Глава 1. КЛАССИЧЕСКИЕ МЕТОДЫ ПОИСКА ЭКСТРЕМУМА

1.1. Основные понятия теории экстремальных задач

Исходными данными при постановке задачи поиска экстремума является множество X и определенная на нем функция $f(x)$. Как уже говорилось, мы будем рассматривать конечномерные задачи, поэтому x является вектором произвольной размерности n, т.е. $x = (x_1, \ldots, x_n)$, а множество X — подмножеством евклидова пространства R^n (возможно совпадающим со всем пространством). Помимо задания X (его называют допустимым множеством) и $f(x)$ (ее называют целевой функцией) необходимо определить, что понимается под решением задачи. Во-первых, речь может идти о нахождении точек максимума (одной или всех), минимума (одной или всех) или тех и других. Во-вторых, необходимо уточнить само понятие максимума (минимума), так как оно может пониматься в глобальном и локальном смысле.

Определение 1. Точка $x^* \in X$ называется точкой глобального (абсолютного) максимума функции $f(x)$ на множестве X, если

$$f(x^*) \geq f(x) \quad \forall x \in X. \tag{1.1}$$

Определение 2. Точка $x^* \in X$ называется точкой локального максимума функции $f(x)$ на множестве X, если $\exists\, \varepsilon > 0$ такое, что

$$f(x^*) \geq f(x) \ \forall x \in X \cap U_\varepsilon(x^*), \tag{1.2}$$

где $U_\varepsilon(x^*)$ — ε-окрестность точки x^* (шар радиусом ε с центром в x^*).

Определения глобального и локального минимумов получаются заменой в (1.1) и (1.2) знаков неравенств на противоположные.

Глобальные максимумы и минимумы называют глобальными экстремумами функции, локальные максимумы и минимумы — локальными экстремумами. Ясно, что глобальные экстремумы являются и локальными, но обратное, вообще говоря, не верно. Для того чтобы найти глобальный максимум, надо найти все локальные максимумы и выбрать из них тот, в котором функция принимает наибольшее значение (аналогично для минимумов).

Если x^* есть точка глобального максимума $f(x)$ на X, то будем использовать запись

$$f(x^*) = \max_{x \in X} f(x), \tag{1.3}$$

т.е. под $\max\limits_{x \in X} f(x)$ понимается максимальное (абсолютное) значение $f(x)$ на X. Для точки глобального максимума принято обозначение

$$x^* = \arg \max_{x \in X} f(x). \tag{1.4}$$

Частным решением задачи поиска максимума называется нахождение одной точки x^* и значения $f(x^*)$, определяемых (1.3), (1.4). Полным решением называется нахождение величины $\max\limits_{x \in X} f(x)$ и всех реализующих ее значений аргумента, множество которых принято обозначать

$$Arg \max_{x \in X} f(x) = \{x^* \in X \mid f(x^*) = \max_{x \in X} f(x)\}. \tag{1.5}$$

Обычно, если это не оговаривается особо, под решением задачи максимизации будем понимать нахождение ее частного решения.

Аналогичные понятия и обозначения используются для задачи поиска глобального минимума (соответственно $\min\limits_{x \in X} f(x)$, $arg \min\limits_{x \in X} f(x)$, $Arg \min\limits_{x \in X} f(x)$).

Множество всех точек локальных максимумов включает в себя множество (1.5), но, вообще говоря, с ним не совпадает. Задачу нахождения всех локальных максимумов мы будем записывать в виде

$$f(x) \to max, x \in X. \tag{1.6}$$

Аналогично для локальных минимумов

$$f(x) \to min, x \in X, \tag{1.7}$$

а для задачи нахождения всех локальных экстремумов (максимумов и минимумов)

$$f(x) \to extr, x \in X. \tag{1.8}$$

Так как, очевидно, точки максимума функции $f(x)$ совпадают с точками минимума функции $-f(x)$, то все свойства и методы решения задачи (1.6) легко переносятся на (1.7), (1.8) (мы будем их рассматривать на примере задачи максимизации).

Экстремальные задачи принято делить на задачи поиска безусловного и условного экстремума. Если $X = R^n$, т.е. рассматриваются точки экстремума $f(x)$ на всем n-мерном пространстве, то имеем задачу на безусловный экстремум (или без ограничений). Если $X \subset R^n$ (собственное подмножество n-мерного пространства), то имеем задачу на условный экстремум (или с ограничениями). Терминология эта возникла в связи с тем, что множество X в последнем случае обычно задается какими-то условиями (обычно ограничениями вида равенств и неравенств). В классическом анализе рассматриваются задачи без ограничений и с ограничениями типа равенств.

Задача поиска экстремума может, вообще говоря, не иметь решения (когда верхняя или нижняя грань функции на множестве не достигается). В курсе математического анализа доказывается следующий важный результат.

Теорема Вейерштрасса. *Если X — компакт в R^n (замкнутое ограниченное множество), $f(x)$ — непрерывная на X функция, то точки глобального максимума и минимума $f(x)$ на X существуют.*

Из условий теоремы Вейерштрасса наиболее обременительным является ограниченность множества X, часто из постановки задачи она не вытекает, а для задачи на безусловный экстремум заведомо не выполняется. Поэтому полезной является следующая ее модификация.

Следствие. Если X — замкнутое непустое множество в R^n, $f(x)$ — непрерывная на X функция и ограничено множество

$$P = \{ x \in X \mid f(x) \geq c\}, \tag{1.9}$$

где c — некоторая константа, меньшая верхней грани $f(x)$ на X, то точка глобального максимума $f(x)$ на X существует.

Доказательство. Множество P является по определению непустым подмножеством множества X, оно ограничено, а в силу непрерывности $f(x)$ и замкнуто. Поэтому по теореме Вейерштрасса существует точка глобального максимума $f(x)$ на P со значением не меньше c. Но на множестве $X \backslash P$ выполняется $f(x) < c$, следовательно, данная точка является и глобальным максимумом $f(x)$ на X, что и требовалось доказать.

Для существования глобального минимума в определении (1.9) множества P знак неравенства следует заменить на противоположный, а константа c должна быть больше нижней грани $f(x)$ на X.

Для рассматриваемых в дальнейшем экстремальных задач будет предполагаться, как правило, выполнение условий теоремы Вейерштрасса или ее следствия. Условие ограниченности множества Р может заменяться на условие $f(x_k) \to -\infty$ при $|x_k| \to \infty$ (для минимума $f(x_k) \to \infty$).

1.2. Условия экстремума в задачах без ограничений

Задача поиска безусловного экстремума (без ограничений) имеет вид

$$f(x) \to extr,\ x \in R^n. \tag{1.10}$$

Классические необходимые условия экстремума (локального или глобального, максимума или минимума) для этой задачи дает

Теорема Ферма. *Пусть функция $f(x)$ дифференцируема в точке решения задачи* (1.10) x^*, *тогда*

$$\frac{\partial f(x^*)}{\partial x_j} = 0,\ j = \overline{1,n}. \tag{1.11}$$

Доказательство. По определению дифференцируемой функции многих переменных для любого вектора $h \in R^n$ имеем

$$f(x^* + dh) = f(x^*) + d < f'(x^*), h > + 0(d), \tag{1.12}$$

где $f'(x^*) = \left(\frac{\partial f(x^*)}{\partial x_1}, \ldots, \frac{\partial f(x^*)}{\partial x_n} \right)$ — градиент функции $f(x)$ в точке x^*, d — достаточно малое положительное число, $0(d)$ — величина более высокого порядка малости, нежели d, $< f'(x^*), h >$ — скалярное произведение векторов $f'(x^*)$ и h.

Предположим противное утверждению (1.11), т.е. что $|f'(x^*)| \neq 0$. Тогда из (1.12) при достаточно малых d имеем $f(x^*+ dh) > f(x^*)$ при $h = f'(x^*)$ и $f(x^*+ dh) < f(x^*)$ при $h = -f'(x^*)$, что противоречит определению экстремумов.

Условие (1.11) или эквивалентное ему

$$f'(x^*) = 0 \tag{1.13}$$

называется условием первого порядка, т.к. использует первые частные производные. Оно является необходимым, но не достаточным, т.е. ему могут удовлетворять и точки, не являющиеся экстремумами (все эти точки называются стационарными).

Для того чтобы сформулировать необходимые и достаточные условия второго порядка (использующие вторые производные), напомним некоторые алгебраические понятия.

Пусть задана квадратная симметрическая матрица $A = (a_{ij})$ размера $n \times n$. Рассмотрим скалярное произведение $< Ah, h >$, где первым вектором является произведение матрицы A на n-мерный вектор h (мы не будем делать различия в обозначениях между вектор-строкой и вектор-столбцом, считая всегда, что вектор задан в такой форме, которая позволяет производить соответствующую операцию умножения на него матрицы).

Это скалярное произведение равно $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_i h_j$, где $a_{i\,j}$ — элементы матрицы A, h_i — компоненты вектора h, т.е. матрица A задает квадратичную форму. Квадратичная форма $< Ah, h >$ называется положительно определенной, если

$$< Ah, h > \; > 0 \text{ при всех ненулевых } h \in R^n. \qquad (1.14)$$

Аналогично вводятся понятия неотрицательной, неположительной и отрицательной определенности (в (1.14) знак > меняется соответственно на $\geq$, $\leq$, $<$). При этом и матрица A называется иногда положительно (неотрицательно, неположительно, отрицательно) определенной.

Непосредственная проверка положительной определенности, т.е. неравенства (1.14), непроста. Поэтому желательно иметь для нее более конструктивные условия. Наиболее общим является **критерий Сильвестра.** Для положительной определенности матрицы А необходимо и достаточно, чтобы были положительными ее главные миноры

$$a_{11}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ldots, \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Необходимым и достаточным условием отрицательной определенности матрицы A является положительность главных миноров четного порядка и отрицательность главных миноров нечетного порядка, т.е.

$$a_{11} < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0 \text{ и т.д.}$$

Вернемся к задаче поиска безусловного экстремума. Так как условия второго порядка для максимумов и минимумов несколько отличаются, то рассмотрим сначала задачу

$$f(x) \to max, \;\; x \in R^n. \qquad (1.15)$$

Будем считать $f(x)$ дважды дифференцируемой и введем матрицу вторых частных производных

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

Это квадратная симметрическая матрица порядка n.

Теорема (необходимые условия максимума второго порядка). *Пусть функция $f(x)$ дважды дифференцируема в точке решения задачи* (1.15) x^*, *тогда матрица* $f''(x^*)$ *неположительно определена, т.е.* $< f''(x^*)h, h > \le 0$ *для любого* $h \in R^n$.

Доказательство. Используя разложение функции многих переменных в ряд Тейлора до квадратичных членов, имеем

$$f(x^*+dh)=f(x^*)+d<f'(x^*),h>+\frac{1}{2}d^2<f''(x^*)h, h> + 0(d^2). \qquad (1.16)$$

Так как x^* — локальный максимум, то по определению $f(x^*+ dh) \le f(x^*)$ при достаточно малом d и любом $h \in R^n$, а по теореме Ферма $f'(x^*) = 0$, поэтому из (1.16) предельным переходом при $d \to 0$ получаем утверждение теоремы.

Условием минимума является неотрицательная определенность матрицы $f''(x^*)$ (доказывается аналогично).

Теорема (достаточные условия максимума второго порядка). *Пусть функция* $f(x)$ *дважды дифференцируема в точке* x^*, $f'(x^*) = 0$ *и матрица* $f''(x^*)$ *отрицательно определена, тогда* x^* — *решение задачи* (1.15).

Доказательство. Используя разложение (1.16) и учитывая, что линейный член в этом разложении равен нулю, а квадратичный член отрицателен для любого ненулевого h , получаем утверждение теоремы.

Задаче на минимум соответствует условие положительной определенности матрицы $f''(x^*)$. Для функции одной переменной эти условия означают просто отрицательность (положительность) второй производной. Для функции двух переменных с учетом критерия Сильвестра имеем условия максимума

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_1}=\frac{\partial f(x_1^*,x_2^*)}{\partial x_2}=0,\ \frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_1^2}<0,$$

$$\frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_1^2}\cdot\frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_2^2}-\left[\frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_1\partial x_2}\right]^2>0$$

(для минимума $\frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_1^2}>0$, а остальное без изменения).

Пример. $f(x_1, x_2) = x_1x_2 - x_1^2-x_2^2 \to max, x \in R^2$.
Условие первого порядка дает уравнения

$$x_2-2x_1 = 0,\ x_1-2x_2 = 0,$$

откуда $x^*_1= x^*_2 = 0$. Матрица вторых производных $\begin{pmatrix}-2 & 1\\ 1 & -2\end{pmatrix}$ отрицательно определена, значит $(0, 0)$ — локальный максимум.

1.3. Условия экстремума в задачах с ограничениями типа равенств

Задача поиска условного экстремума (с ограничениями), вообще говоря, сложнее задачи на безусловный экстремум. Полученные выше условия экстремума сохраняют силу для внутренних точек множества X, так как для такой точки x^* при достаточно малом d вектор x^*+dh для любого h принадлежит X и все доказательства теорем о необходимых и достаточных условиях проходят без изменений. Однако часто экстремум оказывается на границе множества X, а здесь вектора x^*+dh могут не принадлежать множеству X и полученные результаты уже становятся неприменимыми. Здесь мы рассмотрим классическую задачу на условный экстремум, в которой множество X задается конечным числом ограничений типа равенств

$$X=\{ x \in R^n \mid g_i(x) = 0,\ i=\overline{1,m} \}, \tag{1.17}$$

а более общие задачи будут рассмотрены в последующих главах.

Суть используемого подхода продемонстрируем сначала на частном случае задачи с двумя переменными и одним ограничением

$$f(x_1, x_2) \to extr \ \text{ на } \ X = \{ (x_1, x_2) \in R^2 \mid g(x_1, x_2) = 0\}. \tag{1.18}$$

Пусть из уравнения $g(x_1, x_2) = 0$ мы можем однозначно выразить одну переменную через другую, например $x_2=\varphi(x_1)$. Тогда, подставляя эту функциональную зависимость в функцию $f(x_1, x_2)$, получим задачу поиска безусловного экстремума функции одной переменной $\Psi(x_1) = f(x_1, \varphi(x_1))$. Если функции f и φ дифференцируемы, то по теореме Ферма, используя правило дифференцирования сложной функции, имеем для точки экстремума x^*

$$\frac{d\Psi(x_1^*)}{dx_1} = \frac{\partial f(x_1^*,\varphi(x_1^*))}{\partial x_1} + \frac{\partial f(x_1^*,\varphi(x_1^*))}{\partial x_2}\cdot\frac{d\varphi(x_1^*)}{dx_1} = 0. \tag{1.19}$$

Однако условие (1.19) неудобно, так как найти $\varphi(x_1)$ и ее производную в явном виде не всегда просто. Поэтому желательно выразить все через исходные функции. Так как $x_2=\varphi(x_1)$ эквивалентно $g(x_1, x_2) = 0$, имеем $g(x_1, \varphi(x_1)) \equiv 0$, откуда при условии дифференцируемости функции g

$$\frac{\partial g(x_1,\varphi(x_1))}{\partial x_1} + \frac{\partial g(x_1,\varphi(x_1))}{\partial x_2}\cdot\frac{d\varphi(x_1)}{dx_1} \equiv 0. \tag{1.20}$$

Если $\frac{\partial g(x_1^*,\varphi(x_1^*))}{\partial x_2} \neq 0$, то из (1.20) получаем

$$\frac{d\varphi(x_1^*)}{dx_1} = -\left[\frac{\partial g(x_1^*,\varphi(x_1^*))}{\partial x_1}\right]\left[\frac{\partial g(x_1^*,\varphi(x_1^*))}{\partial x_2}\right]^{-1}.$$

Подставляя это выражение в (1.19), имеем в точке экстремума (x_1^*, x_2^*), где $x_2^* = \varphi(x_1^*)$, условие

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_1} - \frac{\partial g(x_1^*,x_2^*)}{\partial x_1}\cdot\frac{\partial f(x_1^*,x_2^*)}{\partial x_2}\left[\frac{\partial g(x_1^*,x_2^*)}{\partial x_2}\right]^{-1} = 0.$$

Если ввести величину $\lambda^* = -\frac{\partial f(x_1^*,x_2^*)}{\partial x_2}\left[\frac{\partial g(x_1^*,x_2^*)}{\partial x_2}\right]^{-1}$, то последнее условие можно записать в виде двух равенств

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_1} + \lambda^*\frac{\partial g(x_1^*,x_2^*)}{\partial x_1} = 0, \tag{1.21}$$

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_2} + \lambda^*\frac{\partial g(x_1^*,x_2^*)}{\partial x_2} = 0.$$

Введем функцию $L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$, тогда из (1.21) получаем

$$\frac{\partial L(x_1^*,x_2^*,\lambda^*)}{\partial x_1}=\frac{\partial L(x_1^*,x_2^*,\lambda^*)}{\partial x_2}=0. \qquad (1.22)$$

Условие (1.22) означает, что (x_1^*,x_2^*) является стационарной точкой функции L. Правда оно дает два уравнения, а неизвестных три: x_1, x_2, λ. Но у нас имеется еще одно условие $g(x_1^*,x_2^*)=0$, т.е. получаем три уравнения с тремя неизвестными, которые в принципе позволяют найти решение задачи (1.18).

Рассмотрим теперь задачу с ограничениями типа равенств в общем виде (1.8), (1.17). Введем для нее аналогичную функцию

$$L(x,\lambda)=f(x)+\sum_{i=1}^{m}\lambda_i g_i(x), \qquad (1.23)$$

которая называется функцией Лагранжа (здесь λ — m-мерный вектор, его компоненты называются множителями Лагранжа). Естественно считать, что неизвестных больше, чем связывающих их уравнений, т.е. $n>m$.

Теорема (принцип Лагранжа). *Пусть функции $f(x)$, $g_1(x),\ldots,g_m(x)$ непрерывно дифференцируемы в некоторой окрестности точки x^*, являющейся решением задачи* (1.8), (1.17)*, и градиенты $g'_1(x^*),\ldots,g'_m(x^*)$ линейно независимы. Тогда существует вектор λ^* такой, что*

$$\frac{\partial L(x^*,\lambda^*)}{\partial x_j}=0,\ j=\overline{1,n}. \qquad (1.24)$$

Доказательство. Так как градиенты $g'_1(x^*),\ldots,g'_m(x^*)$ линейно независимы, то ранг матрицы

$$\begin{pmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_1(x^*)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_m(x^*)}{\partial x_1} & \cdots & \frac{\partial g_m(x^*)}{\partial x_n} \end{pmatrix}$$

равен m. Значит, хотя бы один определитель m-го порядка, составленный из ее столбцов, отличен от нуля. Без ограничения общности можно считать, что

$$\begin{vmatrix} \frac{\partial g_1(x^*)}{\partial x_1} & \cdots & \frac{\partial g_1(x^*)}{\partial x_m} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_m(x^*)}{\partial x_1} & \cdots & \frac{\partial g_m(x^*)}{\partial x_m} \end{vmatrix}\neq 0. \qquad (1.25)$$

Тогда по известной из анализа теореме о неявной функции в некоторой окрестности точки x^* система равенств $g_i(x)=0,\ i=\overline{1,m}$ эквивалентна системе вида

$$x_1=\varphi_1(x_{m+1},\ldots,x_n),\ldots,x_m=\varphi_m(x_{m+1},\ldots,x_n),$$

т.е. одна часть переменных (в количестве m) однозначно выражается через другую часть переменных (в количестве $n-m$), причем функции φ_k имеют непрерывные частные производные. Подставляя эти функции в f, имеем задачу на безусловный экстремум функции

$$\Psi(x_{m+1},\ldots,x_n)=f(\varphi_1(x_{m+1},\ldots,x_n),\ldots,\varphi_m(x_{m+1},\ldots,x_n),x_{m+1},\ldots,x_n).$$

Очевидно, точка $(x^*_{m+1},...,x^*_n)$ является экстремальной для ψ, поэтому

$$\frac{\partial\Psi(x^*_{m+1},...,x^*_n)}{\partial x_j}=0,\ j=\overline{m+1,n}.$$

Вычисляя производные ψ и приравнивая их к нулю (в точке x^*), имеем

$$\frac{\partial\Psi}{\partial x_j}=\frac{\partial f}{\partial x_j}+\sum_{k=1}^{m}\frac{\partial f}{\partial x_k}\frac{\partial\varphi_k}{\partial x_j}=0,\quad j=\overline{m+1,n}. \tag{1.26}$$

Так как $g_i(\varphi_1(x_{m+1},...,x_n),...,\varphi_m(x_{m+1},...,x_n),x_{m+1},...,x_n)\equiv 0,$

$$\frac{\partial g_i}{\partial x_j}+\sum_{k=1}^{m}\frac{\partial g_i}{\partial x_k}\frac{\partial\varphi_k}{\partial x_j}=0,\quad i=\overline{1,m},\quad j=\overline{m+1,n}. \tag{1.27}$$

Умножим равенства (1.27) на λ_i и сложим с (1.26):

$$\left(\frac{\partial f}{\partial x_j}+\sum_{i=1}^{m}\lambda_i\frac{\partial g_i}{\partial x_j}\right)+\sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}+\sum_{i=1}^{m}\lambda_i\frac{\partial g_i}{\partial x_k}\right)\frac{\partial\varphi_k}{\partial x_j}=0,\quad j=\overline{m+1,n}. \tag{1.28}$$

Рассмотрим систему уравнений

$$\frac{\partial f(x^*)}{\partial x_k}+\sum_{i=1}^{m}\lambda_i\frac{\partial g_i(x^*)}{\partial x_k}=0,\quad k=\overline{1,m}, \tag{1.29}$$

относительно неизвестных λ_i. Система (1.29) имеет отличный от нуля определитель (1.25), значит, существует ее единственное решение $\lambda^*_1,\ldots,\lambda^*_m$. При данных λ^*_i из (1.28) и (1.29) следует

$$\frac{\partial f(x^*)}{\partial x_j}+\sum_{i=1}^{m}\lambda^*_i\frac{\partial g_i(x^*)}{\partial x_j}=0,\quad j=\overline{m+1,n}. \tag{1.30}$$

Но (1.29) при $\lambda_i=\lambda^*_i$ и (1.30) с учетом определения функции Лагранжа эквивалентны (1.24), т.е. теорема доказана.

Принцип Лагранжа дает нам $n+m$ уравнений (n уравнений (1.24) и m уравнений в определении (1.17) множества X) для нахождения $n+m$ неизвестных $x_1,\ldots,x_n,\ \lambda_1,\ldots\lambda_m$, что позволяет в простых случаях найти решение задачи.

Пример. $f(x)=x_1x_2x_3\to extr,$

$$X=\{x\in R^3\mid x_1+x_2+x_3-1=0,\ x_1^2+x_2^2+x_3^2-1=0\}.$$

Градиенты ограничений (1,1,1) и $(2x_1,\ 2x_2,\ 2x_3)$ линейно независимы на X (они зависимы при $x_1=x_2=x_3$, но точки такого вида не принадлежат X). Поэтому принцип Лагранжа справедлив. Функция Лагранжа имеет вид

$$L=x_1x_2x_3+\lambda_1(x_1+x_2+x_3-1)+\lambda_2(x_1^2+x_2^2+x_3^2-1).$$

Имеем пять уравнений

$$x_2x_3+\lambda_1+2\lambda_2x_1=0,\quad x_1x_3+\lambda_1+2\lambda_2x_2=0,$$

$$x_1x_2+\lambda_1+2\lambda_2x_3=0,\quad x_1+x_2+x_3=1,\quad x_1^2+x_2^2+x_3^2=1.$$

Эта система имеет две группы симметричных решений (по три точки):

$$x_i^*=x_j^*=\frac{2}{3},\ x_k^*=-\frac{1}{3},\quad \lambda_1^*=\frac{1}{3},\ \lambda_2^*=-\frac{2}{9};$$

$$x_i^*=x_j^*=0,\quad x_k^*=1,\quad \lambda_1^*=\lambda_2^*=0.$$

Для первой значение функции равно $-\frac{4}{27}$, для второй — 0. Так как для данной задачи выполнены условия теоремы Вейерштрасса, то существуют глобальные максимум и минимум. Поскольку они находятся среди локальных, можно сделать вывод, что

$$\max_{x\in X} f(x) = 0,\ \min_{x\in X} f(x) = -\frac{4}{27},$$

а все три локальных максимума и три локальных минимума являются и глобальными.

Условие линейной независимости градиентов $g'_i(x^*)$ называется условием регулярности. Без него принцип Лагранжа в сформулированном виде может не выполняться.

Пример. $f(x) = x_1 \to extr,\ X = \{ x \in R^2 \mid x_1^3 - x_2^2 = 0 \}$.

Здесь из ограничения $x_1 = x_2^{2/3}$, поэтому, очевидно, $f(x)$ имеет единственный локальный (и глобальный) минимум в (0, 0), а локальных (и глобальных) максимумов нет. Функция Лагранжа $L = x_1 + \lambda(x_1^3 - x_2^2)$, поэтому $\frac{\partial L}{\partial x_1} = 1 + 3\lambda x_1^2 \neq 0$ при $x_1 = 0$, т.е. принцип Лагранжа не выполняется (градиент ограничения в решении равен нулю).

Если градиенты $g'_i(x^*)$ линейно зависимы, то по определению существуют такие λ_i, не все равные нулю, что

$$\sum_{i=1}^{m} \lambda_i g'_i(x^*) = 0. \qquad (1.31)$$

Условие (1.31) также выделяет точки, подозрительные на экстремум. Его можно объединить с (1.24), получив обобщенную форму принципа Лагранжа. Для этого введем обобщенную функцию Лагранжа

$$\tilde{L}(x, \lambda) = \lambda_\circ f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

Здесь λ уже $(m+1)$–мерный вектор. Условия стационарности для функции $\tilde{L}(x,\lambda)$, т.е. $\frac{\partial \tilde{L}}{\partial x_j} = 0$, $j = \overline{1,n}$, при $\lambda_o = 1$ превращаются в (1.24), а при $\lambda_o = 0$ — в (1.31). Таким образом, можно утверждать, что если x^* — решение задачи (1.8), (1.17), то существует ненулевой вектор λ такой, что x^* является стационарной точкой обобщенной функции Лагранжа, причем достаточно ограничиться для $\lambda_\circ$ значениями 0 и 1. В рассматриваемом выше примере (без условия регулярности) обобщенная функция Лагранжа

$$\tilde{L} = \lambda_\circ x_1 + \lambda_1 (x_1^3 - x_2^2)$$

и точка $(0, 0)$ является уже для нее стационарной при $\lambda_\circ = 0,\ \lambda_1 = 1$.

Принцип Лагранжа (обычный и обобщенный) представляет собой необходимое условие экстремума первого порядка в задаче с ограничениями типа равенств. Однако стационарные точки функции Лагранжа не обязаны быть решениями исходной задачи. Привлекая вторые производные, можно получить необходимые и достаточные условия второго порядка, аналогичные условиям для безусловного экстремума. Сформулируем (без доказательства) достаточные условия максимума в задаче с ограничениями: если функции $f(x)$, $g_1(x),\dots, g_m(x)$ дважды дифференцируемы в точке x^* и непрерывно дифференцируем в некоторой ее окрестности, градиенты $g'_1(x^*),\dots, g'_m(x^*)$ линейно независимы, в точке x^* выполнены условия (1.24) при некоторых λ_i^* и, кроме того, для матрицы вторых производных по x функции $L(x, \lambda)$ справедливо $\langle L''_{xx}(x^*,\lambda^*)\, h, h\rangle < 0$ для любого ненулевого $h \in R^n$, удовлетворяющего условиям $\langle g'_i(x^*), h\rangle = 0$, $i = \overline{1,m}$, т.е. ортогонального всем градиентам ограничений в точке x^*, то x^* — решение задачи (1.6), (1.17). Для задачи на минимум указанная квадратичная форма должна быть положительной.

Пример. $f(x) = x_1^2 + \frac{1}{2}x_2^2 \to extr,$

$$X = \{ x \in R^2 \mid x_1^3 + x_2^3 - 1 = 0 \}.$$

Здесь условие регулярности выполняется, т. к. $g'(x) = (3x_1^2, 3x_2^2) \neq 0$ на X. Функция Лагранжа $L = x_1^2 + \frac{1}{2}x_2^2 + \lambda(x_1^3 + x_2^3 - 1)$. Имеем систему из трех уравнений

$$2x_1 + 3\lambda x_1^2 = 0, \quad x_2 + 3\lambda x_2^2 = 0, \quad x_1^3 + x_2^3 = 1.$$

Она имеет три решения:

1) $x_1^* = 1, \quad x_2^* = 0, \quad \lambda^* = -\frac{2}{3}$, 2) $x_1^* = 0, \; x_2^* = 1, \quad \lambda^* = -\frac{1}{3}$,

3) $x_1^* = \frac{2}{\sqrt[3]{9}}, \; x_2^* = \frac{1}{\sqrt[3]{9}}, \; \lambda^* = -\frac{\sqrt[3]{9}}{3}$.

Матрица вторых производных функции Лагранжа есть

$$L''_{xx} = \begin{pmatrix} 2 + 6\lambda x_1 & 0 \\ 0 & 1 + 6\lambda_2 x_2 \end{pmatrix}.$$

Проверим достаточные условия для каждого решения.

1) $L''_{xx} = \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}$, $g' = (3,0)$, если $\langle g', h\rangle = 0$, то h_2 — любое, $h_1=0$, $\langle L''_{xx}\, h,h \rangle = h_2^2 > 0$, значит, $(1, 0)$ — точка локального минимума со значением $f = 1$.

2) $L''_{xx} = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$, $g' = (0,3)$, если $\langle g', h\rangle = 0$, то h_1 — любое, $h_2 = 0$, $\langle L''_{xx}h, h\rangle = 2h_1^2 > 0$, значит, (0,1) — точка локального минимума со значением $f = \frac{1}{2}$.

3) $L''_{xx} = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$, $<L''_{xx}\, h,h> \; = -2h_1^2 - h_2^2 < 0$ для любого $h \neq 0$, значит, $\left(\frac{2}{\sqrt[3]{9}}, \frac{1}{\sqrt[3]{9}}\right)$ — точка локального максимума со значением $f = \frac{\sqrt[3]{9}}{2}$.

Далее мы рассмотрим ряд задач из различных разделов математики, которые могут быть решены с помощью классических условий экстремума.

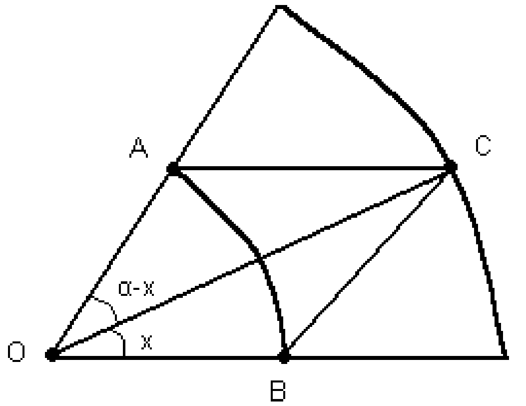
1.4. Использование классических методов поиска экстремума в математических проблемах

В различных разделах математики возникали и возникают проблемы, связанные с решением экстремальных задач. Естественно, что для этих целей могут быть использованы изложенные выше методы. Однако область их применения не ограничивается только такими задачами. Часто проблема в первоначальном виде не связана явно с какой-либо экстремальной задачей, но может быть переформулирована так, что для ее решения оказываются пригодными методы поиска экстремума. Рассмотрим некоторые примеры из разных областей математики.

Геометрия. В геометрии традиционно много экстремальных задач связано с построением фигур с заданными параметрами (вписанных, описанных и т.п.), обладающих наибольшими (наименьшими) объемами (площадями, периметрами и т.д.). Такие задачи могут быть решены чисто геометрическими методами, однако, как правило, при этом требуется проявить догадку и провести какое-то дополнительное построение. Способности и опыт позволяют преодолевать трудности, но часто это связано с затратами времени и труда. В то же время, применение общих методов поиска экстремума позволяет стандартизировать процесс решения и свести все к аналитическим преобразованиям (найденному таким путем решению можно затем дать геометрическую интерпретацию). Рассмотрим для примера следующую задачу.

Задача о двух орбитах

Имеются две концентрические окружности. На окружности меньшего радиуса заданы две точки. Требуется на окружности большего радиуса найти точку, сумма расстояний от которой до заданных точек является минимальной.



Обозначим радиусы окружностей r и R $(R>r)$, угол между радиусами OA и OB (O — центр окружностей, A и B — заданные точки) обозначим α $(0< \alpha <\pi)$. Тогда по теореме косинусов для произвольной точки С на большей окружности сумма расстояний от С до А и В равна

$$\rho = \sqrt{R^2 + r^2 - 2Rr\cos x} + \sqrt{R^2 + r^2 - 2Rr\cos(\alpha - x)},$$

где в качестве неизвестного x взят угол между радиусами OB и OC $(0 \le x \le \alpha)$.

Необходимое условие минимума $\frac{d\rho}{dx} = 0$ дает нам уравнение

$$\frac{\sin^2 x}{R^2 + r^2 - 2Rr\cos x} = \frac{\sin^2(\alpha - x)}{R^2 + r^2 - 2Rr\cos(\alpha - x)}. \tag{1.32}$$

Очевидно, $x = \frac{\alpha}{2}$ является корнем уравнения (1.32), но он может быть не единственным. Из (1.32) имеем

$$(R^2 + r^2)(\sin^2 x - \sin^2(\alpha - x)) = 2Rr(\sin^2 x\cos(\alpha - x) - \cos x\sin^2(\alpha - x))$$

или

$$(R^2 + r^2)(\cos^2(\alpha - x) - \cos^2 x) = 2Rr(\cos(\alpha - x) - \cos^2 x\cos(\alpha - x) -$$
$$- \cos x + \cos x\cos^2(\alpha - x)).$$

Сокращая на разность $\cos(\alpha - x) - \cos x$, не равную нулю при $x \ne \frac{\alpha}{2}$, имеем

$$(R^2 + r^2)(\cos(\alpha - x) + \cos x) = 2Rr(1 + \cos x\cos(\alpha - x)),$$

откуда, используя тригонометрические формулы, получаем

$$(R^2 + r^2)\cos\frac{\alpha}{2}\cos\left(\frac{\alpha}{2} - x\right) = Rr\left(\cos^2\frac{\alpha}{2} + \cos^2\left(\frac{\alpha}{2} - x\right)\right).$$

Решая это квадратное уравнение относительно $\cos(\frac{\alpha}{2} - x)$, получим

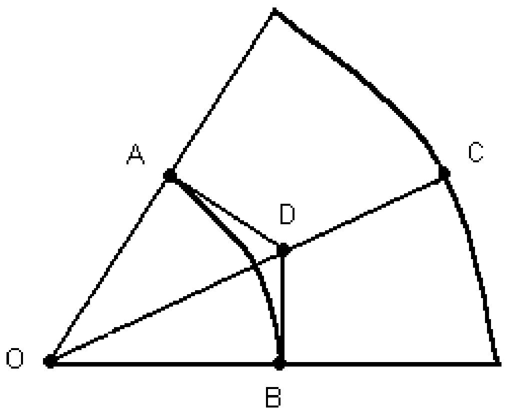
$x = \frac{\alpha}{2} \pm \arccos\left(\frac{R}{r}\cos\frac{\alpha}{2}\right)$ при $\cos\frac{\alpha}{2} < \frac{r}{R}$ (при $\cos\frac{\alpha}{2} > \frac{r}{R}$ уравнение не имеет решения, при $\cos\frac{\alpha}{2} = \frac{r}{R}$ имеем $x = \frac{\alpha}{2}$).

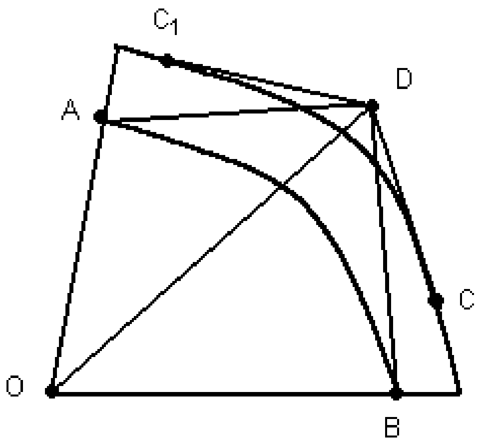
Для нахождения абсолютного минимума на отрезке $[o, \alpha]$ надо сравнить значение целевой функции ρ в точках $0, \frac{\alpha}{2}, \alpha$ при $\cos\frac{\alpha}{2} \ge \frac{r}{R}$ и дополнительно в точках $\frac{\alpha}{2} \pm \arccos\left(\frac{R}{r}\cos\frac{\alpha}{2}\right)$ при $\cos\frac{\alpha}{2} < \frac{r}{R}$.

Легко проверить, что точка абсолютного минимума есть

$$x^* = \begin{cases} \dfrac{\alpha}{2}, & \cos\dfrac{\alpha}{2} \ge \dfrac{r}{R}, \\ \dfrac{\alpha}{2} \pm \arccos\left(\dfrac{R}{r}\cos\dfrac{\alpha}{2}\right), & \cos\dfrac{\alpha}{2} < \dfrac{r}{R}. \end{cases}$$

Нетрудно убедиться, что это решение получается следующим геометрическим построением. Строим касательные к меньшей окружности в точках *A* и *B*, если точка их пересечения *D* принадлежит кругу с радиусом *R*, то искомая точка C лежит на биссектрисе угла *AOB*, если точка *D* лежит вне данного круга, то проводим через нее касательные к большей окружности, получающиеся точки касания C_1 и C_2 являются решением задачи (см. рисунки). При найденном x^* это достаточно очевидно, однако догадаться заранее непросто.

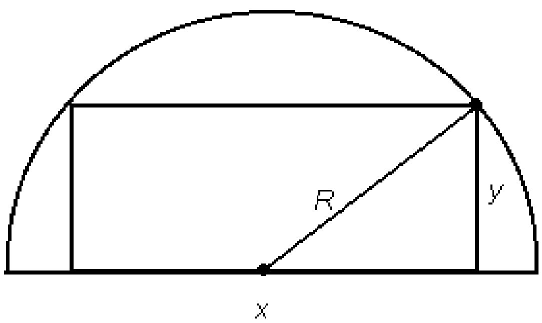




Рассмотрим еще пример геометрической задачи, в которой приходится находить экстремум функции при ограничениях.

Задача о вписанном прямоугольнике

Найти прямоугольник набольшей площади, две вершины которого лежат на диаметре окружности, а две другие — на самой окружности.



Обозначим радиус окружности через *R*, неизвестные длины сторон — через *x* и *y*. Тогда целевая функция есть S = *xy*, причем на *x* и *y* накладывается ограничение $\dfrac{x^2}{4} + y^2 - R^2 = 0$.

Используем принцип Лагранжа. Так как градиент ограничения $(\frac{x}{2}, 2y) \ne 0$, то условие регулярности выполнено. Функция Лагранжа имеет вид

$$L = xy + \lambda\left(\frac{x^2}{4} + y^2 - R^2\right).$$

Имеем систему из трех уравнений:

$$y + \frac{\lambda x}{2} = 0,\ x + 2\lambda y = 0,\ \frac{x^2}{4} + y^2 = R^2.$$

Решение этой системы:

$$x^* = R\sqrt{2},\ y^* = \frac{R\sqrt{2}}{2},\ \lambda^* = -1.$$

Очевидно, точка $(R\sqrt{2},\ \frac{R\sqrt{2}}{2})$ является глобальным максимумом целевой функции (соответствующее ее значение $S^* = R^2$). Таким образом, одна сторона прямоугольника должна быть вдвое больше другой.

Алгебра и анализ. Широкая область применения методов поиска экстремума связана с доказательством неравенств и получением различных оценок. Рассмотрим несколько примеров.

Квадратичные формы

Мы приводили без доказательств один критерий (Сильвестра) положительный (отрицательный) определенности квадратичных форм. Здесь выведем другой критерий. Для этого рассмотрим задачу на условный экстремум квадратичной формы $f(x) = \langle Ax, x \rangle$ (A — симметрическая матрица) на единичной сфере $1 - |x|^2 = 0$. Задача регулярна, функция Лагранжа есть $L = \langle Ax, x \rangle + \lambda - \lambda |x|^2$. Градиент функции Лагранжа L по переменным x равен

$$L'_x = 2(Ax - \lambda x),$$

поэтому для нахождения решений имеем систему уравнений

$$Ax = \lambda x,\ |x| = 1.$$

Таким образом, стационарными точками функции Лагранжа являются собственные вектора матрицы A, а значения функции $f(x)$ в этих точках равны собственным числам матрицы A (как известно, все собственные числа симметрической матрицы являются действительными числами). Перенумеруем эти числа в порядке роста: $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ (их количество, вообще говоря, равно порядку матрицы n, но некоторые могут совпадать).

По теореме Вейерштрасса у функции $f(x)$ на сфере существуют глобальные максимум и минимум, а так как они являются стационарными точками функции Лагранжа, то очевидно,

$$\min_{|x|=1} f(x) = \lambda_1, \ \max_{|x|=1} f(x) = \lambda_n.$$

Отсюда при произвольном x имеем оценки для квадратичной формы

$$\lambda_1 \ |x|^2 \le \langle Ax, x \rangle \le \lambda_n |x|^2 .$$

Получаем следующие критерии: квадратичная форма положительно (неотрицательно) определена тогда и только тогда, когда минимальное собственное число матрицы A положительно (неотрицательно); квадратичная форма отрицательно (неположительно) определена тогда и только тогда, когда максимальное собственное число матрицы A отрицательно (неположительно).

Неравенство между средним арифметическим и средним квадратическим

Средним арифметическим чисел $x_1, \ldots, x_n$ называется, как известно, величина $\frac{1}{n}\sum_{i=1}^{n} x_i$, а средним квадратическим — $\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 \right)$. Для доказательства неравенства

$$\frac{1}{n}\sum_{i=1}^{n} x_i \le \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$$

рассмотрим задачу поиска максимума функции $f(x) = \sum_{i=1}^{n} x_i$ при ограничении $\sum_{i=1}^{n} x_i^2 - B = 0$. Эта задача регулярна и в соответствии с принципом Лагранжа ее решение удовлетворяет системе уравнений

$$1 + 2\lambda x_i = 0, \quad i = \overline{1,n}, \ \sum_{i=1}^{n} x_i^2 = B.$$

Отсюда все x_i^* равны между собой и $x_i^* = \pm\sqrt{\frac{B}{n}}$. Точка максимума $x_i^* = \sqrt{\frac{B}{n}}$, (другая точка — минимум). Значит

$$\frac{1}{n} f(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \le \sqrt{\frac{B}{n}} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}},$$

что и требовалось доказать (заметим, что неравенство превращается в равенство только в случае, когда все x_i равны между собой).

На этой задаче мы продемонстрировали общий прием для доказательства неравенств: одна часть неравенства приравнивается к произвольной величине B и это берется в качестве ограничения, а другая часть неравенства выступает в роли целевой функции, для которой в зависимости от знака неравенства ищется максимум или минимум, найденное экстремальное значение соответственно не больше или не меньше B. Таким стандартным приемом можно доказать все известные неравенства (например, Коши-Буняковского, Гельдера, со всевозможными средними), а также выводить новые неравенства.

Информатика. Здесь также существует множество задач, для решения которых могут быть использованы методы поиска экстремума. Типичные примеры связаны с обработкой данных.

Метод наименьших квадратов

Имеются результаты измерений значений параметров на входе x и на выходе y некоторого устройства:

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n).$$

По этим данным нужно восстановить функциональную зависимость $y = f(x)$, описывающую работу устройства. Уточним постановку задачи. Пусть априори известен класс возможных функциональных зависимостей $f(x, d_1, \ldots, d_k)$ с точностью до параметров $d_1, \ldots, d_k$. По методу наименьших квадратов конкретное значение вектора параметров $d = (d_1, \ldots, d_k)$ выбирается так, чтобы минимизировать сумму квадратов невязок в измерениях, т.е. величину

$$\delta(d) = \sum_{i=1}^{n} [\, y_i - f(x_i, d_1, \ldots, d_k) \,]^2.$$

Рассмотрим для простоты линейную зависимость

$$f(x, a, b) = ax + b.$$

Тогда

$$\delta(a, b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

и надо найти абсолютный минимум этой величины как функции аргументов a и b. Условие экстремума

$$\frac{\partial \delta}{\partial a} = \frac{\partial \delta}{\partial b} = 0$$

дает систему из двух линейных уравнений

$$\begin{cases} a\left(\sum_{i=1}^{n} x_i^2\right) + b\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} x_i y_i, \\ a\left(\sum_{i=1}^{n} x_i\right) + bn = \sum_{i=1}^{n} y_i. \end{cases},$$

Определитель этой системы $D = n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2$. Естественно считать, что не все x_i равны между собой, поэтому из доказанного неравенства между средним арифметическим и средним квадратическим имеем D>0. Значит, система имеет единственное решение

$$a^* = D^{-1}\left[n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\right],$$

$$b^* = D^{-1}\left[\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)\right].$$

Матрица вторых производных δ есть

$$\delta'' = \begin{pmatrix} 2\sum_{i=1}^{n} x_i^2 & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2n \end{pmatrix}.$$

Ее главный минор первого порядка равен $2\sum_{i=1}^{n} x_i^2$, а главный минор второго порядка равен $4D$, т.е. они положительны (если не все x_i равны между собой). Значит, по критерию Сильвестра матрица δ'' положительно определена и в силу теоремы о достаточных условиях экстремума точка (a^*, b^*) доставляет минимум величине δ (по крайней мере локальный).

Но так как выполнены условия следствия из теоремы Вейерштрасса, существует глобальный минимум δ, а в силу единственности экстремума пара (a^*, b^*) является точкой глобального минимума δ.

Мера информации

В теории информации количество информации в сообщении о значении величины, которая может принимать априори n значений с вероятностями $p_1,\dots,p_n$, принято определять величиной

$$\mathrm{H} = -\sum_{i=1}^{n} p_i \log_2 p_i .$$

Для этого имеется ряд оснований, связанных со свойствами функции H (обычно эти свойства постулируются как аксиомы и однозначно определяют вид H). В частности, количество информации в сообщении о значении величины, принимающей априори два значения с равными вероятностями, равно единице (оно и принимается за единицу измерения информации, называемую битом). Для представления единицы информации достаточно одного разряда ячейки памяти машины, работающей с двоичными числами. Поставим вопрос: сколько двоичных разрядов достаточно для представления информации о величине, принимающей n возможных значений с неизвестными вероятностями. Другими словами, какое максимальное количество информации может содержать сообщение о величине с n возможными значениями. Для ответа на этот вопрос надо найти максимум функции H при ограничении $\sum_{i=1}^{n} p_i - 1 = 0$ (вообще говоря, еще надо учитывать, что $p_i \geq 0,\quad i = \overline{1,n}$). Функция H не определена при $p_i = 0$, поэтому доопределим ее по непрерывности при $p_i \to 0$. Используя правило Лопиталя, нетрудно получить $\lim_{p_i \to 0}(p_i \log_2 p_i) = 0$, поэтому при $p_{i_0} = 1, p_i = 0, i \neq i_0$ полагаем $H = 0$ (так как $H \geq 0$, то это минимальное значение). Для нахождения максимума составим функцию Лагранжа

$$L = -\sum_{i=1}^{n} p_i \log_2 p_i + \lambda\left(\sum_{i=1}^{n} p_i - 1\right).$$

Получим систему уравнений

$$\frac{\partial L}{\partial p_i} = -\log_2 p_i - \log_2 e + \lambda = 0,\ i = \overline{1,n},\ \ \sum_{i=1}^{n} p_i = 1,$$

откуда видно, что все p_i равны между собой, т.е.

$$p_i^* = \frac{1}{n},\quad i = \overline{1,n}.$$

Для нахождения глобального максимума, который существует в силу теоремы Вейерштрасса, надо это решение сравнить с граничными точками $p_i=0$, но на границах $H = 0$.

Значит, очевидно, точка $\left(\frac{1}{n},\ldots,\frac{1}{n}\right)$ доставляет глобальный максимум H со значением $\log_2 n$. Это и есть искомое максимальное количество информации. Например, при $n=8$ требуется три разряда.