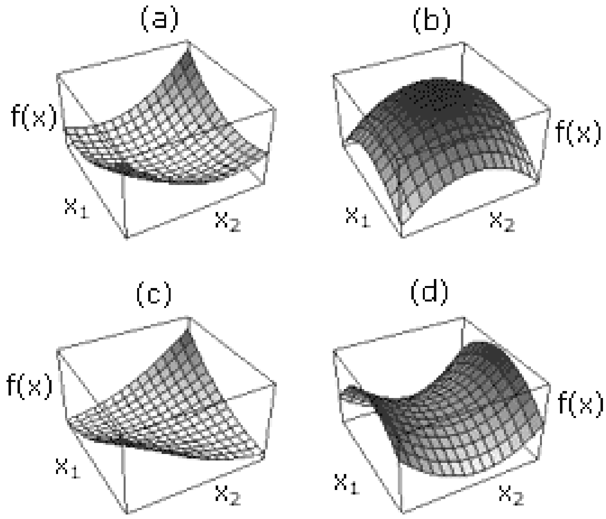


Рис. 20. Квадратичные формы для положительно-определенной, отрицательно-определенной, положительно-полуопределенной, неопределенной матрицы

3.1. Необходимые и достаточные условия безусловного экстремума

<u>Определение</u>. *Функция многих переменных может иметь максимум или минимум (экстремум) только в точках, лежащих внутри области определения функции, в которых все ее частные производные первого порядка равны нулю или не существуют. Такие точки называются* ***критическими*** *(или* ***стационарными****)*.

Теорема (необходимые условия экстремума первого порядка). *Пусть точка $X^* \in E_n$ – есть точка экстремума дифференцируемой функции f(X). Тогда частные производные $\dfrac{\partial f(X^*)}{\partial x_i} = 0$ в этой точке равны нулю (градиент равен нулю).*

Теорема (необходимые условия экстремума второго порядка).

Пусть точка $X^* \in E_n$ *есть точка локального минимума (максимума) дважды дифференцируемой в этой точке функции f(X). Тогда матрица Гессе* $H(X^*)$ *функции f(X), вычисленная в точке* X^*, *является положительно полуопределенной (отрицательно полуопределенной).*

Теорема (достаточные условия экстремума). *Пусть функция f(X) в точке* $X^* \in E_n$ *дважды дифференцируема, ее градиент равен нулю, а матрица Гессе является положительно определенной (отрицательно определенной). Тогда точка* X^* *– точка локального минимума (максимума) функции f(X).*

Критерий Сильвестра проверки достаточных условий экстремума.

1. Для того чтобы матрица Гессе $H(x^*)$ была положительно определена ($H(x^*)>0$) и стационарная точка x^* являлась точкой локального минимума, необходимо и достаточно, чтобы знаки угловых миноров были строго положительны:

$$M_1> 0, M_2> 0,\dots, M_n> 0. \quad (3.3)$$

2. Для того чтобы матрица Гессе $H(x^*)$ была отрицательно определена ($H(x^*)< 0$) и точка x^* являлась точкой локального максимума, необходимо и достаточно, чтобы знаки угловых миноров чередовались, начиная с отрицательного:

$$M_1< 0, M_2> 0, M_3< 0,\dots, (-1)^n M_n> 0. \quad (3.4)$$

Теорема (достаточное условие экстремума функции двух переменных). *Пусть функция* z= *f(x, y):*

а) определена в некоторой окрестности критической точки (x_0, y_0), *в которой* $f'_x(x_0, y_0) = 0$ *и* $f'_y(x_0, y_0) = 0$;

б) имеет в этой точке непрерывные частные производные второго порядка $f''_{xx}(x_0, y_0) = A$; $f'_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0) = B$; $f'_{yy}(x_0, y_0) = C$. *Тогда, если* $\Delta = AC - B^2 > 0$, *то в точке* (x_0, y_0) *функция* z= *f(x, y) имеет экстремум, при-*

чём если $A < 0$ *– максимум, если* $A > 0$ *– минимум. В случае* $\Delta = AC - B^2 < 0$, *функция* $z = f(x, y)$ *экстремума не имеет. Если* $\Delta = AC - B^2 = 0$, *то вопрос о наличии экстремума остаётся открытым (требуется дополнительное исследование).*

Исследование функции двух переменных на экстремум рекомендуется проводить по следующей **схеме:**

1. Найти частные производные функции z'_x и z'_y.

2. Решить систему уравнений $z'_x = 0$, $z'_y = 0$ и найти критические точки функции.

3. Найти частные производные второго порядка, вычислить их значения в каждой критической точке и с помощью достаточного условия сделать вывод о наличии экстремумов.

4. Найти экстремумы функции.

Пример.

Исследовать на экстремум функцию: $Z = F(X, Y) = X^3 + Y^3 - 3XY$.

Решение. Составим систему уравнений:

$$\begin{cases} \dfrac{\partial f}{\partial x} = 3x^2 - 3y = 0 \\ \dfrac{\partial f}{\partial y} = 3y^2 - 3x = 0. \end{cases}$$

Её решением являются пары (0, 0) и (1, 1), т. е. на экстремум надо проверить точки $M_0(0, 0)$ и $M_1(1, 1)$. Частные производные второго порядка имеют вид:

$\dfrac{\partial^2 f}{\partial x^2} = 6x = A$, $\dfrac{\partial^2 f}{\partial x \partial y} = -3 = B$, $\dfrac{\partial^2 f}{\partial y^2} = 6y = C$, определитель $D = 36xy - 9$.

Вычислим Δ в точках M_0 и M_1: $D\big|_{M_0} = -9 < 0$, значит экстремума в этой точке нет; $D\big|_{M_1} = 27 > 0$, при этом $A = 6 > 0$ и, следовательно, в точке M_1 – минимум.

Пример. Исследовать на экстремум функцию $u = x^3 + y^2 - 3x + 4\sqrt{y^5}$.

Решение. Ищем критические точки:

$$\begin{cases} u'_x = 3x^2 - 3 = 0 \\ u'_y = 2y + 2\sqrt{y^3} = 0. \end{cases}$$

Находим $M_0(1, 0)$ и $M_1(-1, 0)$. Эти точки принадлежат области определения исследуемой функции: $-\infty < x < +\infty$, $0 \leq y < +\infty$ (которая представляет половину плоскости *XOН*, лежащую выше оси *Ox*, включая и ось *Ox*), но они расположены не внутри этой области, а на её границе $y = 0$. Поэтому точки M_0 и M_1 не являются критическими. Частные производные по *X* и по *У* существуют во всей области определения функции *u*. Поэтому данная функция, как не имеющая критических точек, не имеет экстремума.

Пример. Исследовать на экстремум функцию $u = (x - y)^2 + (y - 1)^3$.

Решение. Ищем критические точки:

$$\begin{cases} u'_x = 2(x - y) = 0 \\ u'_y = -2(x - y) + 3(y - 1)^2 = 0. \end{cases}$$

Решая систему, найдем критическую точку функции *M*(1; 1).

Далее, чтобы установить, будет ли экстремум в точке *M*, вычисляем

$u''_{xx} = 2$, $u''_{xy} = -2$, $u''_{yy} = 2 + 6(y - 1)$, $A = 2$, $B = -2$, $C = 2$, $D = 0$.

Здесь оказалось, что $D = 0$. Чтобы установить, имеет ли экстремум функция *u* в критической точке *M*, исследуем знак её приращения

$\Delta u = u(M_1) - u(M) = (x - y)^2 + (y - 1)^3$ вблизи точки *M*.

Пусть M_1 лежит на биссектрисе $y = x$. Тогда $\Delta u = (y - 1)^3$. Если M_1 будет ниже *M*, т. е. если $y_{M1} < 1$, то $\Delta u < 0$, а если M_1 будет выше *M*, т. е. если $y_{M1} > 1$, то $\Delta u > 0$. Здесь оказалось, что вблизи точки *M* разность Δu не сохраняет знака, вследствие чего в точке *M* нет экстремума.

3.2 Численные методы, предназначенные для решения задачи нахождения безусловного экстремума функции многих переменных

Определение. Метод, использующий для своей реализации значения $f(X)$, а также ее производных до m-го порядка включительно, называется методом m-го порядка.

Будем рассматривать наиболее применяющиеся методы:

0–го порядка, использующие только значения $f(X)$;

1–го порядка, использующие значения $f(X)$, а также значения ее первых производных;

2–го порядка, использующие значения $f(X)$, а также значения ее первых и вторых производных.

Все рассматриваемые методы характеризуются тем, что строится некоторая последовательность векторов $x^k, k = 0,1,2....$, имеющая для всех методов, кроме метода случайного поиска, вид $X^{k+1} = X^k + \alpha_k * \overline{d}^k$,

k – номер члена последовательности или номер итерации.

Вектор $\overline{d}^k$ определяет направление $(k+1)$-го шага минимизации, α_k – длину этого шага.

Последовательность X^k сходится к точке минимума X^*, если $\lim\limits_{k \to \infty} X^k = X^*$, а метод, реализующий такую последовательность, называется сходящимся.

Естественно, что не всегда метод будет сходящимся, поэтому вопросу сходимости при рассмотрении конкретных методов уделяется одно из центральных мест.

Установление факта сходимости и оценка скорости сходимости дают существенную информацию о выбранном методе минимизации.

Поскольку речь идет о численных методах решения, необходимо задавать точность, с которой ищется решение.

Приближение X^k имеет точность ε, если $\left\|X^k - X^*\right\| \le \varepsilon$,

где $\|B\|$ означает норму вектора $\boldsymbol{B}$. В основном будем использовать норму $\|B\| = \max\limits_i B_i$, где B_i – i-я компонента вектора $\boldsymbol{B}$. В случае использования других норм будут сделаны специальные оговорки.

Итак, если применяется сходящийся численный алгоритм, необходимо выработать критерии остановки вычислительного процесса, гарантирующие выполнение требования по точности.

Обычно на практике применяются следующие критерии остановки:

1) $\left\|X^{k+1} - X^*\right\| \le \varepsilon_1$

2) $\left\|f(X^{k+1}) - f(X^*)\right\| \le \varepsilon_2$

3) $\left\|f'(X^{k+1})\right\| \le \varepsilon_3$

где $\varepsilon_1, \varepsilon_2, \varepsilon_3$ – малые константы;

$f'(X^k)$ – градиент функции $f(X)$, вычисленный в точке X^k.

Градиентом функции $f(X)$ в точке X^k называется вектор, компоненты которого есть частные производные $f(X)$ по компонентам вектора X, вычисленные в точке X^k.

Обычно используется либо один критерий, либо два из предложенных, либо все три критерия.

3.2.1. Методы нулевого порядка

В практических задачах нередки случаи, когда минимизируемая функция либо не обладает нужной гладкостью, либо является гладкой, но вычисление производных с нужной точностью является слишком трудоемким. В таких случаях применяют методы минимизации, которые требуют лишь вычисления значений функций, то есть методы 0-го порядка. Из методов этого типа рассмотрим метод покоординатного спуска.

Метод покоординатного спуска.

Суть метода состоит в том, что, задав начальное приближение, выбирается направление движения по одной из покоординатных осей, причем, на последующих шагах идет циклический перебор направлений по координатным осям.

Наиболее распространенным является метод покоординатного спуска с дроблением шага.

Обозначим через $e_i = (0,...,1,...0)$ – единичный координатный (базисный) вектор, у которого i-я координата равна 1, а остальные равны 0.

Положим $d_k = e_{i_k}$; $i_k = k - n[k/n] + 1$, $[k/n]$ – целая часть числа k/n.

Будем иметь $d_0 = e_1$, $\mathrm{d}_1 = \mathrm{e}_2, ..., \mathrm{d}_{n-1} = \mathrm{e}_n, ...$

Опишем подробно одну итерацию.

Пусть получено X^k. Будем искать X^{k+1}.

Вычислим значение функции $f(X)$ в точке $(X_k + \alpha_k * d_k)$ и проверим неравенство $f(X_k + \alpha_k * d_k) < f(X^k)$.

Если это неравенство выполняется, то полагаем:

$$X^{k+1} = (X_k + \alpha_k * d_k), \quad \alpha_{k+1} = \alpha_k.$$

В случае, если неравенство не выполняется, то вычисляем $(X_k - \alpha_k * d_k)$ и проверим неравенство $f(X_k - \alpha_k * d_k) < f(X^k)$.

В случае выполнения последнего неравенства полагаем

$$X^{k+1} = (X_k - \alpha_k * d_k), \quad \alpha_{k+1} = \alpha_k.$$

В случае невыполнения обоих неравенств полагаем

$$X^{k+1} = X_k, \quad \alpha_{k+1} = \begin{cases} \lambda * \alpha_k & \text{при } \mathrm{i}_k = \mathrm{n}, 2\mathrm{n}, \\ \alpha_k & \text{при } \mathrm{i}_k \neq \mathrm{n}, 2\mathrm{n}, \end{cases}$$

λ – параметр метода; $0 < \lambda < 1$.

Последние условия означают, что если за один цикл из $\boldsymbol{n}$ итераций при переборе направлений всех координатных осей $e_1, e_2, ..., e_n$ с шагом α_k реализовалась хотя бы одна удачная итерация, то длина шага α_k не дробится и сохраняется по крайне мере на протяжении следующего цикла из ***n*** итераций.

Если же среди последних ***n*** итераций не оказалось ни одной удачной, то шаг α_k дробится.

Сходимость метода обеспечена для гладких функций, несмотря на то, что это метод 0-го порядка. Оказывается, что если $f(X)$ не является гладкой, то метод покоординатного спуска может не сходиться к решению.

Другой вариант метода покоординатного спуска может состоять в получении α_k как решения задачи одномерной минимизации: $f(X_k + \alpha_k * d_k) \to \min$. Этот вариант имеет смысл применять в том случае, когда α_k можно найти явно.

Хотя скорость сходимости метода покоординатного спуска невысокая, благодаря его простоте и скромным требованиям к гладкости этот метод довольно широко применяется на практике.

На рис. 21 приведена блок-схема метода покоординатного спуска для функции дух переменных с оптимизацией длины шага.

<u>Пример.</u> Найти минимум $f(x_1, x_2) = 9 * x_1^2 + x_2^2 - 18 * x_1 + 6 * x_2 + 18$

Положим, $X^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$;

$$X^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha_0 * \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \end{pmatrix}.$$

Находим α_0 как решение задачи одномерной минимизации, а именно, ищем α_0, обеспечивающее минимум $9 * (\alpha_0 - 1)^2 + 9$, это будет $\alpha_0 = 1$.

Отсюда $X^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Далее $X^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_1 * \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \alpha_1 \end{pmatrix}$; $f(x^2) = (\alpha_1 + 3)^2 \quad \Rightarrow \alpha_1 = -3$

$$X^2 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

В результате выполнения 2-й итерации получили точное решение.

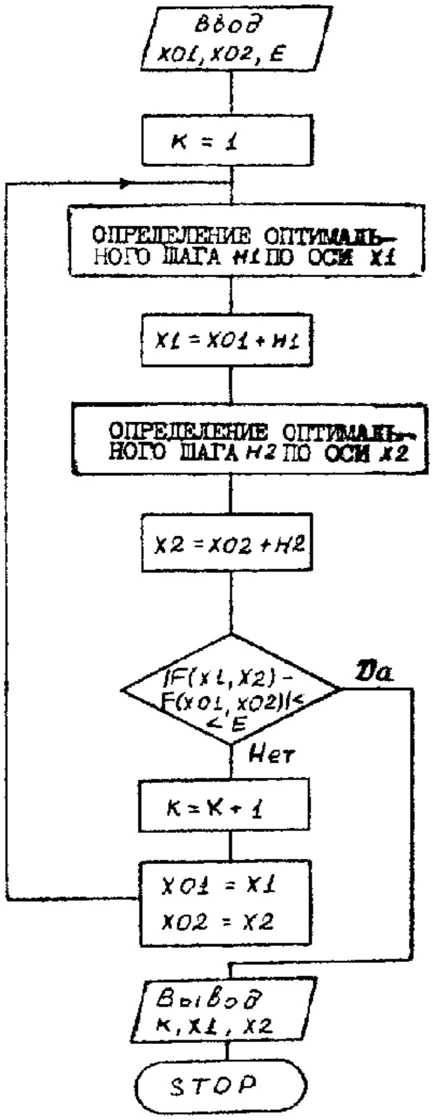


Рис. 21. Блок-схема метода покоординатного спуска

3.2.2 Методы первого порядка.

Общая характеристика градиентных методов.

Градиентные методы представляют собой приближенные (итерационные) методы решения задачи нелинейного программирования и позволяют решить практически любую задачу. Однако при этом определяется локальный экстремум. Поэтому целесообразно применять эти методы для решения задач выпуклого программирования, в которых каждый локальный экстремум является и глобальным.

Основным понятием, используемым во всех градиентных методах, является понятие градиента функции, как направления наискорейшего возрастания функции. Известно, что функция многих переменных $f(X)$ наиболее сильно возрастает в направлении градиента $grad\ f(X)$, а убывает в направлении антиградиента функции в этой точке, то есть в направлении вектора $-grad\ f(X)$.

Процесс решения задачи состоит в том, что, начиная с некоторой точки х (начальной), осуществляется последовательный переход в направлении $grad\ f(X)$, если определяется точка максимума, и $-grad\ f(X)$ (антиградиента), если определяется точка минимума, до точки, являющейся решением задачи.

Простейший алгоритм поиска *minf(X)* записывается в векторной форме следующим образом: $x^{i+1} = x^i - h \cdot gradf(x^i)$

или в скалярном виде: $x_j^{i+1} = x_j^i - h \cdot \dfrac{df}{dx_j^i}$ $\qquad j = 1,\ldots,n.$

Величина рабочего шага *h* в направлении градиента *grad f(x)* зависит от величины градиента, который заранее учесть трудно, и от коэффициента пропорциональности шага *h*, с помощью которого можно управлять эффективностью метода.

Поиск каждой новой точки состоит из двух этапов:

1) оценка градиента *f(X)* путем вычисления частных производных от *f(x)* по каждой переменной *Xj*;

2) рабочий шаг по всем переменным одновременно. Величина *h* сильно влияет на эффективность метода.

Градиентные методы можно разделить на два класса (группы). К первой группе относятся методы, в которых все исследуемые точки принадлежат допустимой области. К таким методам относятся: метод градиента, наискорейшего спуска, Франка-Вулфа и др. Ко второй группе относятся методы, в которых исследуемые точки могут и не принадлежать допустимой области. Общим из таких методов является метод штрафных функций. Все методы штрафных функций отличаются друг от друга способом определения «штрафа».

При определении решения градиентными методами итерационный процесс продолжается до тех пор, пока:

- либо $grad\ f(x^*) = 0$, (точное решение);
- либо $\left| f(x^{(i+1)}) - f(x^{(i)}) \right| < \varepsilon,$ (3.5)

где $x^{(i)}, x^{(i+1)}$ – две последовательные точки, $\varepsilon \geq 0$ – малое число, характеризующее точность решения.

Метод градиента.

Представим человека, стоящего на склоне оврага, которому необходимо спуститься вниз (на дно). Наиболее естественным, кажется, направление в сторону наибольшей крутизны спуска, т.е. направление ($-grad\ f(X)$). Получаемая при этом стратегия, называемая **градиентным методом**, представляет собой последовательность шагов, каждый из которых содержит две операции:

а) определение направления наибольшей крутизны спуска (подъема);

б) перемещение в выбранном направлении на некоторый шаг.

Правильный выбор шага имеет существенное значение. Чем шаг меньше, тем точнее результат, но больше вычислений. Различные модификации градиентного метода и состоят в использовании различных способов определения шага. Если на каком-либо шаге значение $f(X)$ не уменьшилось, это означает, что точку минимума «проскочили», в этом случае необходимо вернуться к предыдущей точке и уменьшить шаг, например, вдвое.

Схема решения.

1. Определение $x^0 = (x_1, x_2, \ldots, x_n)$, принадлежащей допустимой области и $grad\ f(x^0)$.

2. Определение $grad\ f(x^0)$ или $-grad\ f(x^0)$.

3. Выбор шага *h*.

4. Определение следующей точки по формуле $x^{(k+1)} = x^{(k)} \pm h\ grad\ f(x^k)$, («+», если max, «-», если min.

5. Определение $grad\ f(x^{k+1})$ и :

– если $\left| f(x^{(k+1)}) - f(x^{(k)}) \right| < \varepsilon$, решение найдено;

– если нет, то переход к п. 2.

Замечание. Если $grad\ f(x) = 0$, то решение будет точным.

Метод наискорейшего спуска.

В отличие от метода градиента, в котором градиент определяют на каждом шаге, в методе наискорейшего спуска градиент находят в начальной точке и движение в найденном направлении продолжают одинаковыми шагами до тех пор, пока значение функции уменьшается (увеличивается). Если на каком-либо шаге $f(X)$ возросло (уменьшилось), то движение в данном направлении прекращается, последний шаг снимается полностью или наполовину и вычисляется новое значение градиента и новое направление.

Если на каждом шаге выбирать h_k как значение, обеспечивающее $\min\limits_{\alpha} f(X^k - h_k * f'(X^k))$, то мы и получим метод наискорейшего спуска, то есть на каждой итерации необходимо решать задачу одномерной минимизации, которая в основном решается численно.

Итак, для того чтобы использовать метод наискорейшего спуска, необходимо задать правила вычислений $f(X)$, $f'(X)$ (предварительно продифференцировав $f(X)$), выбрать метод одномерной минимизации.

На рис. 22 представлена укрупненная блок-схема такого алгоритма, где критерием окончания счета выбрана близость градиента нулю.

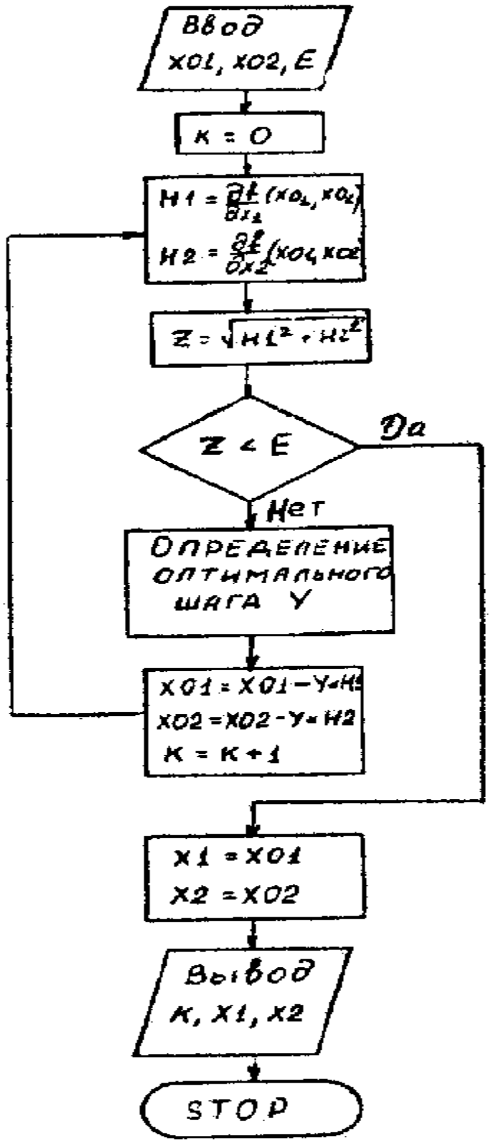


Рис. 22. Блок-схема метода наискорейшего спуска

Пример. Найти минимум $f(x_1, x_2) = 9x_1^2 + x_2^2 - 18x_1 + 6x_2 + 18$.

Методом наискорейшего спуска вычисления будут производиться по расчетной формуле $X^{k+1} = X^k - h_k \cdot \begin{pmatrix} 18x_1^k - 18 \\ 2x_2^k + 6 \end{pmatrix}$.

В качестве начального приближения возьмем $X^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

h_k на каждой итерации находится классическим методом, то есть приравниванием производной к нулю.

Результаты вычислений по итерациям представлены в таблице.

Номер итерации	α_k	x_1^k	x_2^k
0	0,06097	0	0
1	0,2778	1,09756	-0,36585
2	-0,06	0,60976	-1,82928
3	0,322	1,0312	-1,96977
4	0,0592	0,8504	-2,6332
5	0,3218	1,0098	-2,6766
6	0,05578	0,95303	-2,8847
7	0,3782	1,00293	-2,8976
8	0,04596	0,98646	-2,975
9		0,99766	-2,9773

Точным решением этой задачи является $x_1 = 1, \quad x_2 = -3$, проведенные 9 итераций не обеспечили получение приближенного решения с точностью 10^{-3}.

Изменим немного вид исходной функции.

Пусть $f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1 + 6x_2 + 10$.

Нетрудно показать, что точкой минимума и этой функции будет $(1,-3)$.

Применим метод наискорейшего спуска, начав с точки $(0,0)$.

Расчетная формула метода имеет вид

$$X^{k+1} = X^k - h_k \cdot \begin{pmatrix} 2x_1^k - 2 \\ 2x_2^k + 6 \end{pmatrix}, \qquad X^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Найдем h_0 из условия $f(2h_0, -6h_0) \to \min$. Получили $h_0 = 0{,}5$.

Отсюда $x_1 = 1, \quad x_2 = -3$, то есть за один шаг попали в точку минимума.

Чем же разнятся эти задачи, дающие разные по трудоемкости вычислительные процедуры?

Представим линии уровня каждой из функций.

Для $f(x_1, x_2) = 9x_1^2 + x_2^2 - 18x_1 + 6x_2 + 18 = 9(x_1 - 1)^2 + (x_2 + 3)^2$

линиями уровня будут кривые $9(x_1 - 1)^2 + (x_2 + 3)^2 = C$

или $\dfrac{(x_1 - 1)^2}{C/9} + \dfrac{(x_2 + 3)^2}{C} = 1.$

Получили каноническое уравнение эллипса, из которого видно, что одна из полуосей в 3 раза меньше другой, то есть эллипс вытянут вдоль оси x_2.

Для $f(x_1, x_2) = (x_1 - 1)^2 + (x_2 + 3)^2$ линиями уровня будут концентрические окружности с центром в точке $(0,0)$.

На рис. 23 с нанесенными на плоскость линиями уровня представлена траектория движения из точки $(0,0)$ в точку $(1,-3)$, соответствующая методу наискорейшего спуска для функции $f(x_1, x_2) = 9x_1^2 + x_2^2 - 18x_1 + 6x_2 + 18$.

На рис. 23 то же проделано для функции $f(x_1, x_2) = (x_1 - 1)^2 + (x_2 + 3)^2$.

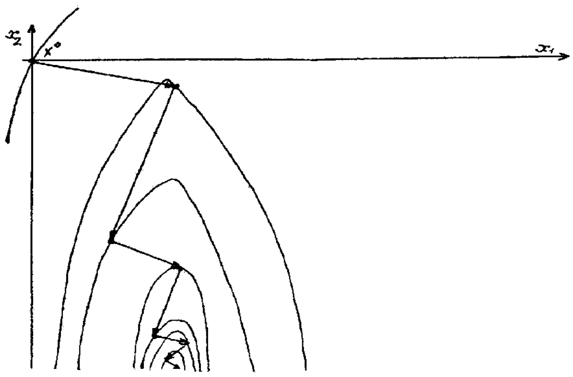


Рис. 23. Траектория движения из точки $(0,0)$ в точку $(1,-3)$ функции

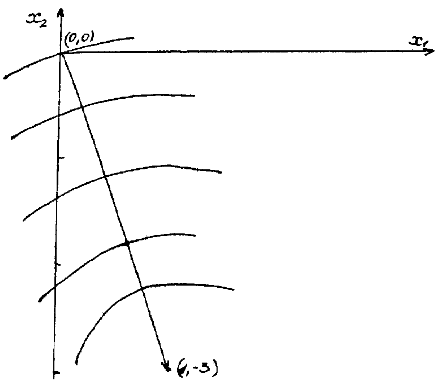


Рис. 24. Траектория движения из точки $(0,0)$ в точку $(1,-3)$ функции $f(x_1, x_2) = (x_1 - 1)^2 + (x_2 + 3)^2$.

Чем больше вытянуты линии уровня, тем сильнее будет эффект «зигзага» траектории. В этом случае функция имеет так называемый «овражный» характер, то есть небольшое изменение переменной x_1 приводит к резкому изменению значений функции, а по переменной x_2 функция меняется незначительно. В процессе реализации градиентного метода очередные приближения будут прыгать со склона на склон, что может сильно замедлить сходимость метода. Этой проблеме уделяется достаточно серьезное внимание, в настоящее время существуют специальные методы для оптимизации овражных функций.

Метод сопряженных градиентов.

Название метода отражает тот факт, что данный метод отыскания безусловного экстремума сочетает в себе понятия градиента целевой функции и сопряженных направлений.

Обозначения, используемые далее.

Скалярное произведение двух векторов записывается $x^{\mathrm{T}}y$ и представляет сумму скаляров: $\sum_{i=1}^{n} x_i y_i$. Заметим, что $x^{\mathrm{T}}y = y^{\mathrm{T}}x$. Если x и y ортогональны, то $x^{\mathrm{T}}y = 0$. В общем, выражения, которые преобразуются к матрице 1x1, такие как $x^{\mathrm{T}}y$ и $x^{\mathrm{T}}Ax$, рассматриваются как скалярные величины.

Квадратичная форма – это просто скаляр, квадратичная функция некого вектора x следующего вида: $f(x) = (1/2)\cdot x^{\mathrm{T}}Ax - b^{\mathrm{T}}x + c$ (3.6).

Определение. Матрица A называется положительно-определенной, если для любого ненулевого вектора x справедливо следующее: $x^{\mathrm{T}}Ax > 0$ (3.7).

Матрица A называется положительно-полуопределенной, если для любого ненулевого вектора x справедливо следующее: $x^{\mathrm{T}}Ax \geq 0$.

Матрица A называется отрицательно-определенной, если для любого ненулевого вектора x справедливо следующее: $x^{\mathrm{T}}Ax < 0$.

Матрица A называется отрицательно-полуопределенной, если для любого ненулевого вектора x справедливо следующее: $x^{\mathrm{T}}Ax \leq 0$.

Матрица A называется неопределенной, если существуют ненулевые векторы x, для которых выполняются неравенства: $x^{\mathrm{T}}Ax < 0$ и $x^{\mathrm{T}}Ax > 0$.

Если матрица A – положительно-определенная, можно найти минимум ее квадратичной функции. Причем, метод сопряженных градиентов сделает это за n или менее шагов, где n – размерность неизвестного вектора x. Так как любая гладкая функция в окрестностях точки своего минимума хорошо аппроксимируется квадратичной, этот же метод можно применить для минимизации и неквадратичных функций. При этом метод перестает быть конечным, а становится итеративным.

Рассмотрение метода сопряженных градиентов целесообразно начать с рассмотрения более простого метода поиска экстремума функции – метода наискорейшего спуска. На рис. 25 изображена траектория движения в точку минимума методом наискорейшего спуска. Суть этого метода:

- в начальной точке $x(0)$ вычисляется градиент, и движение осуществляется в направлении антиградиента до тех пор, пока уменьшается целевая функция;
- в точке, где функция перестает уменьшаться, опять вычисляется градиент, и спуск продолжается в новом направлении;
- процесс повторяется до достижения точки минимума.

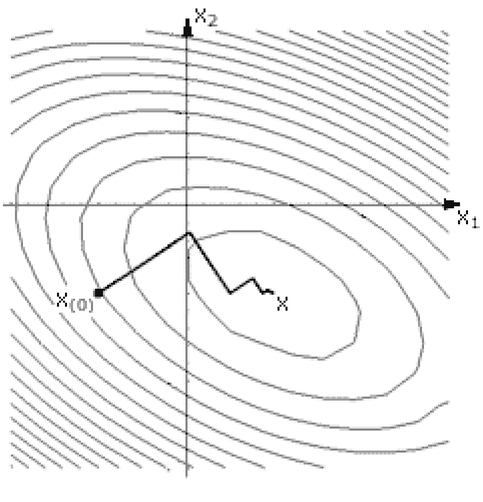


Рис. 25. Траектория движения в точку минимума методом наискорейшего спуска

В данном случае каждое новое направление движения ортогонально предыдущему. Не существует ли более разумного способа выбора нового направления движения? Существует, и он называется метод сопряженных направлений. А метод сопряженных градиентов как раз относится к группе методов сопряженных направлений. На рис. 26 изображена траектория движения в точку минимума при использовании метода сопряженных градиентов.

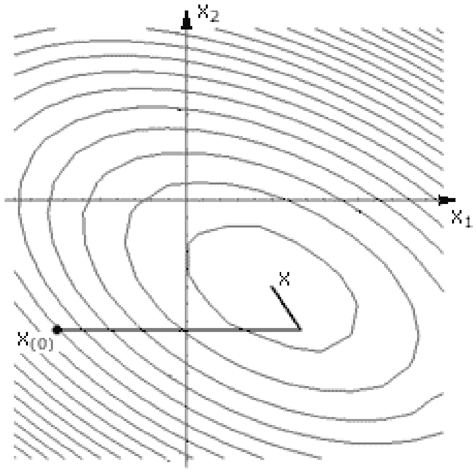


Рис. 26. Траектория движения в точку минимума при использовании метода сопряженных градиентов

Определение сопряженности формулируется следующим образом: два вектора x и y называют A-сопряженными (или сопряженными по отношению к матрице A) или A-ортогональными, если скалярное произведение x и Ay равно нулю, то есть $x^{\mathrm{T}}Ay = 0$.

Сопряженность можно считать обобщением понятия ортогональности. Действительно, когда матрица A – единичная матрица, в соответствии с равенством $x^{\mathrm{T}}Ay = 0$, векторы x и y – ортогональны. Можно и иначе продемонстрировать взаимосвязь понятий ортогональности и сопряженности: мысленно растяните рис. 26 таким образом, чтобы линии равного уровня из эллипсов превратились в окружности, при этом сопряженные направления станут просто ортогональными.

Существуют итеративные способы вычисления сопряженного направления, самый известный – формула Флетчера-Ривса:

$$d_{i+1} = r_{i+1} + \beta_{i+1} d_i \qquad (3.8),$$

где $$\beta_{i+1} = \frac{r_{i+1}^T r_{i+1}}{r_i^T r_i} \qquad (3.9).$$

Формула (3.8) означает, что новое сопряженное направление получается сложением антиградиента в точке поворота и предыдущего направления движения, умноженного на коэффициент, вычисленный по формуле (3.9). Направления, вычисленные по формуле (3.8), оказываются сопряженными, если минимизируемая функция задана в форме (3.5). То есть для квадратичных функций метод сопряженных градиентов находит минимум за n шагов (n – размерность пространства поиска). Для функций общего вида алгоритм перестает быть конечным и становится итеративным. При этом, Флетчер и Ривс предлагают возобновлять алгоритмическую процедуру через каждые n + 1 шагов.

Можно привести еще одну формулу для определения сопряженного направления, формула Полака–Райбера (Polak-Ribiere):

$$\beta_{i+1} = \frac{r_{i+1}^T (r_{i+1} - r_i)}{r_i^T r_i} \quad (3.10).$$

Метод Флетчера-Ривса сходится, если начальная точка достаточно близка к требуемому минимуму, тогда как метод Полака-Райбера может в редких случаях бесконечно циклиться. Однако последний часто сходится быстрее первого метода. К счастью, сходимость метода Полака-Райбера может быть гарантирована выбором $\beta = \max\{\beta, 0\}$. Это эквивалентно рестарту алготрима по условию $\beta \le 0$. Рестарт алгоритмической процедуры необходим, чтобы забыть последнее направление поиска и стартовать алгоритм заново в направлении скорейшего спуска.

Далее приведен алгоритм сопряженных градиентов для минимизации функций общего вида (неквадратичных).

1. Вычисляется антиградиент в произвольной точке x_0 $d_0 = r_0 = -f'(x_0)$.

2. Осуществляется спуск в вычисленном направлении пока функция уменьшается, иными словами, поиск a_i, который минимизирует $f(x_i + a_i d_i)$.

3. Переход в точку, найденную в предыдущем пункте $x_{i+1} = x_i + a_i d_i$.

4. Вычисление антиградиента в этой точке $r_{i+1} = -f'(x_{i+1})$.

5. Вычисления по формуле (3.9) или (3.10). Чтобы осуществить рестарт алгоритма, то есть забыть последнее направление поиска и стартовать алгоритм заново в направлении скорейшего спуска, для формулы Флетчера-Ривса

присваивается 0 через каждые n + 1 шагов, для формулы Полака-Райбера – $\beta_{i+1} = \max\{\beta_{i+1}, 0\}$.

6. Вычисление нового сопряженного направления $d_{i+1} = r_{i+1} + \beta_{i+1} d_i$.

7. Переход на пункт 2.

Из приведенного алгоритма следует, что на шаге 2 осуществляется одномерная минимизация функции. Для этого, в частности, можно воспользоваться методом золотого сечения или методом бисекций. Более быструю сходимость обеспечивает метод Ньютона-Рафсона, но для этого необходимо иметь возможность вычисления матрицы Гессе. В последнем случае, переменная, по которой осуществляется оптимизация, вычисляется на каждом шаге итерации по формуле $\alpha = \dfrac{f'^T(x)d}{d^T f''(x)d}$, где матрица Гессе:

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

Это дает основания некоторым авторам относить метод сопряженных градиентов к методам второго порядка, хотя суть метода вовсе не предполагает необходимым вычисление вторых производных.

По мнению некоторых авторитетных специалистов скорость сходимости алгоритма мало зависит от оптимизационной формулы, применяемой на шаге 2 приведенного выше алгоритма, поэтому можно рекомендовать, например, метод золотого сечения, который не требует вычисления производных.

Метод сопряженных градиентов является методом первого порядка, в то же время скорость его сходимости квадратична. Этим он выгодно отличается от обычных градиентных методов. Например, метод наискорейшего спуска и метод координатного спуска для квадратичной функции сходятся лишь в пределе, в то время как метод сопряженных градиентов оптимизирует квадратичную функцию за конечное число итераций. При оптимизации функций общего вида, метод сопряженных направлений сходится в 4-5 раз быстрее метода наискорейшего спуска. При этом, в отличие от методов второго порядка, не требуется трудоемких вычислений вторых частных производных.

3.2.3. Методы второго порядка

Напомним, что методом второго порядка называется метод, использующий значения минимизируемой функции, а также значения ее первых и вторых производных. Отсюда следует, что для использования методов второго порядка необходимым условием является дважды дифференцируемость $f(X)$.

К методам второго порядка относят:

- метод Ньютона;
- метод Ньютона-Рафсона;
- метод Марквардта и др. .

Центральное место среди методов второго порядка занимает метод Ньютона. Этот метод основан на квадратичной аппроксимации $f(X)$. Матрица вторых производных должна быть невырожденной.

Пусть X^k – k-е приближение к точке минимума. Покажем, как, зная его, можно получить следующее приближение X^{k+1}.

Разложим $f(X)$ в ряд Тейлора в окрестности точки X^k, оставляя члены разложения не выше второго:

$$f(X) = f(X^k) + (f'(X^k) * (X - X^k)) + \frac{1}{2}(f''(X^k) * (X - X^k), (X - X^k)) + o(\left\| X - X^k \right\|^2).$$

В этой формуле использовано традиционное обозначение скалярного произведения векторов a и b как (a,b), $f''(X^k)$ матрица вторых производных $f(X)$, вычисленных в приближении X^k. Разложение в ряд Тейлора используется для того, чтобы через него определить следующее приближение X^{k+1}, как точку минимума $f(X)$, удовлетворяющую соотношению $f'(X^{k+1})=0$ или $f'(X^k)+f''(X^k)*(X^{k+1}-X^k)=0$, откуда $X^{k+1}=X^k-(f''(X^k))^{-1}*f'(X^k)$. Такая формула носит название <u>формулы Ньютона</u>.

Если начальное приближение выбрано достаточно близко к точке минимума, а $f(X)$ – сильно выпуклая функция, метод Ньютона будет сходиться с квадратичной скоростью сходимости, то есть $\left\|X^{k+1}-X^k\right\| \le C*\left\|x^k-x^*\right\|^2$.

Алгоритм метода Ньютона состоит из следующих этапов

1 этап. Задать начальную точку $\bar{x}^0$, погрешности расчета $\varepsilon_1>0$, $\varepsilon_2>0$, M – предельное число итераций.

Найти градиент функции в произвольной точке $\nabla f(\bar{x})=\left(\frac{\partial f(\bar{x})}{\partial x_1},\ldots,\frac{\partial f(\bar{x})}{\partial x_n}\right)^T$ и матрицу Гессе $H(\bar{x})$.

2 этап. Принять $k=0$.

3 этап. Вычислить $\nabla f(\bar{x}^k)$.

4 этап. Проверить выполнение критерия окончания $\left\|\nabla f(\bar{x}^k)\right\|<\varepsilon_1$:

А) если критерий выполнен, расчет закончен $\bar{x}^*=\bar{x}^k$;

Б) если критерий не выполнен, то перейти к этапу 5.

5 этап. Проверить выполнение неравенства $k \ge M$:

А) если неравенство выполнено, то расчет окончен: $\bar{x}^*=\bar{x}^k$;

Б) если нет, то перейти к этапу 6.

6 этап. Вычислить матрицу $H(\bar{x}^k)$.

7 этап. Вычислить матрицу $H^{-1}(\bar{x}^k)$.

8 этап. Проверить выполнение условия $H^{-1}(\bar{x}^k)>0$:

А) если $H^{-1}(\bar{x}^k)>0$, то перейти к этапу 9;

Б) если $H^{-1}(\bar{x}^k)\le 0$, то перейти к этапу 10, приняв $\bar{d}^k=-\nabla f(\bar{x}^k)$.

9 этап. Определить $\bar{d}^k=-H^{-1}(\bar{x}^k)\nabla f(\bar{x}^k)$.

10 этап. Вычислить $\bar{x}^{k+1} = \bar{x}^k + t_k \bar{d}^k$, приняв $t_k = 1$, если $\bar{d}^k = -H^{-1}(\bar{x}^k)\nabla f(\bar{x}^k)$ или выбрав t_k из условия $f(\bar{x}^{k+1}) < f(\bar{x}^k)$, если $\bar{d}^k = -\nabla f(\bar{x}^k)$.

11 этап. Проверить выполнение условий:

$\left\|\bar{x}^{k+1} - \bar{x}^k\right\| \le \varepsilon_2$, $\left\|f(\bar{x}^{k+1}) - f(\bar{x}^k)\right\| \le \varepsilon_2$

А) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен и $\bar{x}^* = \bar{x}^{k+1}$;

Б) если хотя бы одно из условий не выполнено, то принять $k = k + 1$ и перейти к этапу 3.

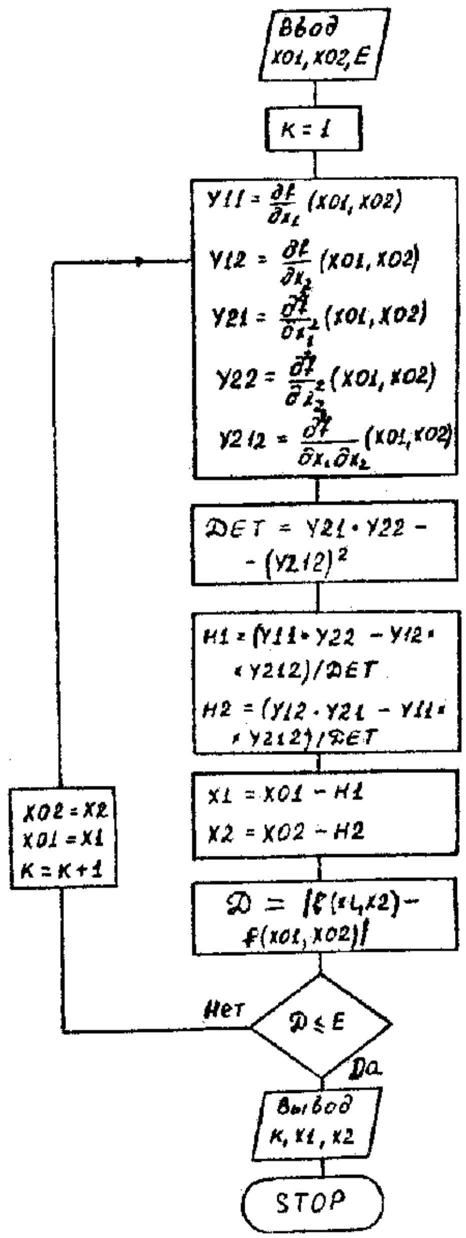


Рис. 27. Блок - схема метода Ньютона для функции двух переменных

Пример. Методом Ньютона найти точку минимума функции $f(X) = 9x_1^2 + x_2^2 - 18x_1 + 6x_2 + 18$ с точностью $\varepsilon = 10^{-3}$.

Решение.

$$f'(X) = \begin{pmatrix} 18x_1 - 18 \\ 2x_2 + 6 \end{pmatrix}$$

$$f''(X) = \begin{pmatrix} 18 & 0 \\ 0 & 2 \end{pmatrix}$$

По критерию Сильвестра матрица $f''(X)$ является положительно определенной.

Напомним, что согласно критерию Сильвестра матрица $(a_{i,j})$ $i = \overline{1,n}$ $j = \overline{1,n}$ положительно определена, если $a_{1,1} > 0$, а также определители всех миноров второго,...., n-го порядка, окаймляющих этот элемент положительны.

$$\begin{matrix} a_{1,1} & a_{1,2} & \ldots\ldots\ldots\ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots\ldots\ldots\ldots & a_{2,n} \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ a_{n,1} & a_{n,2} & \ldots\ldots\ldots\ldots & a_{n,n} \end{matrix}$$

$$f''(X)^{-1} = \begin{pmatrix} 1/18 & 0 \\ 0 & 0{,}5 \end{pmatrix}$$

Для этой задачи расчетная формула метода Ньютона будет иметь вид:

$$X^{k+1} = X^k - \begin{pmatrix} 1/18 & 0 \\ 0 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 18x_1^k - 18 \\ 2x_2^k + 6 \end{pmatrix}.$$

Введем критерии остановки:

$$| x_1^{k+1} - x_1^k | \le 10^{-3}$$

$$| x_2^{k+1} - x_2^k | \le 10^{-3}$$

$$\left\| f(X^{k+1}) - f(X^k) \right\| \le 10^{-3}$$

$|18x_1^{k+1}-18|\le 10^{-4}$

$|2x_2^{k+1}+6|\le 10^{-4}$.

В качестве начального приближения возьмем $X^0=\begin{pmatrix}0\\0\end{pmatrix}$.

Делаем первую итерацию по формуле Ньютона:

$$X^1=\begin{pmatrix}0\\0\end{pmatrix}-\begin{pmatrix}1/18 & 0\\0 & 1/2\end{pmatrix}\cdot\begin{pmatrix}18\cdot 0-18\\2\cdot 0+6\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}-\begin{pmatrix}-1\\3\end{pmatrix}=\begin{pmatrix}1\\-3\end{pmatrix}.$$

Проверяем критерий остановки: $|x_1^1-x_1^0|=|0-1|=1\ge 10^{-3}$, следовательно, необходимо продолжить вычисления.

Делаем вторую итерацию по методу Ньютона:

$$X^2=\begin{pmatrix}1\\-3\end{pmatrix}-\begin{pmatrix}1/18 & 0\\0 & 1/2\end{pmatrix}\cdot\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}-\begin{pmatrix}-1\\3\end{pmatrix}=\begin{pmatrix}1\\-3\end{pmatrix}.$$

Второе приближение совпало с первым, кроме того,

$$\begin{vmatrix}18\cdot 1-18\\2\cdot(-3)+6\end{vmatrix}=\begin{vmatrix}0\\0\end{vmatrix},$$

то есть выполнены все итерации остановки и решением поставленной задачи будет $x_1=1$, x_2 = -3.

Можно показать, что для любой квадратичной функции с положительно определенной матрицей вторых производных метод Ньютона дает точное решение независимо от начального приближения за одну итерацию. На рис. 27 представлена блок-схема метода Ньютона. Итак, метод Ньютона как метод второго порядка имеет быструю сходимость, если начальное приближение близко к точке минимума, однако, он может работать плохо или даже отказать вдали от точки минимума. Нелегкой задачей является процедура нахождения матрицы вторых производных. Обычно метод Ньютона используется на заключительном этапе, когда на начальном этапе работают другие методы. На практике находят широкое применение модификации метода Ньютона, направленные на то, чтобы, сохраняя основное достоинство метода – его быструю сходимость, уменьшить трудоемкость и ослабить требования на выбор начального приближения.

Вот некоторые виды модификаций:

- метод Ньютона с регулировкой шага: $X^{k+1} = X^k + \alpha_k (f''(X^k))^{-1} * f(X^k)$.

Выбор α_k производится либо из условия минимизации функции вдоль выбранного направления, либо путем дробления шага, обеспечивающего монотонное убывание $f(X)$, что обеспечивает сходимость при любой начальной точке;

- пересчет матрицы вторых производных производить не на каждом шаге.

4. Условная оптимизация функции многих переменных

Метод множителей Лагранжа.

Постановка задачи и ее особенности.

Задача математического программирования:

$$F(x_1, x_2, ..., x_n) \to \max(\min),$$
$$\begin{cases} g_i(x_1, x_2, ..., x_n) \le (=, \ge) b_i, \\ x = (x_1, x_2, ..., x_n) \ge 0, ...i = \overline{1, m}. \end{cases} \tag{4.1}$$

в которой либо целевая функция, либо ограничения, либо и то и другое нелинейны, называется задачей нелинейного программирования.

Рассмотрим экстремальную задачу с ограничениями в виде равенств.

Идея метода Лагранжа состоит в сведении задачи поиска условного экстремума целевой функции $f(x_1, x_2, \ldots, x_n)$ на множестве допустимых значений *D*, описываемом системой уравнений

$$D: \begin{cases} g_1(x_1, x_2, ..., x_n) = 0, \\ ... \\ g_m(x_1, x_2, ..., x_n) = 0. \end{cases} \tag{4.2}$$

к задаче безусловной оптимизации функции.

Предполагается, что функции $f(x_1, x_2, \ldots, x_n)$ и $g_i(x)$ непрерывные вместе со своими частными производными.

Эта задача является классической задачей на условный экстремум. Чтобы ее решить, используют функцию Лагранжа:

$$L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = f(x_1,\ldots,x_n) + \lambda_1 \cdot g_1(x_1,\ldots,x_n) + \ldots + \lambda_m \cdot g_m(x_1,\ldots,x_n) \qquad (4.3),$$

где вектор $\lambda \in R^m$ – вектор дополнительных переменных, называемых *множителями Лагранжа.*

Функцию $L(x,\lambda)$, где $x \in R^n$, $\lambda \in R^m$, называют *функцией Лагранжа.*

В случае дифференцируемости функций F и g_i справедлива теорема, определяющая необходимое условие существования точки условного экстремума в данной задаче. Поскольку она непосредственно относится к предмету математического анализа, приведем ее без доказательства.

Теорема 1. *Если x* является точкой условного экстремума функции (4.3) при ограничениях (2) и ранг матрицы первых частных производных функций*

$$J = \left(\frac{\partial g_i(x^*)}{\partial x_j} \right)$$

равен m, то существуют такие $\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*$, *не равные одновременно нулю, при которых*

$$\nabla L(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \cdot \nabla g_i(x^*) = 0 \qquad (4.4).$$

Из теоремы 1 вытекает метод поиска условного экстремума, получивший название метода множителей Лагранжа, или просто *метода Лагранжа.*

Схема применения этого метода следующая.

1. Вводится вектор из m дополнительных новых переменных $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in R^m$ – *множителей Лагранжа.*

2. Определяется *функция Лагранжа*

$$L(x,\lambda) = f(x) + \lambda \cdot g(x)$$

$$L(x_1,...,x_n,\lambda_1,...,\lambda_m) = f(x_1,...,x_n) + \lambda_1 \cdot g_1(x_1,...,x_n) + ... + \lambda_m \cdot g_m(x_1,...,x_n).$$

3. Находятся частные производные функции Лагранжа по x_j, $j = \overline{1,n}$ и λ_i, $i = \overline{1,m}$ и приравниваются к нулю:

$$\begin{cases} \dfrac{\partial f}{\partial x_j}(x_1,...,x_n) - \sum_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_j}(x_1,...,x_n) = 0, \quad j = \overline{1,n}, \\ g_i(x_1,...,x_n) = 0, \quad i = \overline{1,m}. \end{cases} \qquad (4.5)$$

4. Решается система (4.5) относительно $n+m$ неизвестных $x_1,..., x_n, \lambda_1,..., \lambda_m$, т.е. находятся стационарные точки (x^*, λ^*) функции Лагранжа $L(x,\lambda)$.

Отметим, что условия (4.5) дают лишь необходимые условия экстремума. Если же выполнены достаточные условия, то x^* будет точкой локального экстремума.

Пусть $H = \left(\dfrac{\partial^2 L(x,\lambda)}{\partial x_i \partial x_j} \right)$ – матрица Гессе, составленная из вторых частных производных функции Лагранжа по переменным x_j, $J = \left(\dfrac{\partial g_i}{\partial x_j} \right)$ – $m \times n$–матрица, составленная из частных производных функций g_i по переменным x_j, называемая матрицей Якоби. Определим $(m+n)\times(m+n)$-матрицу

$K = \begin{pmatrix} O & J \\ J^{\mathrm{T}} & H \end{pmatrix}$, называемую *окаймленной матрицей Гессе*.

Следующая теорема дает достаточные условия условного экстремума.

Теорема 2. *Пусть (x^*,λ^*) – стационарная точка функции Лагранжа $L(x,\lambda)$, K – окаймленная матрица Гессе в этой точке. Тогда x^* является точкой максимума (минимума), если последние $(n–m)$ главных угловых миноров матрицы K имеют чередующиеся знаки, причем знак первого из них совпадает со знаком* $(-1)^{m+1}$ (*соответственно* $(–1)^m$).

Условия теоремы 2 не являются необходимыми, т.е. стационарная точка, не удовлетворяющая этим условиям, может быть экстремальной.

Рассмотрим теперь важную интерпретацию множителей Лагранжа. Решение системы (4.5) дает, кроме вектора локального экстремума x^*, еще и вектор множителей Лагранжа λ^*, который несет ценную информацию. Оптимальное значение целевой функции $f^* = f(x^*)$ зависит от значений констант ограничений b_i. Можно доказать, что $\lambda_i^* = \dfrac{\partial f^*}{\partial b_i}$, $i = \overline{1,m}$, (4.5*)

т.е. множители Лагранжа, соответствующие решению задачи, измеряют *чувствительность* оптимального значения целевой функции *к изменениям констант ограничений*. В экономических задачах распределения ресурсов целевая функция имеет размерность стоимости, а множители Лагранжа – размерность цены (т.е. стоимости единицы ресурса). По этой причине множители Лагранжа часто называют *теневыми ценами* соответствующих ресурсов.

Отметим, что основное практическое значение метода Лагранжа заключается в том, что он *позволяет перейти от условной оптимизации к безусловной* и, соответственно, расширить арсенал доступных средств решения проблемы. Однако нетрудно заметить, что задача решения системы уравнений (4.5), к которой сводится данный метод, в общем случае не проще исходной проблемы поиска экстремума задачи (4.1). Методы, подразумевающие такое решение, называются *непрямыми*. Они могут быть применены для весьма узкого класса задач, для которых удается получить линейную или сводящуюся к линейной систему уравнений. *Прямые* методы основаны на итеративных процессах вычисления и сравнения значений оптимизируемых функций.

Метод множителей Лагранжа можно использовать при построении критериев оптимальности для задач с ограничениями в виде неравенств. Условия Куна-Таккера.

Распространим метод множителей Лагранжа для решения задач нелинейного программирования с ограничениями в форме неравенств:

найти $\max\limits_x f(x)$ при условии $g(x) \le b$. (4.6)

По аналогии с предыдущим пунктом определим для задачи (4.6) функцию

Лагранжа: $L(x,\lambda) = f(x) - \lambda \cdot g(x)$ (4.7)

или в развернутом виде:

$$L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = f(x_1,\ldots,x_n) - \lambda_1 \cdot g_1(x_1,\ldots,x_n) - \ldots - \lambda_m \cdot g_m(x_1,\ldots,x_n).$$

Определение. *Пара векторов* $(\bar{x},\bar{\lambda})$ *, максимизирующая функцию* $L(x,\lambda)$ *по совокупности всех неотрицательных переменных* $(\bar{x})$ *и минимизирующая ее по совокупности всех неотрицательных множителей Лагранжа* $(\bar{\lambda})$ *, называется* **седловой точкой** *функции* $L(x, \lambda)$ *в некоторой области* X x Λ, *если для любых* $x \in X$ *и* $\lambda \in \Lambda$

$$L(x,\bar{\lambda}) \le L(\bar{x},\bar{\lambda}) \le L(\bar{x},\lambda). \qquad (4.8)$$

Неравенства (4.8) также называют *неравенствами седловой точки.*

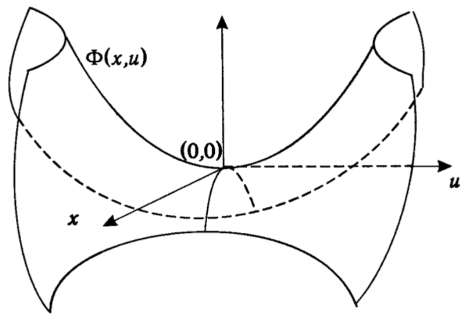


Рис. 28. Иллюстрация седловой точки

В качестве примера седловой точки может быть приведена точка (0, 0) для функции $L(x, \lambda) = -x^2 + \lambda^2$, определенной на множестве R x R. В самом деле, $\Phi(0,0)=0$, $\Phi(x,0)=-x^2$, $\Phi(0, \lambda) = \lambda^2$, а для любых $x \in R$ и $\lambda \in R$ выполняются неравенства $-x^2 \le 0$ и $0 \le \lambda^2$.

На рис. 28 изображен график функции $L(x, \lambda)$ (гиперболический параболоид), и, как видно, в окрестности точки (0,0) он действительно по форме напоминает седло, чем и объясняется происхождение соответствующего термина.

Рассмотрим экстремальную задачу с ограничением в виде неравенств.

Ограничения-неравенства можно преобразовать к виду равенств путем введения дополнительных неотрицательных переменных s_i^2 , $i = \overline{1,m}$.

Пусть $s = \left(s_1, s_2, \ldots, s_m\right)^{\text{T}}$, $s^2 = \left(s_1^2, s_2^2, \ldots, s_m^2\right)^{\text{T}}$.

Тогда ограничения запишутся в виде $g(x)+s^2 = b$, и функция Лагранжа будет иметь вид $L_1(x, s, \lambda) = f(x) + \lambda(b - g(x) - s^2)$. (4.9)

Необходимым условием оптимальности в задаче (4.6) является неотрицательность λ^*. Действительно, согласно (4.5*) $\lambda^* = \dfrac{\partial f^*}{\partial b}$, но при увеличении b допустимое множество расширяется, и, следовательно, максимум целевой функции не может уменьшиться. Поэтому $\lambda^* \geq 0$. Аналогично в задаче минимизации $\lambda^* \leq 0$. Если же ограничения заданы в виде равенства $g(x) = b$, то на знак λ никаких условий не накладывается.

Запишем уже известные нам необходимые условия экстремума, приравняв к нулю частные производные по x, s и λ функции Лагранжа (4.9).

$$\frac{\partial L_1}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0, \tag{4.10}$$

$$\frac{\partial L_1}{\partial s_i} = -2\lambda_i s_i = 0, \quad i = \overline{1, m}, \tag{4.11}$$

$$\frac{\partial L_1}{\partial \lambda} = b - g(x) - \mathrm{s}^2 = 0. \tag{4.12}$$

После очевидных преобразований, получим:

$$\lambda_i (b_i - g_i(x)) = 0, \quad i = \overline{1, m}, \tag{4.13}$$

1. Если $\lambda_i > 0$, то $g_i(x) = b_i$.
2. Если $g_i(x) < b_i$, то $\lambda_i = 0$.

Заметим, что совокупность уравнений (4.13) при условиях $\lambda \geq 0$ и $g(x) \leq b$ равносильна одному уравнению:

$$\sum_{i=1}^{m} \lambda_i (b_i - g_i(x)) = 0 \text{ или } \lambda(b - g(x)) = 0, \tag{4.14}$$

в последней сумме все слагаемые неотрицательны и равенство ее нулю равносильно равенству каждого слагаемого.

Условия (4.13) и (4.14) называются **условиями дополняющей нежесткости**.

Таким образом, совокупность необходимых условий экстремума в задаче (4.6) может быть записана в виде системы:

$$\begin{cases} \dfrac{\partial f}{\partial x} - \lambda \dfrac{\partial g}{\partial x} = 0, \\ \lambda \geq 0, \\ g(x) \leq b, \\ \lambda(b - g(x)) = 0. \end{cases} \tag{4.15}$$

Эти условия называются **условиями Куна-Таккера**.

Рассмотрим теперь задачу (4.6) при дополнительном условии неотрицательности переменных $x \geq 0$.

Заметим, что эта задача легко сводится к уже рассмотренной задаче (4.6), если переписать условие $x \geq 0$ в виде $h(x) \leq 0$, где $h = (h_1,h_2,...,h_n)^T$ и $h_j(x) = -x_j$, $j = \overline{1,n}$, т. е. мы имеем задачу с $m+n$ ограничениями: $g(x) \leq b$, $h(x) \leq 0$. Введя аналогичную s переменную $t = (t_1, t_2, ..., t_n)^T$, запишем функцию Лагранжа для этой задачи:

$$L_2(x, s, t, \lambda, \mu) = f(x) - \lambda\,(b - g(x) - s^2) + \mu\,(x - t^2).$$

Запишем теперь условия Куна-Таккера:

$$\begin{cases} \dfrac{\partial L_2}{\partial x} = \dfrac{\partial f}{\partial x} - \lambda \dfrac{\partial g}{\partial x} + \mu = 0 \\ \lambda \geq 0 \\ g(x) \leq b \\ \lambda(b - g(x)) = 0 \\ \mu \geq 0 \\ x \geq 0 \\ \mu x = 0. \end{cases}$$

Исключив из этих уравнений μ, получим систему:

$$
\begin{cases}
\lambda \geq 0 \\
g(x) \leq b \\
\lambda(b - g(x)) = 0 \\
\dfrac{\partial f}{\partial x} - \lambda \dfrac{\partial g}{\partial x} \leq 0 \\
x \geq 0 \\
\left(\dfrac{\partial f}{\partial x} - \lambda \dfrac{\partial g}{\partial x} \right) x = 0.
\end{cases}
\tag{4.16}
$$

Эти условия называются условиями Куна-Таккера для расширенной задачи (4.6). В развернутом виде эти условия представляют собой систему из $2m+2n+2$ соотношений:

$$
\begin{cases}
\lambda_i \geq 0, \quad i = \overline{1, m}, \\
\dfrac{\partial L}{\partial \lambda_i} = b_i - g_i(x) \geq 0, \quad i = \overline{1, m}, \\
\sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial L}{\partial \lambda_i} = \sum\limits_{i=1}^{m} \lambda_i (b_i - g_i(x) = 0, \\
\dfrac{\partial L}{\partial x_j} = \dfrac{\partial f}{\partial x_j} - \sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_j} \leq 0, \quad j = \overline{1, n}, \\
x_i \geq 0, \quad j = \overline{1, n}, \\
\sum\limits_{j=1}^{n} \dfrac{\partial L}{\partial x_j} x_j = \sum\limits_{j=1}^{n} \left(\dfrac{\partial f}{\partial x_j} - \sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_j} \right) x_j = 0,
\end{cases}
\tag{4.17}
$$

где L – функция Лагранжа.

В случае задачи минимизации знаки неравенств в первом и четвертом из этих условий заменяются противоположными.

Решение системы, порожденной условиями Куна-Таккера, сопряжено со значительными трудностями. В большинстве случаев описанный метод не подходит для численных расчетов. Тем не менее, он имеет важное теоретическое значение для построения алгоритмов решения задач нелинейного программирования. На основе полученных результатов можно построить достаточные условия глобального экстремума задачи (4.1).

Задачи оптимизации с ограничениями в форме равенств и неравенств. Штрафные и барьерные функции.

Суть используемых здесь методов заключается в замене исходной задачи эквивалентной задачей безусловной оптимизации или последовательностью задач безусловной оптимизации.

Рассматриваются два альтернативных подхода:

- первый называется методом штрафных функций и заключается в следующем: к целевой функции исходной задачи добавляется функция, интерпретируемая как штраф за нарушение каждого из ограничений. Метод генерирует последовательность недопустимых точек, которая сходится к оптимальному решению исходной задачи.
- второй подход называется методом барьеров. В этом методе к целевой функции исходной задачи добавляется барьерный член, который не позволяет генерируемым точкам выходить за пределы допустимой области, и эта последовательность точек сходится к оптимальному решению исходной задачи. Этот метод (барьер) может использоваться только в задачах с ограничениями в виде неравенств.

Метод штрафных функций.

В этом методе с помощью функций, задающих ограничения, строится так называемый штраф, который добавляется к целевой функции исходной задачи так, что нарушение какого-либо из ограничений становится невыгодным с точки зрения полученной задачи безусловной оптимизации.

Обычно подходящая штрафная функция должна определять положительный штраф в недопустимых точках и не штрафовать допустимые точки.

В этом методе с помощью функций, задающих ограничения, строится так называемый штраф, который добавляется к целевой функции исходной задачи так, что нарушение какого-либо из ограничений становится невыгодным с точки зрения полученной задачи безусловной оптимизации.

Обычно подходящая штрафная функция должна определять положительный штраф в недопустимых точках и не штрафовать допустимые точки.

Если ограничения имеют форму: $j = \overline{1, n}$

$g_i(x) \le 0, \quad i = \overline{1, m}$

$h_j(x) = 0, \quad j = \overline{1, l}$

$x \in R^n$,

то целесообразно использовать штрафную функцию следующего вида:

$$\alpha(x) = \sum_{i=1}^{m} \varphi(g_i(x)) + \sum_{j=1}^{l} \psi(h_j(x)) \qquad (4.18)$$

$\varphi(y) = 0$, если $y \le 0$ и $\varphi(y) > 0$, если $y > 0$,

где φ и ψ – непрерывные функции , удовлетворяющие условиям:

$\psi(y) = 0$, если $y = 0$ и $\psi(y) > 0$, если $y \ne 0$.

Типичными являются следующие формы функций φ и ψ:

$\varphi(y) = [\max\{0, y\}]^p$ и $\psi(y) = |y|^d$, где p, d – целые положительные числа.

Часто выбирают $p = d$.

Таким образом, штрафная функция имеет вид:

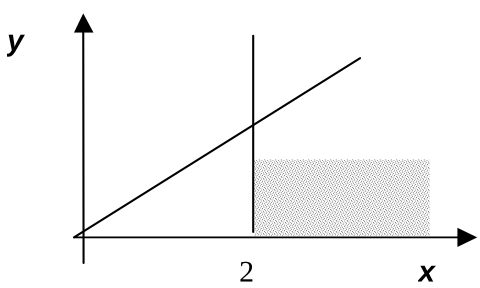
$$\alpha(x) = \sum_{i=1}^{m} [\max\{0, g_i(x)\}]^p + \sum_{j=1}^{l} |h_j(x)|^p .$$

Функцию $f(x) + \mu \cdot \alpha(x)$ называют <u>вспомогательной</u>.

<u>Пример.</u>

Рассмотрим задачу: $x \to \min$.

$-x + 2 \le 0$

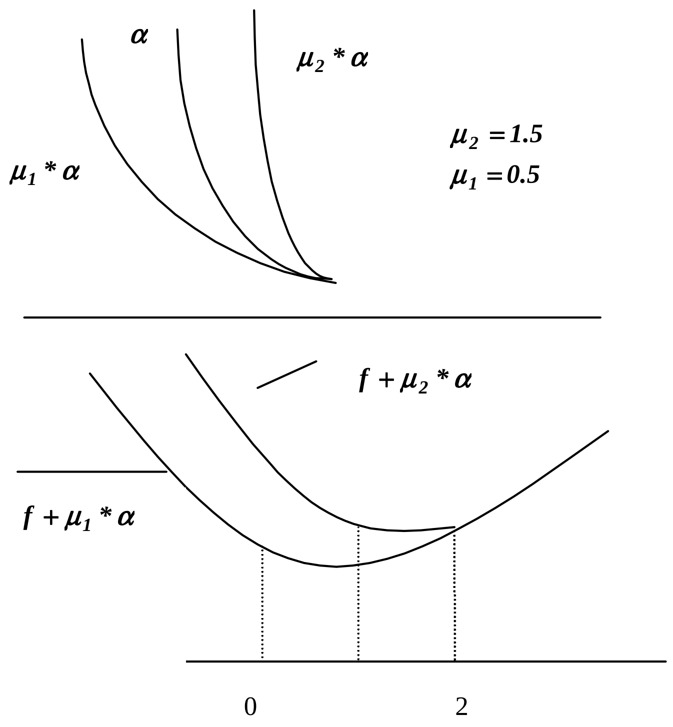


Решение $x = 2, \quad y = 2$.

Положим $\alpha(x) = [\max\{0, g(x)\}]^2$, тогда $\alpha(x) = \begin{cases} 0, & \text{если } x \geq 2 \\ (-x+2)^2, & \text{если } x < 2 \end{cases}$.

Минимум $f(x) + \mu \cdot \alpha(x)$ достигается в точке $2 - \dfrac{1}{2 \cdot \mu}$. При $\mu \to \infty$ последовательность таких точек стремится к точке $x = 2$, являющейся точкой минимума исходной задачи.

График:



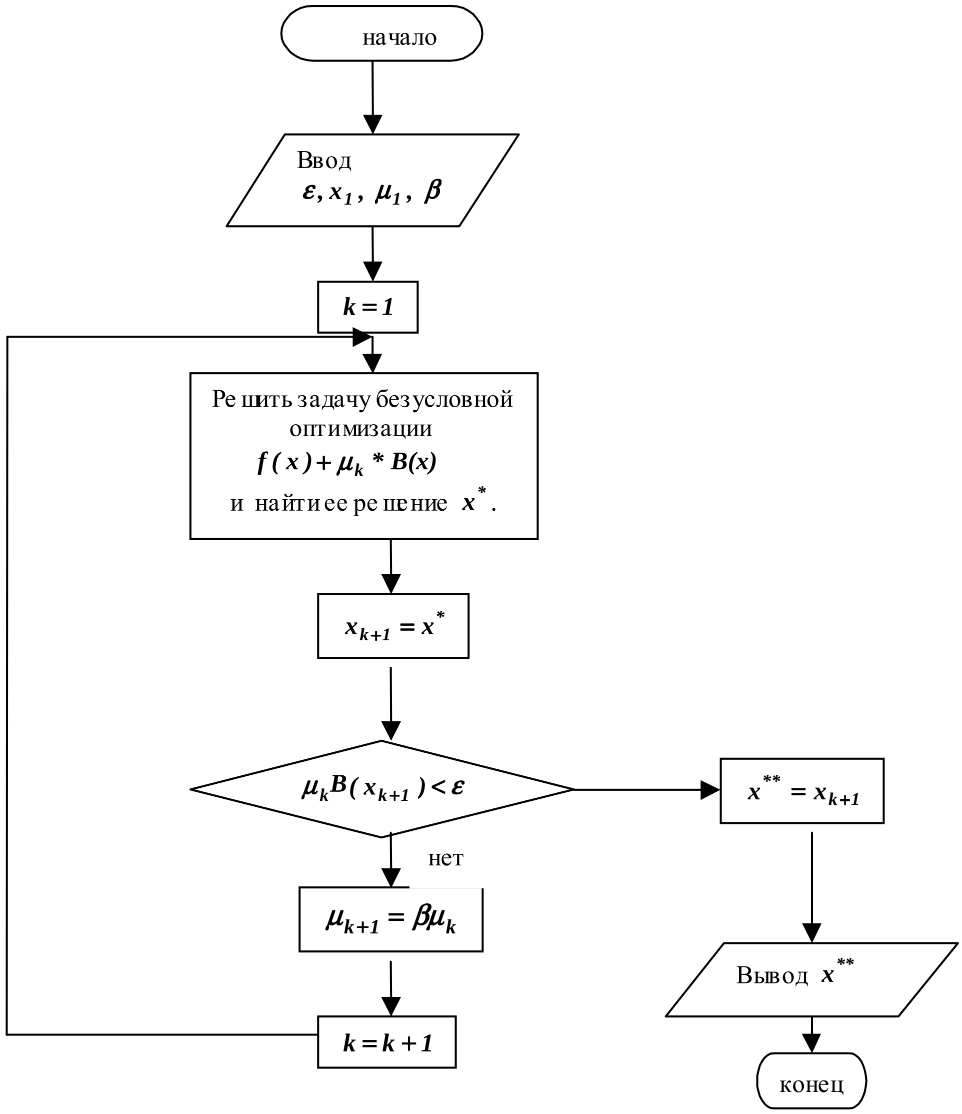
Алгоритм метода штрафных функций

$f(x) \to \min$

$g_i(x) \leq 0, \quad i = \overline{1, m}$

$h_j(x) = 0, \quad j = \overline{1, l}$

$x \in R^n$



ε – точность вычисления;

β – некое число, $\beta > 1$;

x_1 – начальная точка;

μ_1 – штрафной параметр, $\mu_1 > 0$;

k – параметр цикла;

x^* – оптимальное решение задачи безусловной оптимизации (на каждой итерации свое);

x^{**} – оптимальное решение исходной задачи.

Метод барьерных функций или метод внутренних штрафных функций

В этом методе к целевой функции исходной задачи добавляется барьерная функция, которая не позволяет генерируемым точкам выходить за пределы допустимой области. Эта последовательность точек сходится к оптимальному решению исходной задачи.

Барьерные функции используются, также, как и штрафные, для преобразования задачи с ограничениями в задачу безусловной оптимизации или в последовательность таких задач. Барьерные функции как бы препятствуют выходу из допустимой области. Ограничения должны быть только в форме неравенств.

Исходная задача

$f(x) \to \min$

$g_i(x) \le 0$, $i = 1,...,m$

$x \in R^n$

преобразуется в задачу безусловной оптимизации:

$$f_1(x) = f(x) + \mu \cdot B(x) \to \min, \quad x \in R^n, \quad \mu \ge 0,$$

где $B(x)$ – барьерная функция, которая в общем виде записывается как:

$$B(x) = \sum_{i=1}^{m} \varphi_1(g_i(x)),$$

где φ_1 – функция одной переменной, удовлетворяющая условиям:

$\varphi_1(y) \ge 0$, если $y < 0$ и $\lim\limits_{y \to 0} \varphi_1(y) = \infty$.

$B(x)$ конструируется таким образом, чтобы она была неотрицательна и непрерывна в области $\{x : g_i(x) < 0\}$ и стремилась к бесконечности при приближении из внутренней точки к границе области.

Типичная барьерная функция имеет вид: $B(x) = \sum\limits_{i=1}^{m} \dfrac{(-1)}{g_i(x)}$ («минус», так как задача на *min* и $g_i(x) < 0$).

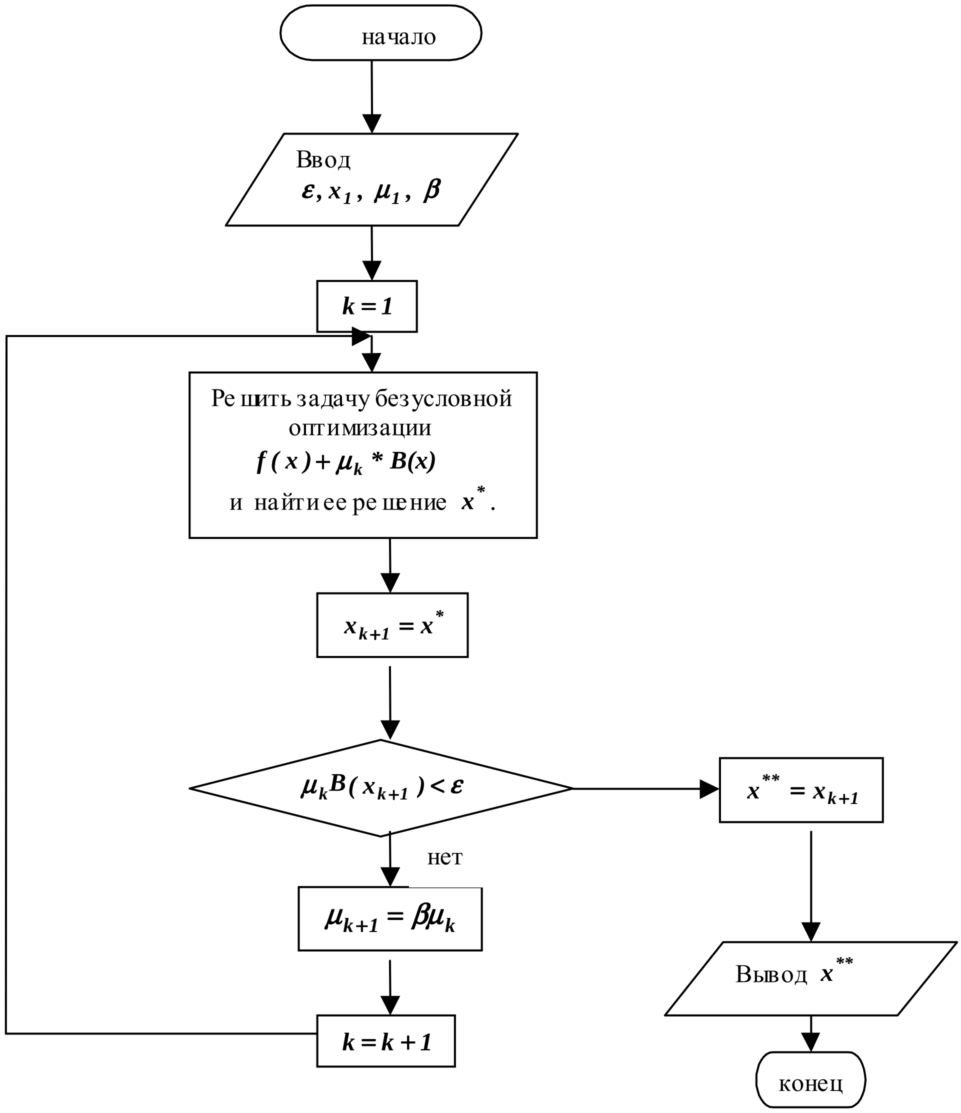
Функцию $f(x) + \mu \cdot B(x)$ называют <u>вспомогательной конструкцией</u>.

Алгоритм метода барьеров

$f(x) \to \min$

$g_i(x) \le 0, \quad i = \overline{1, m}$

$x \in R^n$



ε – точность вычисления; k – параметр цикла;

β – некое число, $\beta \in (0,1)$;

x_1 – начальная точка;

μ_1 – штрафной параметр, $\mu_1 > 0$;

x^* – оптимальное решение задачи безусловной оптимизации (на каждой итерации свое); x^{**} – оптимальное решение исходной задачи.

5. **Примеры задач для практических занятий**

Задача 1.

Показать, что функция $f(x) = x^4 - 10x^3 + 36x^2 + 5x$ унимодальна на отрезке [3;5].

Решение. Найдем последовательно первую и вторую производные функции

$f'(x) = 4x^3 - 30x^2 + 72x + 5$,

$f''(x) = 12x^2 - 60x + 72$.

Решим уравнение $f''(x) = 0$. Корни полученного квадратного трехчлена x_1=2 и x_2=3. Следовательно, $f''(x) \geq 0$, если $x \geq 3$ и, в частности, при $x \in$ [3; 5]. Используя дифференциальный критерий унимодальности, получаем, что $f(x)$ выпуклая на отрезке [3; 5] и, значит, унимодальная.

Задача 2.

Выяснить, является ли функция $f(x) = x^2 - 3x + x\ln x$ на отрезке [1;2] унимодальной.

Решение. Найдем первую и вторую производные функции:

$f'(x) = 2x + \ln x + 4$,

$f''(x) = 2 + \frac{1}{x}$.

Решим уравнение $f''(x) = 0$. Полученное уравнение имеет единственный корень x=-0,5. Следовательно, $f''(x) \geq 0$, если $x \geq -0{,}5$ и, в частности, при $x \in$ [1; 2]. Используя дифференциальный критерий унимодальности, получаем, что $f(x)$ выпуклая на отрезке [1; 2] и, значит, унимодальная.

Задача 3.

Определить направление выпуклости и точки перегиба кривой $f(x) = 3x^5 - 5x^4 + 4$.

Решение. Ищем точки x, в которых $f''(x) = 0$ или не существует, а кривая непрерывна и которые лежат внутри области расположения кривой:

$f'(x) = 15x^4 - 20x^3$,

$f''(x) = 60x^3 - 60x^2 = 60x^2(x-1)$.

$f''(x) = 0$ в точках $x = 0$, $x = 1$. Эти точки являются искомыми, так как область расположения и область непрерывности данной кривой есть вся ось абсцисс. Других точек x, которые могли бы быть абсциссами точек перегиба, нет, так как $f''(x)$ существует всюду.

Исследуем найденные точки, определяя знак второй производной слева и справа от каждой из них. Запишем это исследование в таблице:

x	-1	0	1/2	1	10
y''	-	*0*	-	*0*	+
y	*Выпукла*	*Нет перегиба*	*Выпукла*	*Перегиб*	*Вогнута*

Из таблицы следует, что x=1 есть абсцисса точки перегиба кривой: $y(1)$ = 2. Поскольку эта кривая непрерывная, то во всем интервале *(-∞, 1)* она выпукла, а во всем интервале (1, +∞) – вогнута.

Задача 4.

Найти наибольшее и наименьшее значение функции $f(x) = x^3 - 3x^2 - 9x + 35$ на отрезке [-4; 4].

Решение. Найдем критические точки функции *u(x)*, лежащие внутри отрезка [-4; 4], и вычислим ее значения в этих точках: $f'(x) = 3x^2 - 6x - 9 = 0$ в точках x = -1, x= 3. Эти точки лежат внутри рассматриваемого отрезка и являются критическими.

Вычислим значения функции на концах отрезка и в критических точках:

$f(-1) = 40$, $f(3) = 8$, $f(-4) = -41$, $f(4) = 15$.

Сравнивая все вычисленные значения функции во внутренних критических точках и на концах отрезка, заключаем: наибольшее значение функции *u* на отрезке [-4; 4] равно 40 и достигается ею во внутренней критической точке x = -1, а ее наименьшее значение равно –41 и достигается на левой границе отрезка x=-4.

Задача 5. Классический метод минимизации.

Решить задачу $f(x) = x^3 - 3x + 1 \to \min$, $x \in [-2; 2]$.

Шаг 1. Находим корни уравнения $f'(x) = 3x^2 - 3 = 0$ из интервала $(-2; 2)$:

$x_1 = -1$, $x_2 = 1$. Полагаем $x_0 = -2$, $x_3 = 2$.

Шаг 2. Вычисляем значения $f(x)$ в точках x_i, $i = 0, .., 3$: $f(x_0) = -17$, $f(x_1) = 3$, $f(x_2) = -1$, $f(x_3) = 1$.

Шаг 3. Находим $f^* = \min(-17, 3, -1, 1) = -17 = f(x_0)$.

Поэтому $x^* = x_0 = -2$, $f^* = -17$.

Задача 6.

Найти и идентифицировать оптимумы функции

$$f(x) = 5x^6 - 36x^5 + \frac{165}{2}x^4 - 60x^3 + 36.$$

Решение. Сначала найдем первую производную функции:

$$f'(x) = 30x^5 - 180x^4 + 330x^3 - 180x^2 = 30x^2(x-1)(x-2)(x-3).$$

Найдем стационарные точки. Для этого решим уравнение $f'(x) = 0$:

Следовательно, стационарные точки: $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, $x_4 = 3$.

Найдем вторую производную $f''(x) = 150x^4 - 720x^3 + 990x^2 - 360x$.

Для идентификации точек оптимума, вычислим значение второй производной в стационарных точках.

x	$f(x)$	$f''(x)$
0	36	0
1	27,5	60
2	44	-120
3	5,5	540

Значит, $x=1$ $x=3$ – точки локальных минимумов, $x=2$ – точка локального максимума.

$$\left.\frac{d^3 f}{dx^3}\right|_{x=0} = (600x^3 - 2160x^2 + 1980x - 360)\big|_{x=0} = -360.$$

Чтобы идентифицировать точку $x=\boldsymbol{0}$, найдем и вычислим третью производную.

Так как $\dfrac{d^3 f}{dx^3} \neq 0$ и n=3 – нечетное, то x*=0 – точка перегиба.

Задача 7.

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

Решение. Будем разбивать первоначальный и новый, удовлетворяющий нас, интервал на 4 части, при этом новый шаг рассчитываем по формуле:

$$h = \frac{|a_i - b_i|}{n},$$

где n – количество частей деления интервала, a_i, b_i – концы интервала, в котором содержится максимальное значение функции, погрешность $\varepsilon = |a_i - b_i|$, где i – номер итерации.

Номер п.	***Шаг***	***Концы новых интервалов***	***Значение функции***	***Погрешность***	***Примечание***
1	**2**	**3**	**4**	**5**	**6**
1.	0,1250	1,5000	-89,4375	0,2500	Точность не достигнута
1,6250	-91,5427				
1,7500	-92,1211				
1,8750	-90,9998				
2,0000	-88,0000				
2.	0,0625	1,6250	-91,5427	0,1250	Точность не достигнута
1,6875	-92,0334				
1,7500	-92,1211				
1,8125	-91,7839				

1,8750	-90,9998				
3.	0,0313	1,6875	-92,0334	0,0625	Точность не достигнута
1,7188	-92,1290				
1,7500	-92,1211				
1,7813	-92,0070				
1,8125	-91,7839				
4.	0,0156	1,6875	-92,0334	0,0313	Точность достигнута
1,7031	-92,0940				
1,7188	-92,1290				
1,7344	-92,1381				
1,7500	-92,1211				

Ответ: $x^* \approx 1{,}7344$, $f^* \approx -92{,}1381$.

Задача 8.

Методом деления отрезка пополам найти $f(x) = x^4 + e^{-x} \to \min$, $x \in [0;1]$, ε=0,1. Выберем δ=0,02.

Итерация 1.

Шаг 1. $x_1 = 0{,}49$, $x_2 = 0{,}51$. $f(x_1) = 0{,}670$, $f(x_2) = 0{,}688$.

Шаг 2. $f(x_1) > f(x_2)$, поэтому полагаем $a = x_1 = 0{,}49$.

Шаг 3. $(b\text{–}a)/2 = 0{,}255 > 0{,}1$, т.е. переходим к следующей итерации. Результаты вычислений на остальных итерациях записаны в таблице:

Номер итерации	a	b	$\frac{u-a}{2}$	x_1	x_2	$f(x_1)$	$f(x_2)$	Сравнение $f(x_1)$ и $f(x_2)$
2	0,49	1	0,26	0,735	0,755	0,771	0,792	$f(x_1) < f(x_2)$
3	0,49	0,755	0,13	0,613	0,633	0,683	0,691	$f(x_1) < f(x_2)$
4	0.49	0,633	0,07	0,07 < 0,1 – точность достигнута				

Таким образом, $x^* \approx \dfrac{0{,}49 + 0{,}633}{2} \approx 0{,}56$, $f^* \approx f(0{,}56) \approx 0{,}67$

Задача 9.

Методом золотого сечения найти минимум функции

$f(x) = x^4 + e^{-x} \to \min$, $x \in [0; 1]$, $\varepsilon = 0{,}1$.

Итерация 1.

Шаг 1. Находим: $x_1 = 0{,}382$, $x_2 = 0{,}618$, $f(x_1) = 0{,}704$, $f(x_2) = 0{,}685$, $\varepsilon_n = 0{,}5$.

Шаг 2. $\varepsilon_n = 0{,}5 > \varepsilon = 0{,}1$, поэтому переходим к шагу 3.

Шаг 3. $f(x_1) > f(x_2)$, поэтому полагаем $a = 0{,}382$, $x_1 = 0{,}618$, $f(x_1) = 0{,}685$, $x_2 = 0{,}764$, $\varepsilon_n = 0{,}309$ и вычисляем $f(x_2) = 0{,}807$. Переходим к следующей итерации, начиная с шага 2.

Результаты вычислений на остальных итерациях представлены в таблице (стрелки указывают значения, переходящие на данную итерацию с предыдущей).

Таблица

Номер итерации	a	b	ε_n	x_1	x_2	$f(x_1)$	$f(x_2)$	Сравнение $f(x_1)$ и $f(x_2)$
2	0,382	1,000	0,309	0,618	0,764	0,685	0,807	$f(x_1) < f(x_2)$
3	0,382	0,764	0,191	0,528	0,618	0,668	0,685	$f(x_1) < f(x_2)$
4	0,382	0,618	0,118	0,472	0,528	0,673	0,668	$f(x_1) > f(x_2)$
5	0,472	0,618	0,073	0,073 < 0,1 – точность достигнута				

Таким образом, $x^* \approx \dfrac{0{,}472 + 0{,}618}{2} \approx 0{,}55$, $f^* \approx f(0{,}55) = 0{,}67$.

Замечание. Число итераций, необходимое для достижения заданной точности ε, можно найти из условия $\varepsilon_n \le \varepsilon$ с учетом соотношения: $n \ge \ln\left(\dfrac{2\varepsilon}{b-a}\right) \Big/ \ln \approx -2{,}1\ln\left(\dfrac{2\varepsilon}{b-a}\right)$.

Так как $\boldsymbol{N}$ вычислений $f(x)$ позволяют выполнить $\boldsymbol{N} - 1$ итераций метода золотого сечения, то достигнутая в результате этих вычислений точность определения x^* составляет

$$\varepsilon(N)=\varepsilon_{N-1}=\frac{1}{2}\left(\frac{\sqrt{5}-1}{2}\right)^{N-1}(b-a).$$

Задача 10.

Методом золотого сечения найти минимальное значение f^* и точку минимума x^* функции $f(x)=x^4+8x^3-6x^2-72x$ на отрезке [1,5; 2]. Точку x^* найти с погрешностью ε=0,05.

Решение. Вычисления проведем по формулам, представив результаты в таблице:

N	εn	*An*	*Bn*	*X1(n)*	*X2(n)*	*F(x1(n))*	*F(x2(n))*	**Примечание**
1	0,309	1,500	2,000	1,691	1,809	-92,049	-91,814	$f(x_1^{(1)})<f(x_2^{(1)}), b_2=x_2^{(1)}$
2	0,191	1,500	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	-92,084	$f(x_1^{(4)})<f(x_2^{(4)}), b_5=x_2^{(4)}$
5	0,045				1,736		-92,138	$\varepsilon_n<\varepsilon$, *точность_достигнута*

Первоначальные значения x_1 *и* x_2 находим по формулам

$$x_1=a+\frac{3-\sqrt{5}}{2}(b-a),\; x_2=a+\frac{\sqrt{5}-1}{2}(b-a),$$

а значения точности ε по формуле: $|x^*-x_n|\le\varepsilon_n=\left(\frac{\sqrt{5}-1}{2}\right)(b-a)$.

Из таблицы получаем $x^*\approx\bar{x}_5=1{,}736,\, f^*\approx-92{,}138$.

Заметим, что если воспользоваться формулой:

$n\ge\ln\left(\frac{\varepsilon}{b-a}\right)\Big/\ln\left(\frac{\sqrt{5}-1}{2}\right)\approx-2.1\cdot\ln\left(\frac{\varepsilon}{b-a}\right)$, то необходимое число шагов N можно определить заранее. В нашем случае N=4,79, т. е. N= 5, и отпадает необходимость во втором столбце таблицы.

Задача 11.

Найти частные производные первого и второго порядков от функции $z = \ln(x^2 - y^2)$.

Считая последовательно постоянной y, затем x, и применяя правило дифференцирования сложной функции, получим:

$$\frac{\partial z}{\partial x} = \frac{1}{(x^2 - y^2)} \cdot (x^2 - y^2)'_x = \frac{2x}{x^2 - y^2},$$

$$\frac{\partial z}{\partial y} = \frac{1}{(x^2 - y^2)} \cdot (x^2 - y^2)'_y = -\frac{2y}{x^2 - y^2}.$$ Дифференцируя вторично, получим:

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{2x}{x^2 - y^2}\right) = 2\frac{(x^2 - y^2) - x \cdot 2x}{(x^2 - y^2)^2} = -2\frac{x^2 + y^2}{(x^2 - y^2)^2},$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}\left(\frac{2x}{x^2 - y^2}\right) = 2x\left(-\frac{(-2y)}{(x^2 - y^2)^2}\right) = \frac{4xy}{(x^2 - y^2)^2},$$

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}\left(-\frac{2y}{x^2 - y^2}\right) = -2y\left(-\frac{2x}{(x^2 - y^2)^2}\right) = \frac{4xy}{(x^2 - y^2)^2},$$

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial y}\left(-\frac{2y}{x^2 - y^2}\right) = -2\frac{(x^2 - y^2) - y(-2y)}{(x^2 - y^2)^2} = -2\frac{x^2 + y^2}{(x^2 - y^2)^2}.$$

Задача 12.

Найти дифференциал функции $f(x, y, z) = x^2 y^3 / z^4$.

Первый способ. По формуле (5.4): $\frac{\partial f}{\partial x} = \frac{2xy^3}{z}, \frac{\partial f}{\partial y} = \frac{3x^2 y^2}{z^4}, \frac{\partial f}{\partial z} = -\frac{4x^2 y^3}{z^5}$,

$$df(x, y, z) = \frac{2xy^3}{z^4}dx + 3\frac{3x^2 y^2}{z^4}dy - \frac{4x^2 y^3}{z^5}dz = xy^2(2yzdx + 3xzdy - 4xydz)/z^5.$$

Второй способ. Применяем правила дифференцирования (5.5):

$$df(x, y, z) = d(x^2 y^3 \cdot \frac{1}{z^4}) = \frac{1}{z^4}d(x^2 y^3) + x^2 y^3 d(\frac{1}{z^4}) = (y^3 \cdot 2xdx + x^2 \cdot 3y^2 dy)/z^4 +$$

$$+ x^2 y^3 \cdot (-4dz/z^5) = xy^2(2yzdx + 3xzdy - 4xydz)/z^5.$$

Задача 13.

Найти дифференциалы 1-го, 2-го и 3-го порядков для функции $f(x,y)$.

По формуле (5.4): $df = f'_x dx + f'_y dy$. По формуле (5.6) при m = 2 и m = 3, считая dx и dy постоянными, последовательно находим (смешанные частные производные не зависят от порядка дифференцирования):

$$d^2 f = d(df) = d(f'_x dx + f'_y dy) = (f'_x dx + f'_y dy)'_x dx + (f'_x dx + f'_y dy)'_y dy =$$

$$= f''_{xx}(dx)^2 + 2 f''_{xy} dxdy + f''_{yy}(dy)^2;$$

$$d^3 f = d(d^2 f) = (f''_{xx}(dx)^2 + 2 f''_{xy} dxdy + f''_{yy}(dy)^2)'_x dx + (f''_{xx}(dx)^2 + 2 f''_{xy} dxdy +$$

$$+ f''_{yy}(dy)^2)'_y dy = f'''_{xxx}(dx)^3 + 3 f'''_{xxy}(dx)^2 dy + 3 f'''_{xyy} dx(dy)^2 + f'''_{yyy}(dy)^3.$$

Задача 14.

Даны функция $u=x^2+y^2+4x-6y+1$, точка $M_0(-1,2)$ и вектор $\vec{a} = 4\vec{i} - 3\vec{j}$.

1. Найти градиент функции в точке M_0 и наибольшую скорость изменения функции в точке M_0. Построить градиент.
2. Вычислить производную функции в точке M_0 по направлению вектора $\vec{a}$.
3. Составить уравнение линии уровня функции и построить ее график при a=4.

Решение.

Градиентом функции u=f(x,y) в точке $M_0(x_0,y_0)$ называется вектор, координаты которого равны значениям частных производных функции в точке M_0:

$$gra\overline{d}f(M_0) = \left\{ \frac{du}{dx}\Big|_{M_0}; \frac{du}{dy}\Big|_{M_0} \right\}.$$

Найдем значение частных производных функции в точке $M_0(-1, 2)$:

$$\frac{du}{dx} = (x^2 + y^2 + 4x - 6y + 1)'_x = 2x + 4; \quad \frac{du}{dx}\Big|_{M_0(-1,2)} = 2;$$

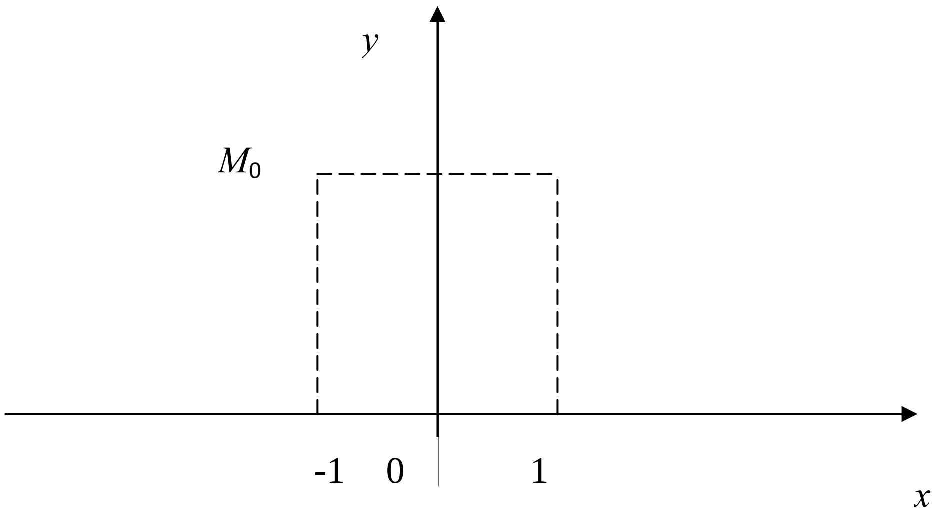
$$\frac{du}{dy} = (x^2 + y^2 + 4x - 6y + 1)'_y = 2y - 6; \quad \frac{du}{dy}\Big|_{M_0(-1,2)} = -2;$$

Вектор $gra\overline{d}f(M_0) = 2\vec{i} - 2\vec{j}$ указывает направление наискорейшего возрастания функции f в точке M_0.

Наибольшая скорость возрастания функции f равна модулю градиента:

$$\left|grad\,\overline{f}\,(M_0)\right| = \sqrt{\left(\left.\frac{du}{dx}\right|_{M_0}\right)^2 + \left(\left.\frac{du}{dy}\right|_{M_0}\right)^2} = \sqrt{2^2 + (-2)^2} = 2\sqrt{2}.$$

Построим градиент, начало которого находится в точке $M_0(-1, 2)$:



Задача 15.

Классический метод минимизации.

Решить задачу $f(x) = x^2_1 + x^2_2 + x^2_3 + x_1 - x_3 - x_2 x_3 \to \min$.

Шаг 1. Запишем систему:

$$\frac{df}{dx_1} = 2x_1 + 1 = 0;\ \frac{df}{dx_2} = 2x_2 - x_3 = 0;\ \frac{df}{dx_3} = 2x_3 - 1 - x_2 = 0.$$

Решив ее, получим стационарную точку $x^0 = \left(-\frac{1}{2}, \frac{2}{3}, \frac{1}{3}\right)$.

Шаг 2. Находим гессиан $f\,''(x^0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Так как, согласно критерию Сильвестра, эта матрица положительно определена, заключаем, что x^0 является точкой минимума функции $f(x)$.

Минимальное значение $f^* \approx f(x^0) = -19/12$.

Задача 16.

Найти локальный экстремум функции $z = x^3 + y^3 - 3xy$.

Решение. Находим частные производные функции:

$$z'_x = 3x^2 - 3y; \quad z'_y = 3y^2 - 3x$$

Приравниваем частные производные к нулю:

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}.$$

Решаем систему уравнений:

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ 3x^4 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ 3x\left(x^3 - 1\right) = 0 \end{cases} \Leftrightarrow \left[\begin{array}{l} \begin{cases} x = 0 \\ y = 0 \end{cases} \\ \begin{cases} x = 1 \\ y = 1 \end{cases} \end{array} \right..$$

Имеем две стационарные точки (0,0) и (1,1).

Найдем вторые частные производные:

$$z''_{x^2} = 6x, \;\; z''_{xy} = -3, \;\; z''_{yx} = -3, \;\; z''_{y^2} = 6y.$$

Вычисляем значения вторых частных производных в каждой стационарной точке, составляем определитель Δ и применяем достаточные условия экстремума.

$$a_{11} = z''_{x^2}(0,0) = 0, \;\; a_{12} = z''_{xy}(0,0) = -3, \;\; a_{21} = z''_{yx}(0,0) = -3, \;\; a_{22} = z''_{x^2}(0,0) = 0$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 0 & -3 \\ -3 & 0 \end{vmatrix} = 0 \cdot 0 - (-3) \cdot (-3) = -9$$

$$a_{11} = z''_{x^2}(1,1) = 6, \;\; a_{12} = z''_{xy}(1,1) = -3, \;\; a_{21} = z''_{yx}(1,1) = -3, \;\; a_{22} = z''_{x^2}(1,1) = 6$$

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 6 & -3 \\ -3 & 6 \end{vmatrix} = 6 \cdot 6 - (-3) \cdot (-3) = 27$$

Достаточные условия экстремума функции двух переменных:

А) Если $\Delta>0$ и $a_{11}<0$ ($a_{22}<0$), то в точке функция имеет максимум; если $\Delta>0$ ($a_{22}>0$), то в точке минимум.

Б) Если Δ<0, то экстремума нет.

В) Если Δ=0, то вопрос об экстремуме остается открытым.

В точке (0,0) Δ<0, значит, экстремума нет. В точке (1,1) Δ>0 и a_{11}>0, следовательно, точка (1,1) является точкой минимума функции. Вычислим значение функции в этой точке.

$$z(1,1) = 1 + 1 - 3 = -1$$

Ответ: (1,1) – точка минимума, $z(1,1) = -1$.

Задача 17.

Исследовать на экстремум функцию $z = x^3 + y^3 - 3xy$.

Из необходимого условия экстремума функции (теорема 9.7) имеем систему $\begin{cases} z'_x = 3x^2 - 3y = 0, \\ z'_y = 3y^2 - 3x = 0, \end{cases}$ решая которую получаем критические точки $M_1(0;0)$, $M_2(1;1)$. Определим характер критических точек по достаточным условиям экстремума. Находим $z''_{xx}(x,y) = 6x, z''_{xy}(x,y) = -3, \; z''_{yy}(x,y) = 6y$. В точке $M_1(0;0)$: $z''_{xx}(M_1) = 0$, $z''_{xy}(M_1) = -3$, $z''_{yy}(M_1) = 0$, $\Delta = z_{xx} \cdot z_{yy} - (z_{xy})^2\Big|_{M_1(0;0)} = -9 < 0$. Следовательно, $M_1(0;0)$ – седловая точка. В точке $M_2(1;1)$: $z''_{xx}(M_2) = 6, \; z''_{xy}(M_2) = -3$, $z''_{yy}(M_2) = 6$, $\Delta = 6 \cdot 6 - (-3)^2 = 27 > 0$, поэтому $M_2(1;1)$ – точка минимума функции z; $z_{\min} = z(M_2) = -1$.

Задача 18. (минимизация функции нескольких переменных методом Ньютона).

Минимизировать функцию $f(x)$ методом Ньютона с заданной точностью $\varepsilon = 10^{-3}$.

$$f(x) = x_1^2 + 2x_2^2 + x_1^2 x_2^2 + x_3^2 + \exp(x_2^2 + x_3^2) - x_2 + x_3$$

Решение.

Итерационная формула метода Ньютона для минимизации функции трех переменных имеет вид:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[j+1]} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[j]} - H^{-1}\left(f(x_1, x_2, x_3)^{[j]}\right)\nabla f(x_1, x_2, x_3)^{[j]}.$$

Найдем градиент и матрицу Гессе функции

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{pmatrix}.$$

$$\frac{\partial f}{\partial x_1} = 2x_1 + 2x_1 x_2^2$$

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2x_1^2 x_2 + 2x_2 \cdot \exp\left(x_2^2 + x_3^2\right) - 1$$

$$\frac{\partial f}{\partial x_3} = 2x_3 + 2x_3 \cdot \exp\left(x_2^2 + x_3^2\right) + 1$$

Составим градиент

$$\nabla f = \begin{pmatrix} 2x_1 + 2x_1 x_2^2 \\ 4x_2 + 2x_1^2 x_2 + 2x_2 \cdot \exp\left(x_2^2 + x_3^2\right) - 1 \\ 2x_3 + 2x_3 \cdot \exp\left(x_2^2 + x_3^2\right) + 1 \end{pmatrix}.$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2 + 2x_2^2$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4 + 2x_1^2 + 2 \cdot \exp\left(x_2^2 + x_3^2\right) + 4x_2^2 \cdot \exp\left(x_2^2 + x_3^2\right)$$

$$\frac{\partial^2 f}{\partial x_3^2} = 2 \cdot \exp\left(x_2^2 + x_3^2\right) + 4x_3^2 \cdot \exp\left(x_2^2 + x_3^2\right) + 2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 4x_1x_2, \qquad \frac{\partial^2 f}{\partial x_1 \partial x_3} = 0$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_3} = 4x_2x_3 \cdot \exp\left(x_2^2 + x_3^2\right).$$

$$H(f) = \begin{pmatrix} 2+2x_2^2 & 4x_1x_2 & 0 \\ 4x_1x_2 & 4+2x_1^2+2\cdot\exp\left(x_2^2+x_3^2\right)+4x_2^2\cdot\exp\left(x_2^2+x_3^2\right) & 4x_2x_3\cdot\exp\left(x_2^2+x_3^2\right) \\ 0 & 4x_2x_3\cdot\exp\left(x_2^2+x_3^2\right) & 2\cdot\exp\left(x_2^2+x_3^2\right)+4x_3^2\cdot\exp\left(x_2^2+x_3^2\right)+2 \end{pmatrix}$$

Первая итерация.

Пусть $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[0]} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}^{[0]}$

$$H^{[0]} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4+2 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

$$\left(H^{[0]}\right)^{-1} = \frac{1}{48}\begin{pmatrix} 24 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 2 \end{pmatrix} \approx \begin{pmatrix} 0{,}5 & 0 & 0 \\ 0 & 0{,}16667 & 0 \\ 0 & 0 & 0{,}25 \end{pmatrix}$$

$$\nabla f^{[0]} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

Подставим

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[1]} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[0]} - H^{-1}\left(f(x_1, x_2, x_3)^{[0]}\right)\nabla f(x_1, x_2, x_3)^{[0]}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[1]} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}^{[0]} - \frac{1}{48}\begin{pmatrix} 24 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[1]} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}^{[0]} - \frac{1}{48} \cdot \begin{pmatrix} 0 \\ -8 \\ -0{,}25 \end{pmatrix}$$

Проверим

$$\left| \frac{\partial f\left(x^{[k]}\right)}{\partial x_i} \right| \le 0{,}001$$

$$\frac{\partial f\left(x^{[1]}\right)}{\partial x_i} = \begin{pmatrix} 2 \cdot 0 + 2 \cdot 0 \cdot \frac{1}{6} \\ 4 \cdot \frac{1}{6} + 2 \cdot 0 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} \cdot \exp\left(\frac{1}{36} + \frac{1}{16} \right) - 1 \\ -2 \cdot \frac{1}{4} - 2 \cdot \frac{1}{4} \cdot \exp\left(\frac{1}{36} + \frac{1}{16} \right) + 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{3} \cdot \exp\left(\frac{13}{144} \right) - \frac{1}{3} \\ \frac{1}{2} - \frac{1}{2} \cdot \exp\left(\frac{13}{144} \right) \end{pmatrix}$$

$$\left| \frac{\partial f\left(x^{[1]}\right)}{\partial x_i} \right| = \sqrt{0^2 + \left(\frac{1}{3} \cdot \left(-1 + \exp\left(\frac{13}{144} \right) \right) \right)^2 + \left(\frac{1}{2} \cdot \left(1 - \exp\left(\frac{13}{144} \right) \right) \right)^2} =$$

$$= \sqrt{\frac{1}{9} \cdot \left(1 - \exp\left(\frac{13}{144} \right) \right)^2 + \frac{1}{4} \cdot \left(1 - \exp\left(\frac{13}{144} \right) \right)^2} =$$

$$= \sqrt{\left(\frac{1}{9} + \frac{1}{4} \right) \left(1 - \exp\left(\frac{13}{144} \right) \right)^2} \approx 0{,}056774 \ge 0{,}001.$$

Вторая итерация.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[1]} = \begin{pmatrix} 0 \\ \frac{1}{6} \\ -0{,}25 \end{pmatrix}$$

$$H^{[1]} = \begin{pmatrix} 2{,}0555 & 0 & 0 \\ 0 & 6{,}3105 & -0{,}1824 \\ 0 & -0{,}1824 & 4{,}46257 \end{pmatrix}$$

$$\left(H^{[1]}\right)^{-1} = \begin{pmatrix} 0{,}486486 & 0 & 0 \\ 0 & 0{,}15865 & 0{,}006485 \\ 0 & 0{,}006485 & 0{,}224351 \end{pmatrix}$$

$$\nabla f^{[1]} = \begin{pmatrix} 0 \\ 0{,}031493 \\ -0{,}04724 \end{pmatrix}$$

Подставим

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[2]} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[1]} - H^{-1}\left(f(x_1, x_2, x_3)^{[1]}\right)\nabla f(x_1, x_2, x_3)^{[1]}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[2]} = \begin{pmatrix} 0 \\ \frac{1}{6} \\ -0{,}25 \end{pmatrix}^{[1]} - \begin{pmatrix} 0{,}486486 & 0 & 0 \\ 0 & 0{,}15865 & 0{,}006485 \\ 0 & 0{,}006485 & 0{,}224351 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0{,}031493 \\ -0{,}04724 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[2]} = \begin{pmatrix} 0 \\ 1 \\ -0{,}25 \end{pmatrix}^{[1]} - \begin{pmatrix} 0 \\ 0{,}00469 \\ -0{,}01039 \end{pmatrix} = \begin{pmatrix} 0 \\ 0{,}161977 \\ -0{,}23961 \end{pmatrix}$$

Проверим

$$\left| \frac{\partial f\left(x^{[k]}\right)}{\partial x_i} \right| \le 0{,}001$$

$$\frac{\partial f\left(x^{2}\right)}{\partial x_i} = \begin{pmatrix} 0 \\ 0{,}000123 \\ -0{,}00023 \end{pmatrix}$$

$$\left| \frac{\partial f\left(x^{[2]}\right)}{\partial x_i} \right| = \sqrt{0^2 + 0{,}000123^2 + 0{,}00023^2} \approx 0{,}000264 < 0{,}001$$

Необходимое условие выполнено.

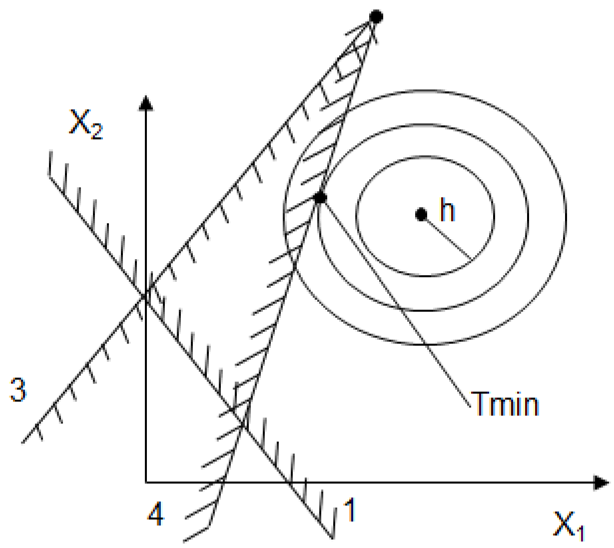
Ответ: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^{[2]} = \begin{pmatrix} 0 \\ 0{,}161977 \\ -0{,}23961 \end{pmatrix}$, $f(x) = 0{,}795547$.

Задача 19. Геометрический способ решения ЗНЛП

$$F = (x_1 - 3)^2 + (x_2 - 4)^2 \to \min, \max$$

$$\begin{cases} 3x_1 + 2x_2 \geq 7 \\ 10x_1 - x_2 \leq 8 \\ -18x_1 + 4x_2 \leq 12 \end{cases}$$

$$x_1, x_2 \geq 0$$



$$(x_1 - 3)^2 + (x_2 - 4)^2 = h$$

$$2(x_1 - 3) + 2(x_2 - 4)x_2' = 0$$

$$x_2' = \frac{-2(x_1 - 3)}{2(x_2 - 4)} = \frac{x_1 - 3}{4 - x_2}$$

$$\begin{cases} x_1 - 3 = 10(4 - x_2) \\ 10x_1 - x_2 = 8 \end{cases}$$

$$x_1^* = \frac{123}{101}; x_2^* = \frac{422}{101}$$

$$f(\min) = \frac{324}{101}$$

$$x^{**}{}_1 = 2, x^{**}{}_2 = 12,$$

$$f(\max) = 65$$

Задача 20. (метод Лагранжа)

Методом Лагранжа найти экстремум функции $u = x + y + z^2$ при условиях связи $\begin{cases} z - x = 1 \\ y - xz = 1. \end{cases}$

Решение. Составим функцию Лагранжа

$L(x, y, z, \lambda_1, \lambda_2) = x + y + z^2 + \lambda_1(z - x - 1) + \lambda_2(y - xz - 1)$ и рассмотрим систему уравнений

$$\begin{cases} \dfrac{\partial L}{\partial x} = 1 - \lambda_1 - \lambda_2 z = 0, \\ \dfrac{\partial L}{\partial y} = 1 + \lambda_2 = 0, \\ \dfrac{\partial L}{\partial z} = 2z + \lambda_1 - \lambda_2 x = 0, \\ \varphi_1 = z - x - 1 = 0, \\ \varphi_2 = y - xz - 1 = 0. \end{cases}$$

Она имеет единственное решение $x = -1$, $y = 1$, $z = 0$, $\lambda_1 = 1$, $\lambda_2 = -1$ то есть $M_0 = (-1,1,0)$ – единственная точка возможного экстремума функции при заданных условиях связи. Вычислим второй дифференциал функции Лагранжа $d^2F = 2(dz)^2 - 2\lambda_2 dxdz$ и подставляя $\lambda_2 = -1$ и $dz = dx$, найденное из первого уравнения связи, получаем положительно определенную квадратичную форму от переменной dx: $4(dx)^2 > 0$ при $dx \neq 0$. Отсюда следует, что функция при заданных условиях связи имеет в точке $M_0 = (-1,1,0)$ условный минимум.

Задача 21. (показывает, что в правиле множителей Лагранжа не всегда можно полагать $\lambda_0 = 1$**).**

$$f_0(x_1, x_2) = x_1 \to \inf; \quad f_1(x_1, x_2) = x_1^3 - x_2^2 = 0\,.$$

Решение. Функции f_0 и f_1 непрерывно дифференцируемы. Из условия $x_1^3 - x_2^2 = 0$ следует, что $x_1^3 = x_2^2 \geq 0 \Rightarrow x_1 \geq 0$. Поэтому очевидно, что решение задачи $\hat{x} = (0,\ 0)$. Если прямо следовать Лагранжу, то надо положить $\lambda_0 = 1$, составить сумму $L = x_1 + \lambda\left(x_1^3 - x_2^2\right)$ и далее решать систему уравнений $\begin{cases} L_{x_1} = 0 \\ L_{x_2} = 0 \end{cases} \Rightarrow \begin{cases} 1 + 3\lambda x_1^2 = 0 \\ -2\lambda x_2 = 0 \end{cases}$. Из последней системы следует, что $\lambda \neq 0$ (так как в противном случае не будет удовлетворяться первое уравнение системы), тогда решение системы имеет вид: $x_1 = \sqrt{-\dfrac{1}{3\lambda}}$, $x_2 = 0$. Но при этих значениях x_1, x_2 не будет удовлетворяться уравнение связи $x_1^3 - x_2^2 = 0$. Таким образом, получим, что решения нет, а это неверно.

Задача 22. (показывает, что экстремум функции Лагранжа как задачи без ограничений может не совпадать с экстремумом исходной задачи с ограничениями**).**

$$f_0\left(x_1;\ x_2\right) = x_2^2 - x_1 \to \inf;\ \ f_1\left(x_1;\ x_2\right) = x_1 + x_1^3 = 0.$$

Решение. Очевидно, что решение задачи $\hat{x} = (0,\ 0)$ (так как из условия $x_1 + x_1^3 = 0$ следует, что $\hat{x}_1 = 0$, а решение задачи $x_2^2 \to \inf$ имеет вид $\hat{x}_2 = 0$). Функция Лагранжа записывается в виде: $L = \lambda_0\left(x_2^2 - x_1\right) + \lambda\left(x_1 + x_1^3\right)$. Необходимое условие экстремума: $\begin{cases} L_{x_1} = 0 \\ L_{x_2} = 0 \end{cases} \Rightarrow \begin{cases} -\lambda_0 + \lambda\left(1 + 3x_1^2\right) = 0 \\ 2\lambda_0 x_2 = 0 \end{cases}$. Если $\lambda_0 = 0$, то $\lambda \neq 0$ и из первого уравнения системы следует, что $1 + 3x_1^2 = 0$ – противоречие. Значит, $\lambda_0 \neq 0$. Положим $\lambda_0 = 1$. Тогда функция Лагранжа примет вид: $L == x_2^2 - x_1 + \lambda\left(x_1 + x_1^3\right)$. Очевидно, что $L(0,\ 0) = 0$. Пусть $\varepsilon \neq 0 -$ произвольное действительное число. Тогда $L(\varepsilon,\ 0) = -\varepsilon + \lambda\left(\varepsilon + \varepsilon^3\right) = \ = \varepsilon\left(\lambda\varepsilon^2 + \lambda - 1\right)$.

1) Если $\lambda < 0$, то $\lambda\varepsilon^2 + \lambda - 1 < 0$. Тогда $L(\varepsilon,\ 0) < 0$, при $\varepsilon > 0$ и $L(\varepsilon,\ 0) > 0$ при $\varepsilon < 0$.

2) Если $\lambda = 0$, то $L(\varepsilon,\ 0) = -\varepsilon$. Тогда $L(\varepsilon,\ 0) < 0$ при $\varepsilon > 0$ и $L(\varepsilon,\ 0) > 0$ при $\varepsilon < 0$.

3) Если $0<\lambda<1$, то $L(\varepsilon,\ 0)>0$ при $\varepsilon\in\left(-\sqrt{\frac{1-\lambda}{\lambda}};\ 0\right)$ и $L(\varepsilon,\ 0)<0$ при $\varepsilon\in\left(0;\ \sqrt{\frac{1-\lambda}{\lambda}}\right)$.

4) Если $\lambda=1$, то $L(\varepsilon,\ 0)=\varepsilon^3$ и, следовательно, $L(\varepsilon,\ 0)<0$ при $\varepsilon<0$ и $L(\varepsilon,\ 0)>0$ при $\varepsilon>0$.

5) Если $\lambda>1$, то $\lambda\varepsilon^2+\lambda-1>0$ и, следовательно, $L(\varepsilon,\ 0)<0$ при $\varepsilon<0$ и $L(\varepsilon,\ 0)>0$ при $\varepsilon>0$.

Таким образом, при любых λ функция L принимает в любой достаточно малой окрестности точки $(0,\ 0)$ как положительные значения, так и отрицательные значения. А это означает, что ни при каких λ эта функция в точке $\hat{x}=(0,\ 0)$ не имеет даже локального минимума. Значит, точка $\hat{x}=(0,\ 0)$ не является решением задачи, а это неверно.

Задача 5. На развитие двух предприятий выделено 2 млн. рублей. Если первому предприятию дадут x_1 млн. рублей, то прибыль, полученная от этого предприятия, будет равна $2\sqrt{x_1}$ млн. рублей, если x_2 млн. дадут второму, то прибыль от него будет равна $3\sqrt{x_2}$ млн. рублей. Определить, как следует распределить средства между предприятиями, чтобы суммарная прибыль была максимальной. Решим эту задачу методом множителей Лагранжа.

Задача состоит в отыскании точки глобального максимума функции $f=2\sqrt{x_1}+3\sqrt{x_2}$ при ограничении $x_1+x_2=2$.

Точку возможного максимума найдем методом множителей Лагранжа.

Функция Лагранжа имеет вид:

$L(x_1,x_2,\lambda)=2\sqrt{x_1}+3\sqrt{x_2}+\lambda(x_1+x_2-2)$.

Для отыскания точек возможных экстремумов составим систему:

$$\begin{cases} L'_{x_1} = \dfrac{1}{\sqrt{x_1}} + \lambda = 0 \\ L'_{x_2} = \dfrac{3}{2\sqrt{x_2}} + \lambda = 0 \\ L'_{\lambda} = x_1 + x_2 - 2 = 0 \end{cases}$$

Найдем ее решение. $-\lambda = \dfrac{1}{\sqrt{x_1}} = \dfrac{3}{2\sqrt{x_2}} \Rightarrow 2\sqrt{x_2} = 3\sqrt{x_1} \Rightarrow x_2 = \dfrac{9}{4}x_1$.

Подставим найденное соотношение $x_2 = \dfrac{9}{4}x_1$ в уравнение (3), получим

$x_1 + \dfrac{9}{4}x_1 - 2 = 0 \Rightarrow \dfrac{13}{4}x_1 = 2 \Rightarrow x_1 = \dfrac{8}{13}$ и тогда $x_2 = \dfrac{18}{13}$. Находим λ:

$$\lambda = -\frac{1}{\sqrt{x_1}} = -\sqrt{\frac{13}{8}}.$$

Итак, система имеет одно решение

$$P_0\left(\frac{8}{13};\frac{18}{13}\right);\ \lambda_0 = -\sqrt{\frac{13}{8}}$$

Исследуем найденную точку на локальный условный экстремум с помощью определителя $\Delta(L)$

$$y(x_1, x_2) = x_1 + x_2 - 2$$

$$y'_{x_1} = 1; y'_{x_2} = 1$$

$$L'_{x_1} = \frac{1}{\sqrt{x_1}} + \lambda$$

$$L''_{x_1^2} = -\frac{1}{2\sqrt{x_1^{\,3}}}; L''_{x_1^2}(P_0\lambda_0) = -\frac{13}{32}\cdot\sqrt{\frac{13}{2}}$$

$$L''_{x_1x_2} = 0$$

$$L'_{x_2} = \frac{3}{2\sqrt{x_2}} + \lambda.$$

$$L''_{x_2^2} = -\frac{3}{4\sqrt{x_2^{\,3}}}; L''_{x_2^2}(P_0\lambda_0) = -\frac{13}{72}\cdot\sqrt{\frac{13}{2}}$$

Подставив все в формулу, получаем

$$\Delta = -1\cdot\begin{vmatrix} 0 & 1 & 1 \\ 1 & -\frac{13}{32}\cdot\sqrt{\frac{13}{2}} & 0 \\ 1 & 0 & -\frac{13}{72}\cdot\sqrt{\frac{13}{2}} \end{vmatrix} = -\left(-1\cdot\begin{vmatrix} 1 & 0 \\ 1 & -\frac{13}{72}\cdot\sqrt{\frac{13}{2}} \end{vmatrix} + 1\cdot\begin{vmatrix} 1 & -\frac{13}{32}\cdot\sqrt{\frac{13}{2}} \\ 1 & 0 \end{vmatrix}\right) = -\left(\frac{13}{72}\cdot\sqrt{\frac{13}{2}} + \frac{13}{32}\cdot\sqrt{\frac{13}{2}}\right)\langle 0$$

Так как $\Delta < 0$, то $Po\left(\frac{8}{13};\frac{18}{13}\right)$ – точка локального условного максимума.

Чтобы показать, что именно в точке P_o достигается и глобальный максимум, перейдём к задаче на отыскивание безусловного максимума функции одной переменной. С помощью задачи $x_1 + x_2 = 2$, запишем условную функцию в виде:

$$f(x_1, x_2) = 2\sqrt{x_1} + 3\sqrt{x_2} = 2\sqrt{x_1} + 3\sqrt{2 - x_1} = y(x_1).$$

Требуется найти такую точку, где достигается наибольшее значение функции.

Область возможного изменения оставшейся переменной отрезок $[0;2]$.

Непрерывная функция на замкнутом отрезке обязательно достигает своего наибольшего значения либо в критических точках внутри отрезка, либо на концах отрезка: $y'(x_1) = \frac{1}{\sqrt{x_1}} - \frac{3}{2\sqrt{2 - x_1}} = \frac{2\sqrt{2 - x_1} - 3\sqrt{x_1}}{2\sqrt{x_1\sqrt{2 - x_1}}}$.

Из условия $y'(x_1) = 0$ находим стационарную точку

$$2\sqrt{2 - x_1} - 3\sqrt{x_1} = 0 \Rightarrow 2\sqrt{2 - x_1} = 3\sqrt{x_1} \Rightarrow 4(2 - x_1) = 9x_1 \Rightarrow 8 - 4x_1 = 9x_1 \Rightarrow 13x_1 = 8 \Rightarrow x_1 = \frac{8}{13} \in (0;2)$$

Точек, где производная не существует, внутри отрезка нет. Находим значение целевой функции в стационарной точке и на концах отрезка.

$$y(\frac{8}{13}) = 2\sqrt{\frac{18}{13}} + 3\sqrt{2 - \frac{8}{13}} = 2\sqrt{\frac{8}{13}} + 3\sqrt{\frac{18}{13}} \approx 1{,}56 + 3{,}35 = 5{,}09$$

$$y(0) = 3\sqrt{2} \approx 4{,}24 \qquad y(2) = 2\sqrt{2} \approx 2{,}83$$

Мы видим, что наибольшее значение достигается в точке $x_1 = \frac{8}{13}$.

Итак, глобальный максимум достигается при $x_1 = \frac{8}{13}$ *млн.руб.*, $x_2 = \frac{18}{13}$ *млн.руб.*

$$y = 2\sqrt{\frac{8}{13}} + 3\sqrt{\frac{18}{13}} = \frac{2\cdot 2\sqrt{2}}{\sqrt{13}} + \frac{3\cdot 3\sqrt{2}}{\sqrt{13}} = \frac{13\sqrt{2}}{\sqrt{13}} = \sqrt{26} \approx 5{,}09\text{млн.руб.}$$

Задача 6. (метод штрафных функций)

$x_1^2 + x_2^2 \to \min$

$x_1 + x_2 - 1 = 0$

Оптимум достигается в точке $(\frac{1}{2}, \frac{1}{2})$ и равен $\frac{1}{2}$.

Задача со штрафом при достаточно большом μ:

$x_1^2 + x_2^2 + \mu * (x_1 + x_2 - 1)^2, \qquad (x_1, x_2) \in R^2$

x_1 и x_2 – координаты вектора x.

При $\forall \mu \geq 0$ целевая функция этой задачи выпуклая. Необходимым и достаточным условием оптимальности является равенство нулю градиента функции

$grad\ f(X) = x_1^2 + x_2^2 + \mu * (x_1 + x_2 - 1)^2,$

то есть частные производные

$x_1 + \mu * (x_1 + x_2 - 1) = 0$

$x_1 + \mu * (x_1 + x_2 - 1) = 0.$

Решая эту систему из двух уравнений, получаем:

$x_1 = x_2 = \frac{\mu}{1 + 2 * \mu}$ при $\mu \to \infty$ $x_1 = x_2 = \frac{1}{2}$.

6. Задания для самостоятельной работы

Раздел 1. Оптимизационные задачи нелинейного программирования. Их классификация.

Задача 1.

Найти множество точек минимума U^* функции *f(x)* на множестве *U*.

$$f(x) = \begin{cases} x & при \quad |x| > 1, \\ 1 & при \quad |x| \leq 1, \end{cases} \qquad U = R$$

Задача 2.

Убедиться в унимодальности функций ***f(x)*** на указанных отрезках [a; b].

a) $f(x)=\frac{1}{4}x^4+x^2-8x+12, \quad [0;2]$

b) $f(x)=\frac{1}{2}x^2-\sin x, \quad [0;1]$

Задача 3.

На какие три части следует разбить отрезок [-1; 2], чтобы на каждой из них функция $f(x)=||x(x-1)|-1$ была унимодальной?

Задача 4.

Найти максимальное значение b, при котором функция $f(x)=-x^2+5x-6$ унимодальна на отрезке [-5; b].

Задача 5.

Показать, что если *f(x)* – выпуклая дифференцируемая функция, то любая касательная к графику *f(x)* лежит не выше этого графика.

Задача 6.

Показать, что выпуклая дифференцируемая на отрезке [a; b] функция унимодальна на этом отрезке.

Задача 7.

Показать, что если *f(x)* – выпуклая дифференцируемая функция, то любая касательная к графику *f(x)* лежит не выше этого графика.

Задача 8.

Показать, что если *f(x)* – выпуклая дифференцируемая функция, то любая касательная к графику *f(x)* лежит не выше этого графика.

Раздел 2. Безусловная нелинейная оптимизация функций одной переменной. Классификация методов.

Задача 1.

Методом перебора для функции $f(x)=\sqrt{1+x^2}+e^{-2x}, \quad [0;1] \quad \varepsilon=0.1$ найти точку минимума x^* функции на отрезке [a; b] с точностью ε и минимум f^*.

Задача 2.

Показать, что для $f(x) \in Q[a;b]$ n шагов метода деления отрезка пополам обеспечивает вычисление точки минимума x^* на отрезке [a; b] с абсолютной погрешностью, не превосходящей $\varepsilon_n = \dfrac{b-a}{2^{n+1}} + \dfrac{\delta}{2}\left(1 - \dfrac{1}{2^n}\right)$.

Задача 3.

Найти число шагов n метода деления отрезка пополам, необходимое для определения точки минимума функции $f(x) \in Q[a;b]$ на отрезке [a; b] с точностью ε > 0.

Задача 4.

Достаточно ли вычисления 10 значений функции $f(x) \in Q[a;b]$ для определения ее точки минимума на отрезке [0; 1] с точностью ε = 0.02 методом деления отрезка пополам?

Задача 5.

Сравнить необходимые количества вычисленных значений $N_д$ и $N_п$ функции *f(x)* при поиске её точки минимума на отрезке длины 1 с точностью $\varepsilon = 10^{-5}$ методами деления отрезка пополам и перебора соответственно.

Задача 6.

Сравнить необходимые количества вычисленных значений $N_д$ и $N_с$ функции *f(x)* при её минимизации на отрезке длины 1 с точностью $\varepsilon = 10^{-5}$ методами деления отрезка пополам и золотого сечения соответственно.

Задача 7.

Методом золотого сечения найти точку минимума x^* функции *f(x)* на отрезке [a;b] с точностью ε и минимум f^*, если $f(x) = x^4 + 2x^2 + 4x + 1, \quad [-1;0] \quad \varepsilon = 0{,}1.$

Задача 8.

Показать, что если функция $f(x) = \dfrac{1}{3}x^3 + 2x^2 - 5x + 6$ удовлетворяет условию Липшица, то она непрерывна на [a; b].

Найти наименьшую из констант Липшица функции на отрезке: а) [0; 1];

б)[0; 10].

Задача 9.

Убедившись в выпуклости функции на отрезке [a;b] $f(x) = x - \ln x, \quad [0{,}1;2]$, найти её точку минимума x^* и минимальное значение f^* методом касательных, используя в качестве условия достижения требуемой точности неравенство $\left| f'(c_n^*) \right| \le 0{,}01$.

Задача 10

1. Для заданной целевой функции $f(x) = x^4 + e^{-x}$ произвести графический анализ функции. Найти промежуток ($X \subset R$), на котором функция унимодальна. Построить графики первой и второй ее производных.
2. Найти методом золотого сечения точку минимума на этом отрезке с точностью ε=0,001. Вычисление вести с одним запасным знаком.
3. Найти минимум функции методом средней точки для заданной точности $\varepsilon=0{,}02$ на интервале [0,1].
4. Найти минимум функции методом хорд для точности $\varepsilon=0{,}05$.
5. Найти минимум функции методом Ньютона.
6. Для каждого метода изучить зависимость скорости работы (числа вычислений функции) от значения точности.
7. Проверить результаты вычислений с использованием надстроек Excel «Подбор параметра» и «Поиск решения».
8. Сделать выводы об эффективности работы методов.
9. Выполнить задания для функций по вариантам и оформить отчет. Отчет должен включать следующие пункты:
   1) Постановка задачи
   2) Проверка унимодальности аналитическим методом и по графику (Если функция не является унимодальной, то измените интервал так, чтобы условие унимодальности было выполнено)

3) Точный метод поиска экстремума

4) Приближенные методы поиска экстремума:

   - ✓ Для каждого метода описание алгоритма
   - ✓ Блок-схема
   - ✓ расчеты

5) Сравнение методов по количеству обращений к функции и по точности

6) Вывод: какой метод более эффективен.

Задачи по вариантам:

1. $y = x^2 - 2x + e^{-x}$ на отрезке $[0; \pi/4]$
2. $y = x^3 - 2x^2 - 7x + 4$ на отрезке $[-0{,}5; 3]$
3. $y = \dfrac{e^x}{x}$ на отрезке $[0{,}1; 2]$
4. $y = e^{\sqrt[3]{x^2}}$ на отрезке $[-1; 2]$
5. $y = \dfrac{x}{\ln x}$ на отрезке $[1{,}1; 4]$
6. $y = (2 + x)e^{-x}$ на отрезке $[-2; 1]$
7. $y = \ln(2 - Cosx)$ на отрезке $[-1; 3]$
8. $y = 3x^2 - 6xCosx$ на отрезке $[0; 3]$
9. $y = x \cdot Sinx + 2 \cdot Cosx$ на отрезке $[-5; -4]$
10. $y = 5x^2 - 8x^{\frac{5}{4}} - 20x$ на отрезке $[3; 3{,}5]$
11. $y = \dfrac{1 - x - x^2}{1 + x - x^2}$ на отрезке $\left[0; \dfrac{1 + \sqrt{5}}{2}\right]$
12. $y = 2 \cdot Sin2x + 3 \cdot Cos2x$ на отрезке $[0; \pi/3]$
13. $y = x \cdot Cos\, x$ на отрезке $[0; \pi]$
14. $y = x \cdot Sin\pi x$ на отрезке $[1{,}5; 2{,}5]$
15. $y = \sqrt{1 + x^2} + e^{-2x}$ на отрезке $[0; 1]$
16. $y = \dfrac{1}{\left|(x - 3)^3\right|}$ на отрезке $[2; 4]$

17. $y=\begin{cases} -4x & 0\le x\le 2 \\ x-4 & 2\le x\le 4 \end{cases}$

18. $y=\dfrac{x^3}{3}-5x-x\cdot\ln x$ на отрезке $[1{,}5;\ 2]$

19. $y=x^4+8x^3-6x^2-72x+90$ на отрезке $[1{,}5;\ 2]$

20. $y=x^6+3x^2+6x-1$ на отрезке $[-1;\ 0]$

21. $y=\dfrac{x^7}{7}-x^3-\dfrac{x^2}{2}-x$ на отрезке $[1;\ 1{,}5]$

22. $y=\sqrt{\dfrac{1+x}{\ln x}}$ на отрезке $[1{,}1;\ e]$

23. $y=e^{-x}-Cos\,x$ на отрезке $[-2;\ 1]$

24. $y=x^3-2x^2-7x+4$ на отрезке $[-0{,}5;\ 3]$

25. $y=\dfrac{e^{-x}}{x}$ на отрезке $[0{,}1;\ 2]$

26. $y=10e^{\sqrt[3]{x^2}}$ на отрезке $[-1;\ 2]$

27. $y=\dfrac{x}{\ln x}$ на отрезке $[1{,}1;\ 4]$

28. $y=(2+e^{-x})x$ на отрезке $[-1;\ 2]$

29. $y=\ln(2-Cosx)$ на отрезке $[-1;\ 2]$

30. $y=3x^2-6xSinx$ на отрезке $[0;3]$

31. $y=2+3\cdot Cos2x\cdot Sin2x$ на отрезке $[-5;\ -4]$

32. $y=x^4+2x^2-8x+16$ на отрезке $[3;\ 9]$

Раздел 3. Безусловная оптимизация функции многих переменных.

Задача 1.

Установить, являются ли выпуклыми множества U.

1) $U=\{(x_1,x_2)\,|\,x_1x_2>1,\quad x_1>0\}$

2) $U=\{(x_1,x_2)\,|\,x_2\ge x_1^2\}$

3) $U=\{(x_1,x_2)\,|\,x_1x_2<1,\quad x_1>0,\quad x_2>0\}$

4) $U=\{(x_1,x_2)\,|\,x_1-x_2\le 2,\quad x_1^2+x_2^2\le 4\}$

5) $U = \{(x_1, x_2, x_3) \mid x_3 \ge x_1^2 + x_2^2\}$

6) $U = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 \le 1\}$

7) $U = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 \le 1, \quad x_1 \ge 0, \quad x_2 \ge 0\}$

8) $U = \{(x_1, x_2, x_3) \mid x_1^2 + \dfrac{x_2^2}{2} + \dfrac{x_3^2}{3} \ge 1\}$

Задача 2.

Убедиться в выпуклости функции *f(x)* во всем пространстве E_n.

1) $f(x_1, x_2) = 4x_1^2 + x_2^2 - 2x_1x_2 + 6x_1 - x_2 - 2$

2) $f(x_1, x_2) = \sqrt{1 + x_1^2 + x_2^2}$

3) $f(x_1, x_2) = x_1^2 + x_2^2 - \cos\dfrac{x_1 - x_2}{2}$

Задача 3.

Указать множества *U*, на которых функции *f(x)* являются выпуклыми.

1) $f(x) = \dfrac{x_1^2}{x_2}$

2) $f(x) = x_1^2 + 2x_2^2 - \sin(x_1 - x_2)$

Задача 4.

При каких a, b и c функция $f(x) = ax_1^2 + bx_1x_2 + cx$ является выпуклой.

Задача 5.

При каких значениях a функция $f(x) = x_1^2 + x_2^2 + x_3^2 + ax_1x$ выпукла в E^3.

Задача 6.

Выписать матрицу Q квадратичной функции *f(x)*, найти её градиент $\nabla f(x^{(0)})$ в точке $x^{(0)}$ и убедиться в выпуклости *f(x)* в E^n.

$f(x) = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 - x_2x_3 + 2x_1 + x_3, \qquad x^{(0)} = (1;0;-1)$

Задача 7.

Найти экстремумы функций 2-х переменных:

1) $z = x^2 + xy + y^2 - 3x - 6y$

2) $z = xy^2(1 - x - y) \quad (x > 0, y > 0)$

3) $z = xy + 50/x + 20/y \ (x > 0, y > 0)$

4) $z = x^3 + 3xy^2 - 15x - 12y$

5) $z = (2x^2 + y^2)e^{-(x^2+y^2)}$.

Задача 8.

Совершить один шаг градиентного спуска из точки $x^{(0)}$ с шагом α_0 и сравнить значения $f(x^{(0)})$ и $f(x^{(1)})$.

$$f(x) = x_1^2 + 2x_2^2 + e^{x_1 + x_2}, \quad x^{(0)} = (1;1), \quad а)\alpha_0 = 0{,}1; \quad б)\alpha_0 = 0{,}265; \quad в)\alpha_0 = 0{,}5.$$

Задача 9.

Показать, что градиенты $\nabla f(x^{(k)})$ и $\nabla f(x^{(k+1)})$ в последовательных точках итерационного процесса метода наискорейшего спуска ортогональны, т. е. $(f'(x^{(k)}), f'(x^{(k+1)})) = 0, \quad k = 0,1,...$

Задача 10.

Для функции $f(x)$ найти величину шага α_0 метода наискорейшего спуска из точки $x^{(0)}$, если $f(x) = 2x_1^2 + x_2^2 + x_1x_2 + x_1 + x_2, \quad x^{(0)} = (0;0)$.

Задача 11.

Минимизировать методом наискорейшего спуска квадратичную функцию $f(x) = 7x_1^2 + 2x_1x_2 + 5x_2^2 + x_1 - 10x_{2.}$, заканчивая вычисления при $\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 0{,}01, \quad i = 1,2,...,n$.

Задача 12.

Минимизировать функцию $f(x) = x_1^4 + 2x_2^4 + x_1^2x_2^2 + 2x_1 + x_2$. методом сопряженных направлений, заканчивая вычисления при $\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 10^{-3}, \quad i = 1,2,...,n$.

Задача 13.

Показать, что точка минимума выпуклой квадратичной функции находится с помощью одной итерации метода Ньютона из произвольного начального приближения $x^{(0)} \in R^n$.

Задача 14.

Построить линии уровня и траектории подъема.

1) $f(x) = 6x_1 + 32x_2 - x_1^2 - 4x_2^2 \to \min, \quad x^0 = (7;4)$

2) $f(x) = 2x_2 - x_1^2, \quad x^0 = (0;-1)$

3) $f(x) = 2(x_1 - 1)^2 + 3(x_2 - 2)^2, \quad x^0 = (3;3)$

4) $f(x) = (x_1 - 2)^2 - (x_2 - 3)^2, \quad x^0 = (6;4)$

Задача 15. (по вариантам)

1. Построить график поверхности заданной функции в трехмерной системе координат. Графически отобразить линии уровня функции.
2. Найти точку минимума аналитически.
3. Методом покоординатного спуска с точностью $\varepsilon = 10^{-3}$ и $\varepsilon = 10^{-5}$.
4. Методом наискорейшего спуска.
5. Методом сопряженных градиентов.
6. Методом Ньютона.
7. Проверить вычисления при различных начальных векторах X_0 и проследить зависимость числа итераций от выбора X_0.
8. Графически представить траектории движения к экстремуму, полученные соответствующими методами.
9. Сравнить эффективность численных методов по числу итераций.
10. Выполнить задания для функций по вариантам и оформить отчет: (постановка проблемы, описание всех методов, результаты, выводы).

Варианты:

1. $f(X) = 64x_1^2 + 126x_1x_2 + 64x_2^2 - 10x_1 + 30x_2 + 13$

2. $f(X) = 2x_1^2 - x_1x_2 + x_2^2 - x_1 - x_2 + 1$

3. $f(X) = 129x_1^2 - 256x_1x_2 + 129x_2^2 - 51x_1 - 149x_2 - 27$

4. $f(X) = x_1^4 - 2x_1x_2 + x_2^4 - x_1^2 - x_2^2$

5. $f(X) = 254x_1^2 + 506x_1x_2 + 254x_2^2 + 50x_1 + 130x_2 - 111$

6. $f(X) = (x_1 - 4)^2 + 10(x_2 - 5)^2 - 5$

7. $f(X) = 151x_1^2 - 300x_1x_2 + 151x_2^2 + 33x_1 + 99x_2 + 48$

8. $f(X) = x_1^2 + 2x_2^2 - 4x_1 - 4x_2$

9. $f(X) = 85x_1^2 + 168x_1x_2 + 85x_2^2 + 29x_1 - 51x_2 + 83$

10. $f(X) = 16(x_1 + 5)^4 + 3(x_2 - 1)^2$

11. $f(X) = 211x_1^2 - 420x_1x_2 + 211x_2^2 - 192x_1 + 50x_2 - 25$

12. $f(X) = x_1^2 + 10x_2^2 - 4x_1 - 11x_2$

13. $f(X) = 194x_1^2 + 376x_1x_2 + 194x_2^2 + 31x_1 - 229x_2 + 4$

14. $f(X) = (x_1 - 2)^4 + 300(x_2 + 2)^2$

15. $f(X) = 45x_1^2 - 88x_1x_2 + 45x_2^2 + 102x_1 + 268x_2 - 21$

16. $f(X) = 99x_1^2 + 196x_1x_2 + 99x_2^2 - 95x_1 - 9x_2 + 91$

17. $f(X) = x_1^3 + x_2^3 - 3x_1x_2$

18. $f(X) = 20(x_1 - 4)^4 + 300(x_2 - 5)^2$

19. $f(X) = x_1^2 + x_1x_2 + x_2^2 + x_1 - x_2 + 1$

20. $f(X) = 2x_1^2 + 4x_1x_2 + 8x_2^2 + 100$

21. $f(X) = x_1^4 - 2x_1x_2 + x_2^4 + x_1^2 - x_2^2$

Раздел 4. Условная оптимизация функции многих переменных.

Задание 1. Определить глобальные экстремумы функций на множестве:

Решить задачи по вариантам.

1. $z = \sqrt{x_1^2 + x_2^2}$

при условиях:
$$\begin{cases} 3x_1 + 4x_2 \le 24 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

2. $z = (x_1 - 2)^2 + (x_2 - 3)^2$

при условиях:
$$\begin{cases} x_1 + 2x_2 \le 12 \\ x_1 + x_2 \le 9 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

3. $z = (x_1 - 4)^2 + (x_2 - 6)^2$

при условиях:
$$\begin{cases} 2x_1 + 3x_2 \le 12 \\ x_1 + x_2 \ge 1 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

4. $z = x_1 \cdot x_2$

при условиях:
$$\begin{cases} 2x_1 + x_2 \le 10 \\ x_1 + x_2 \le 6 \\ x_1 + 2x_2 \ge 2 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

5. $z = 2x_1 + 3x_2 - 2x_1^2$

при условиях:
$$\begin{cases} x_1 + 2x_2 \le 4 \\ x_1 + x_2 \le 2 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

6. $z = (x_1 - 6)^2 + (x_2 - 2)^2$

при условиях:
$$\begin{cases} x_1 + 2x_2 \le 8 \\ 3x_1 + x_2 \le 15 \\ x_1 + x_2 \ge 1 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

7. $z = x_1 + 3x_2$

при условиях:
$$\begin{cases} (x_1 - 5)^2 + (x_2 - 3)^2 \ge 9 \\ (x_1 - 5)^2 + (x_2 - 3)^2 \le 36 \\ x_1 + x_2 \ge 8 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

8. $z = 2(x_1 - 5)^2 + (x_2 - 7)^2$

при условиях:
$$\begin{cases} x_1 + x_2 \le 9 \\ x_1 + 2x_2 \le 19 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

9. $z = 2x_1 + x_2 - x_1^2$
при условиях:

$$\begin{cases} 3x_1 + 2x_2 \le 12 \\ x_1 \ge 0 \\ 0 \le x_2 \le 3 \end{cases}$$

10. $z = 2(x_1 - 7)^2 + 4(x_2 - 3)^2$
при условиях:

$$\begin{cases} x_1 + 2x_2 \ge 2 \\ x_1 + x_2 \le 6 \\ 2x_1 + x_2 \le 10 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

11. $z = x_1 + x_2$
при условиях:

$$\begin{cases} (x_1 - 5)^2 + (x_2 - 3)^2 \ge 9 \\ (x_1 - 5)^2 + (x_2 - 3)^2 \le 36 \\ x_1 + x_2 \ge 8 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

12. $z = x_1 \cdot x_2$
при условиях:

$$\begin{cases} x_1 + 2x_2 \ge 2 \\ x_1 + x_2 \le 6 \\ 2x_1 + x_2 \le 10 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

13. $z = 2x_1 + 3x_2 - 0{,}2x_1^2 - 0{,}2x_2^2$
при условиях:

$$\begin{cases} x_1 + 3x_2 \le 13 \\ 2x_1 + x_2 \le 10 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

14. $z = x_1 + x_2$
при условиях:

$$\begin{cases} (x_1 - 5)^2 + (x_2 - 3)^2 \ge 9 \\ (x_1 - 5)^2 + (x_2 - 3)^2 \le 36 \\ x_1 + x_2 \ge 8 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

15. $z = 3x_1 + 6x_2 - 0{,}3x_1^2 - 0{,}3x_2^2$
при условиях:

$$\begin{cases} 9x_1 + 8x_2 \le 72 \\ x_1 + 2x_2 \le 10 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

16. $z = 3x + y$
при условиях:

$$\begin{cases} x^2 + y^2 \le 40 \\ x^2 + y^2 \ge 4 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

17. $z = x_1 \cdot x_2$

при условиях:

$$\begin{cases} (x_1-5)^2+(x_2-3)^2 \ge 9 \\ (x_1-5)^2+(x_2-3)^2 \le 36 \\ x_1+x_2 \ge 8 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

18. $z = (x_1-3)^2 + 2(x_2-2)^2$

при условиях:

$$\begin{cases} x_1+4x_2 \le 16 \\ 3x_1+x_2 \le 15 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

19. $z = 2 - x_1^2 - x_2^2$

при условиях:

$$\begin{cases} x_2 \le 4 - x_1^2 \\ x_1+x_2 \ge 1 \\ x_1 \ge 0 \\ x_2 \ge 0 \end{cases}$$

20. $z = x_2 - x_1^2$

при условиях:

$$\begin{cases} 2x_1+3x_2^2 \le 3 \\ x_2 \ge 0 \\ 0 \le x_1 \le \dfrac{2}{3} \end{cases}$$

Задача 2.

Исследуйте на условный экстремум методом исключения части переменных функцию:

а) $u = x^2 + y^2$ при условии связи $\dfrac{x}{a} + \dfrac{y}{b} = 1$;

б) $u = x + y$ при условии связи $\dfrac{1}{x^2} + \dfrac{1}{y^2} = \dfrac{1}{a^2}$;

в) $u = x \cdot y$ при условии связи $x^2 + y^2 = 1$;

г) $u = x^2 + y^2 + 2z$ при условии связи $x - y + z = 1$;

д) $u = x - 2y + z$ при условии связи $x + y^2 - z^2 = 1$;

е) $u = x \cdot y^2 \cdot z^3$ при условии связи $x + 2y + 3z = 6$, $(x > 0, y > 0, z > 0)$;

ж) $u = x - 2y + 2z$ при условии связи $x^2 + y^2 + z^2 = 1$.

Ответ: а) $u_{\min} = u(\dfrac{ab^2}{a^2+b^2}, \dfrac{a^2b}{a^2+b^2}) = \dfrac{a^2b^2}{a^2+b^2}$;

б) $u_{\min} = u(\sqrt{2}a, \sqrt{2}a) = 2\sqrt{2}a$; $u_{\max} = u(-\sqrt{2}a, -\sqrt{2}a) = -2\sqrt{2}a$

в) $u_{\max} = 0{,}5$ в точках $(1/\sqrt{2}, 1/\sqrt{2})$ и $(-1/\sqrt{2}, -1/\sqrt{2})$;

$u_{\min} = -0{,}5$ в точках $(-1/\sqrt{2}, 1/\sqrt{2})$ и $(1/\sqrt{2}, -1/\sqrt{2})$;

г) $u_{\min} = u(0{,}4; -0{,}4, 0{,}2) = 0{,}4$; д) нет точек экстремума;

е) $u_{\max} = u(1, 1, 1) = 1$;ж) $u_{\min} = u(-1/3, 2/3, -2/3) = -3$;

$u_{\max} = u(1/3, -2/3, 2/3) = 3$

2. Исследуйте на условный экстремум методом Лагранжа:

а) функцию $u = x \cdot y \cdot z$ при условии связи $x^2 + y^2 + z^2 = 3$

б) функцию $u = x \cdot y \cdot z$ при условиях связи

$x^2 + y^2 + z^2 = 1$, $x + y + z = 0$.

Ответ: а) $u_{\min} = -1$ в точках (-1,1,1),(1,-1,1),(1,1,-1),(-1,-1,-1),

$u_{\min} = -1$ в точках (1,1,1),(-1,-1,1),(-1,1,-1),(1,-1,-1);

б) $u_{\min} = -1/(3\sqrt{6})$ в точках $(1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$, $(1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})$,

$(-2/\sqrt{6}, 1/\sqrt{6}, 1/\sqrt{6})$

$u_{\max} = 1/(3\sqrt{6})$ в точках $(-1/\sqrt{6}, -/\sqrt{6}, 2/\sqrt{6})$, $(-1/\sqrt{6}, 2/\sqrt{6}, -1/\sqrt{6})$,

$(2/\sqrt{6}, -1/\sqrt{6}, -1/\sqrt{6})$

Библиография

1. Кремер Н.Ш. Исследование операций для экономистов. – М.: ЮНИТИ, 2006.
2. Таха Х. Введение в исследование операций. – М.: Вильямс, 2007.
3. Гончаров В.А. Методы оптимизации: учеб. пособие для студентов вузов по спец. 010501(010200) «Приклад. математика и информатика», 230105(220400) «Програм. обеспечение вычислит. техники и автоматизир. систем» / В. А. Гончаров. – М. : Юрайт : Высш. образование, 2010.
4. Базара, Шетти. Нелинейное программирование. Теория и алгоритмы. –М.: Мир, 1982.
5. Бейко И.В., Бублик Б.Н., Зинько П.Н. Методы и алгоритмы оптимизации. – М.: Высш. шк., 1983.
6. Пантелеев А. В., Летова Т.А. Методы оптимизации в примерах и задачах: учебное пособие, 2-е издание – М.: Высш. шк. , 2005, – 544 с.
7. Аттетков А. В. , Зарубин В. С., Канатников А. Н. Введение в методы оптимизации : учебное пособие. – М. : Финансы и статистика, 2014.
8. Васильева О. А. , Ларионов Е. А., Лемин А. Ю., Макаров В. И. Методы оптимизации : учебное пособие. – М. : Московский государственный строительный университет, ЭБС АСВ, 2014.
9. Ф. Гилл, У. Мюррей, М. Райт: Практическая оптимизация .– М.: Мир, 1985.
10. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах: Учебное пособие. – М.: Высш. шк., 2002, -544 с.

Прокопенко Наталья Юрьевна

МЕТОДЫ ОПТИМИЗАЦИИ

Учебное пособие

Подписано в печать Формат 60х90 1/16 Бумага газетная. Печать трафаретная.
Уч. изд. л. 6,8. Усл. печ. л. 7,3. Тираж 300 экз. Заказ №

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Нижегородский государственный архитектурно-строительный университет»
603950, Нижний Новгород, ул. Ильинская, 65.
http://www. nngasu.ru, srec@nngasu.ru