Санкт-Петербургский государственный университет

Кафедра математического анализа

Замена переменных и симметризация в задачах об экстремуме

методические указания

О. Л. Семенова, Т. П. Дубова

САНКТ-ПЕТЕРБУРГ

2024

0. Аннотация

Исследование на экстремум (условный или безусловный) функции нескольких переменных, как и исследование на экстремум функции одной переменной, обычно производится в два этапа. На первом этапе отыскиваются подозрительные на экстремум точки, для этого обычно используется необходимое условие соответствующего варианта экстремума. На втором этапе производится классификация подозрительных точек. При классификации подозрительных точек обычно стараются избегать применения достаточного условия условного экстремума через знак второго дифференциала функции Лагранжа, потому что этот способ решения в общей ситуации приводит к большому, по сравнению с альтернативными методами, объему вычислений. К альтернативным методам, применимым в тех или иных ситуациях, можно отнести достаточное условие глобального экстремума на компакте (следствие теоремы Вейерштрасса, см. [2]) и оценки значений функций, вытекающие из неравенств между средними, или полученными еще как-либо. В [2] мы подробно останавливались на этих методах, и там приведен ряд примеров решения задач с их помощью. Однако альтернативные методы классификации имеют ограниченную область применения. Кроме того, в случае нулевых смешанных частных производных второго порядка вычисления в ходе исследования знака второго дифференциала заметно сокращаются — настолько, что некоторые задачи о (безусловном) экстремуме оказывается выгодным свести к задачам об условном экстремуме и далее использовать именно достаточное условие.

В данном пособии основное внимание мы уделяем тем задачам, в которых применение достаточного условия условного экстремума оправданно. Также мы рассматриваем задачи, в которых в том или ином виде присутствует симметрия, а в других задачах симметризация достигается за счет подходящей замены переменных. Зачастую именно в симметричных задачах достаточное условие проявляет себя эффективным средством. Мы приводим некоторые утверждения, обосновывающие замену переменных при исследовании на экстремум, и доказываем достаточное условие условного экстремума.

Наше пособие адресовано, в первую очередь, студентам специальностей «фундаментальная математика», «технологии программирования», «фундаментальная механика», «механика и математическое моделирование» математико-механического факультета СПбГУ.

1. Терминология задач об экстремуме и об условном экстремуме. Замена переменных при исследовании на экстремум

Пусть $f : X \mapsto \mathbb{R}$ — функция, заданная в топологическом пространстве X. Мы называем ***точкой (локального) минимума*** функции f произвольную точку $a \in X$, для которой существует такая окрестность U, что для всякого $x \in U$ выполняется неравенство $f(x) \geq f(a)$. Читатель, знакомый с дифференциальным исчислением функции одной переменной, легко модифицирует это определение на случай точки максимума, точки экстремума, а также для строгих вариантов этих понятий.

Если исследуется функция $f : E \mapsto \mathbb{R}$, заданная на подмножестве E пространства $\mathbb{R}^n$, то в этой ситуации применяется предыдущее определение, где в качестве топологического пространства выступает множество E с индуцированной из $\mathbb{R}^n$ топологией (подробнее: $a \in E$ есть точка минимума функции f тогда и только тогда, когда существует такая окрестность U этой точки, что для всякого $x \in U \cap E$ справедливо $f(x) \geq f(a)$). Таким образом, в отличие от некоторых источников (например, [1]) нами, вообще говоря, не предполагается, что точка экстремума — это обязательно внутренняя относительно $\mathbb{R}^n$ точка множества E.

Первое утверждение показывает, что непрерывная замена переменных не приводит к потере точки экстремума в новых переменных.

Утверждение 1. *Пусть X и Y — топологические пространства, $f : Y \mapsto \mathbb{R}$ — произвольная функция, а $\Phi : X \mapsto Y$ — непрерывное отображение. Если точка $y = \Phi(x)$ есть точка минимума (максимума) для функции f, то x есть точка минимума (максимума) для $f \circ \Phi$.*

Это утверждение, как и его первое следствие, вытекает из определения экстремума.

Следствие 1. *Пусть X и Y — топологические пространства, $f : Y \mapsto \mathbb{R}$ — произвольная функция, $\Phi : X \mapsto Y$ — локальный гомеоморфизм, $y = \Phi(x)$. Тогда y есть точка минимума (максимума) функции f, если и только если x есть точка минимума (максимума) для композиции $f \circ \Phi$.*

Заметим, что утверждение 1 нарушается в отношении строгих экстремумов: если точка $y = \Phi(x)$ есть точка строгого минимума (максимума) для функции f, $y = \Phi(x)$, то x не обязательно точка именно строго минимума (максимума) для $f \circ \Phi$ (например, для постоянного отображения $\Phi \equiv y$). А вот в условиях следствия 1 (случае локального гомеоморфизма) прообразы строгих экстремумов — это обязательно именно строгие экстремумы.

Все рассматриваемые нами в этом пособии примеры относятся к ситуации функции, заданной или исследуемой на некотором подмножестве E пространства $\mathbb{R}^n$ (второй случай может быть сведен к первому, так как можно рассматривать сужение исходной функции на E). Множество E зачастую задается как множество решений некоторой системы неравенств

$$F_1(x) \geq 0, F_2(x) \geq 0, \ldots, F_m(x) \geq 0, \tag{1}$$

или, как частный случай, — некоторой системы равенств

$$F_1(x) = 0, F_2(x) = 0, \ldots, F_m(x) = 0, \tag{2}$$

или в векторной форме

$$F(x) = \mathbf{0}_m, \tag{3}$$

здесь $F_1, F_2, \ldots, F_m$ — функции переменной $x = (x_1, \ldots, x_n)$, в этом контексте часто называемые ***функциями связи***, и $F = (F_1, F_2, \ldots, F_m)$ — ***отображение связи***.

Предположим, что точка $a \in E$ является точкой экстремума (минимума или максимума, строгого минимума или строгого максимума) сужения функции f на множество E. Тогда точка a называется точкой ***условного*** (или относительного) ***экстремума*** функции f соответствующего типа ***при условиях связи*** (3).

Следующее следствие утверждения 1 типично применимо в ситуации наличия какой-то формы симметрии — как у исследуемой функции, так и у множества, на котором функция исследуется. Например, если исследуемая функция и все функции, задающие множество (системой равенств или системой неравенств), четны относительно какой-либо из переменных; или принимают равные значения в точках, симметричных относительно начала координат (или какой-то другой точки); или симметричным образом зависят от какой-либо пары переменных; и т.п. Уточним понятия.

Для произвольной биекции U в $\mathbb{R}^n$ называем U-***инвариантным***
1) множество $E \subseteq \mathbb{R}^n$, если $U(E) = E$,
2) функцию f, если $f \circ U \equiv f$, то есть области определения функций f и $f \circ U$ совпадают, и в каждой точке области определения совпадают их значения.

Замечание 1. Как для множества E, так и для функции f их U-инвариантность равносильна U^{-1}-инвариантности (поскольку $f = f \circ U \circ U^{-1} = f \circ U^{-1}$ и области определения также совпадают).

Замечание 2. Если множество E задается системой неравенств (1) (или системой равенств (3)) и все функции $F_1, F_2, \ldots, F_m$ U-инвариантны, то E также U-инвариантно.

Так, к примеру, если все функции $f, F_1, F_2, \ldots, F_m$ четны относительно какой-либо из переменных x_i, то функция и множество инвариантны к отражению относительно гиперплоскости $\{x_i = 0\}$.

Следствие 2. *Пусть функция f и подмножество E ее области определения инвариантны относительно некоторой изометрии U пространства $\mathbb{R}^n$. Тогда для всякой точки x_* экстремума (минимума, строгого минимума, и т.п.) сужения f на E при любом целом k точка $U^k(x_*)$ (образ точки x_* под действием k-й степени относительно композиции изометрии U) является для f точкой экстремума того же типа, что и x_*.*

Замечание 3. Пусть в условиях U-инвариантности множества E, на котором исследуется функция f, эта функция «антиинвариантна» относительно U, то есть $f \equiv -f \circ U$. Тогда точки x_* и $U(x_*)$ также являются точками условного экстремума f на E одновременно, но тогда при наличии экстремума эти экстремумы противоположных типов (если одна — точка минимума, то другая — точка максимума).

Пример 1. Исследовать на условный экстремум функцию $f(x, y) = xy$ при условии $x^3 + y^3 - 3axy = 0$, где $a > 0$.

Решение. Кривая, заданная уравнением связи $\{(x, y) \ : \ x^3 + y^3 - 3axy = 0\}$ данной задачи, известна под названием декартов лист. Определение в качестве параметра отношение переменных $t = y/x$ позволяет получить одну из параметризаций кривой:

$$x(t) = \frac{3at}{t^3 + 1}, \quad y(t) = \frac{3at^2}{t^3 + 1}, \quad t \in D := \mathbb{R} \setminus \{-1\}. \tag{4}$$

Легко видеть, что для функций $x(t)$ и $y(t)$, определяемых равенствами (4), справедливы тождества

$$x\,(1/t) \equiv y(t), \quad y\,(1/t) \equiv x(t), \tag{5}$$

то есть при любом допустимом t точки $(x(t), y(t))$ и $(x(1/t), y(1/t))$ расположены симметрично относительно прямой $y = x$. Это означает, что декартов лист состоит из

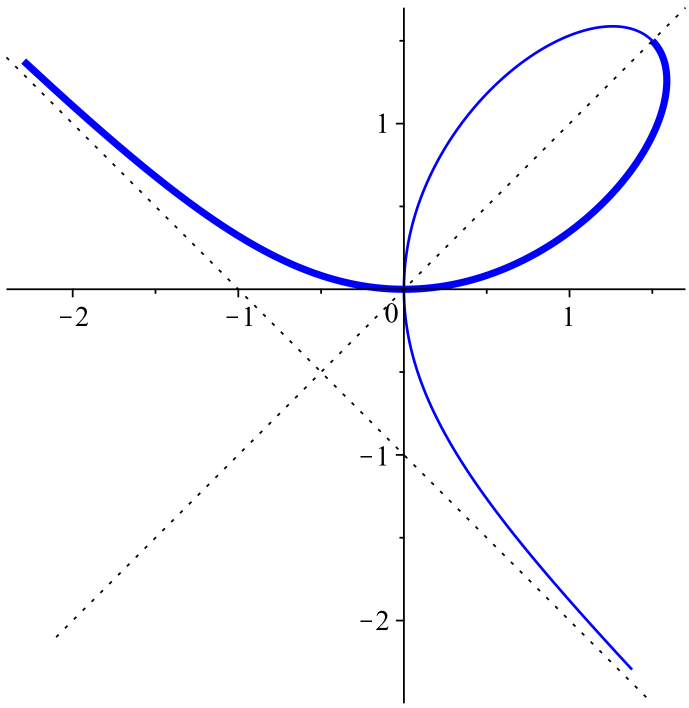

Рис. 1. Декартов лист с осью симметрии $y = x$ и асимптотой $x+y = -1$ при $a = 1$. Жирным выделена «половина» кривой Δ, отвечающая промежутку $(-1; 1]$ при параметризации (4).

двух гладких «половин», симметричных относительно прямой $\{y = x\}$, из которых одна, обозначим ее Δ, отвечает множеству $\{t \in D \ : \ |t| \le 1\}$, то есть промежутку $(-1; 1]$, при параметризации (4), а вторая — множеству $\{t \in D \ : \ |t| \ge 1\}$ (см. рис. 1). Поскольку исследуемая функция f, также как и функция связи, симметрично зависит от переменных x и y, достаточно исследовать ее на экстремум на Δ. В силу следствия 1 утверждения 1 задача сводится к исследованию функции $\widetilde{f}(t) = f(x(t), y(t)) = x(t) \cdot y(t) = \frac{3at}{t^3+1} \cdot \frac{3at^2}{t^3+1} = \frac{9a^2t^3}{(t^3+1)^2}$ на промежутке $(-1; 1]$. Эту задачу удобно решать, сделав еще одну замену $s = t^3$ (она сохраняет промежуток $(-1; 1]$), тогда $\frac{t^3}{(t^3+1)^2} = \frac{s}{(s+1)^2}$. Легко проверить, что дробь $\frac{s}{(s+1)^2}$ строго возрастает на $(-1; 1]$, то есть единственная точка экстремума этой функции — концевая точка $s = 1$, и она является точкой глобального максимума. Тогда у исходной функции $f(x, y)$ также имеется единственная точка условного экстремума — точка глобального максимума $(x(t(s)), y(t(s)))|_{s=1} = \left(\frac{3a}{2}, \frac{3a}{2}\right)$.

Ответ: $\left(\frac{3a}{2}, \frac{3a}{2}\right)$ — точка строгого условного максимума.

2. Необходимое условие условного экстремума (метод множителей Лагранжа)

В [2] мы довольно подробно рассмотрели необходимое условие условного экстремума, приведя его графические иллюстрации и несколько примеров использования. Поэтому здесь мы довольно сжато излагаем две формулировки необходимого условия (геометрическую и с помощью функции Лагранжа). А затем рассматриваем пример, где исследование на условный экстремум производится с помощью данного условия и замены переменных.

Необходимое условие условного экстремума (геометрическая формулировка). *Пусть $m, n \in \mathbb{N}$, $m < n$, множество O открыто в $\mathbb{R}^n$, $f \in C^1(O)$, $F \in C^1(O \mapsto \mathbb{R}^m)$, $\operatorname{rang} F'(a) = m$ (то есть ранг в точке a максимален). Если точка a является точкой условного экстремума функции f при условии (3), то в точке a градиент функции f представим в виде линейной комбинации градиентов функций $F_1, \ldots, F_m$. Иначе говоря, существуют такие вещественные числа $\lambda_1, \ldots, \lambda_m$, что*

$$\nabla_a f = \sum_{i=1}^{m} \lambda_i \cdot \nabla_a F_i. \tag{6}$$

(Доказательство необходимого условия условного экстремума можно найти, например, в [4].

На практике часто используют формулировку необходимого условия, связанную с понятием функции Лагранжа.

Функцией Лагранжа (отвечающей функции $f(x)$ и системе связи (3)) называют функцию

$$L(x) = f(x) - \sum_{i=1}^{m} \lambda_i \cdot F_i(x). \tag{7}$$

Необходимое условие условного экстремума как условие стационарности функции Лагранжа. *Пусть $m, n \in \mathbb{N}$, $m < n$, множество O открыто в $\mathbb{R}^n$, $f \in C^1(O)$, $F \in C^1(O \mapsto \mathbb{R}^m)$, $\operatorname{rang} F'(a) = m$ (максимален). Если точка a является точкой условного экстремума функции f при условии (3), то найдется вектор $\lambda \in \mathbb{R}^m$, для которого справедлива следующая система уравнений Лагранжа относительно точки a:*

$$\begin{cases} \nabla_a L = \mathbf{0}_n, \\ F(a) = \mathbf{0}_m. \end{cases} \tag{8}$$

Система (8) состоит из двух условий: условия стационарности функции Лагранжа L и условия связи. В виде системы $n + m$ скалярных уравнений эта система записывается так:

$$\begin{cases} \frac{\partial f}{\partial x_1}(a) - \sum\limits_{i=1}^{m} \lambda_i \frac{\partial F_i}{\partial x_1}(a) = 0 \left(= \frac{\partial L}{\partial x_1}(a) \right), \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{\partial f}{\partial x_n}(a) - \sum\limits_{i=1}^{m} \lambda_i \frac{\partial F_i}{\partial x_n}(a) = 0 \left(= \frac{\partial L}{\partial x_n}(a) \right), \\ F_1(a) = 0, \\ \ldots\ldots\ldots \\ F_m(a) = 0. \end{cases} \tag{9}$$

Конечно, не всякая точка, удовлетворяющая необходимому условию условного экстремума, является точкой экстремума. Точки, в которых выполняется необходимое условие экстремума (в данном случае, условного), но не являющиеся точками экстремума, называют ***седловыми*** точками.

Следующий пример является незначительной модификацией задачи из [4] (см. III §3).

Пример 2. Пусть R — наибольшее или наименьшее расстояние от начала координат до точки поверхности

$$\left(\sum_{k=1}^{n} x_k^2\right)^2 = \sum_{k=1}^{n} a_k^2 x_k^2 > 0, \tag{10}$$

$$\sum_{k=1}^{n} A_k x_k = 0, \tag{11}$$

где a_k, A_k для $k = 1, \ldots, n$ — отличные от нуля параметры, и модули $|a_1|, \ldots, |a_n|$ попарно различны. Доказать равенство:

$$\sum_{k=1}^{n} \frac{A_k^2}{R^2 - a_k^2} = 0. \tag{12}$$

Доказательство. Речь идет о глобально экстремальных значениях функции $f(x) = \sum_{k=1}^{n} x_k^2 = r^2$, где $x = (x_1, \ldots, x_n)$, и $r = \sqrt{\sum_{k=1}^{n} x_k^2}$, при условиях (10), (11). Отметим, что по условию задачи r не ноль, и равенство (10) допускает сокращение на r^2. Удобно ввести в рассмотрение следующие переменные:

$$u_k = \frac{x_k}{r}, \quad \text{для} \quad k = 1, \ldots, n, \tag{13}$$

после чего оказывается, что в точках поверхности (10) исследуемая функция $f = r^2$ удовлетворяет равенству

$$r^2 = \sum_{k=1}^{n} a_k^2 u_k^2, \tag{14}$$

Относительно новых переменных набор уравнений связи приобретает вид:

$$\sum_{k=1}^{n} u_k^2 = 1, \tag{15}$$

$$\sum_{k=1}^{n} A_k u_k = 0 \tag{16}$$

(всякая точка x, удовлетворяющая системе (10) и (11), определяет точку u, удовлетворяющую системе (15) и (16), и обратно — если r определяется равенством (14)).

Таким образом, в новых переменных задача нахождения экстремальных значений величины r^2 превращается в задачу о нахождении экстремальных значений квадратичной формы $r^2(u) = \sum_{k=1}^{n} a_k^2 u_k^2$ на $(n-2)$-мерной сфере, заданной системой уравнений (15) и (16) в пространстве $\mathbb{R}^n$. Если бы заданное уравнением (16) подпространство пространства $\mathbb{R}^n$ было инвариантно относительно матрицы этой формы, то можно было бы рассмотреть её сужение на данное подпространство и воспользоваться известным решением об экстремумах квадратичной формы на сфере в $\mathbb{R}^{n-2}$, и тогда экстремальными значениями r^2 являлись бы некоторые собственные числа ее матрицы, то есть какие-то из чисел $a_1^2, \ldots, a_n^2$. Однако, как мы далее покажем, в заданных условиях (попарной различности модулей параметров $|a_1|, \ldots, |a_n|$) ни одно из собственных чисел не совпадает с экстремальным значением R^2 квадратичной формы $r^2 = \sum_{k=1}^{n} a_k^2 u_k^2$.

Для решения задачи в общем случае сначала отметим, что на заданной системой (15) и (16) сфере отсутствуют точки нарушения регулярности. Действительно, в таких точках u градиенты функций связи, то есть векторы $2u$ и A, должны быть линейно зависимы, следовательно, коллинеарны. Такие u не удовлетворяют системе уравнений связи (15) и (16), поскольку все числа $A_1, \dots, A_n$ отличны от нуля.

Далее рассмотрим функцию Лагранжа с двумя параметрами λ и μ:

$$L(u) = \sum_{k=1}^{n} a_k^2 u_k^2 - \lambda \left(\sum_{k=1}^{n} u_k^2 - 1 \right) - \mu \sum_{k=1}^{n} A_k u_k,$$

и получим систему уравнений Лагранжа, добавив к системе

$$0 = L'_{u_k} = 2u_k(a_k^2 - \lambda) - \mu A_k, \qquad k = 1, \dots, n, \tag{17}$$

уравнения связи (15) и (16).

Домножение равенств (17) на u_k с тем же номером k, а затем суммирование получившихся равенств по всем k с учетом (15) и (16) дает

$$0 = \sum_{k=1}^{n} \big(2u_k^2(a_k^2 - \lambda) - \mu A_k u_k\big) = 2\sum_{k=1}^{n} u_k^2 a_k^2 - 2\lambda \sum_{k=1}^{n} u_k^2 - \mu \sum_{k=1}^{n} A_k u_k = 2(r^2 - \lambda) - 0,$$

или, в итоге,

$$r^2 = \lambda. \tag{18}$$

Предположим, что в какой-то точке u условного экстремума функции $r^2 = \sum_{k=1}^{n} a_k^2 u_k^2$ система (17) выполняется при $\mu = 0$. Тогда система (17) — это ничто иное как система отыскания собственного вектора u диагональной матрицы с числами $a_1^2, \dots, a_n^2$ на диагонали. Собственные числа такой матрицы — это числа, стоящие на диагонали, а раз они по условию попарно различны, то всякий собственный вектор такой матрицы коллинеарен одному из элементов стандартного базиса пространства $\mathbb{R}^n$, то есть у такого вектора не более одной ненулевой координаты. Но если такой вектор не вырожден, он не удовлетворяет (16), ибо по условию задачи вектор A не имеет нулевых компонент. То есть в точках u условного экстремума функции $r^2(u)$ условия (17) выполняются лишь при $\mu \neq 0$. А тогда, опять же, поскольку вектор A не имеет нулевых координат, должны быть отличными от нуля и все числа u_k и $a_k^2 - \lambda$, для которых верны равенства (17). В таком случае эти равенства допускают такую манипуляцию, как умножение каждого из равенств на дробь $\frac{A_k}{a_k^2 - \lambda}$ с тем же номером k и последующим суммированием по всем k, в результате чего получается

$$0 = 2\sum_{k=1}^{n} A_k u_k - \mu \sum_{k=1}^{n} \frac{A_k^2}{a_k^2 - \lambda} = 0 - \mu \sum_{k=1}^{n} \frac{A_k^2}{a_k^2 - r^2}. \tag{19}$$

Поскольку число μ, как показано выше, отлично от нуля, (19) означает, что требуемое равенство (12) выполняется, даже если R^2 интерпретировать как значение r^2 в произвольной стационарной точке функции Лагранжа. Тем более, равенство верно в точках глобальных условных экстремумов (экстремальные значения достигаются, поскольку множество связи есть компакт).

Следующая задача есть незначительная модификация задачи из [3] (для сокращения письма мы рассматриваем n-мерную ситуацию вместо 3-мерной и изменили шкалу времени).

Пример 3. Точка $q(t) = t \cdot (a_1, \dots, a_n)$ в момент времени $t = \frac{1}{\sqrt{n}}$ стартует с эллипсоида $\sum_{k=1}^n \frac{x_k^2}{a_k^2} = 1$. Пусть $p(t)$ — точка того же эллипсоида, ближайшая к $q(t)$ в момент времени t. Найдите предельное положение точки $p(t)$ при $t \to \infty$.

Решение. Точка $p = p(t)$ есть точка минимума функции $f(x) = ||x - q||^2 = \sum_{k=1}^n (x_k - ta_k)^2$ при условии $\sum_{k=1}^n \frac{x_k^2}{a_k^2} = 1$. В переменных

$$u_k = \frac{x_k}{a_k}, \qquad k = 1, \dots, n,$$

минимизируемая функция принимает вид $\sum_{k=1}^n a_k^2 (u_k - t)^2$, а ее минимум отыскивается на единичной сфере $\sum_{k=1}^n u_k^2 = 1$. Подозрительные на условный экстремум в момент времени t точки $u = u(t)$ — точки стационарности функции Лагранжа

$$L(u) = \sum_{k=1}^n a_k^2 (u_k - t)^2 - \lambda \left(\sum_{k=1}^n u_k^2 \right),$$

Зависящей от параметра t (производная по t не участвует в условии стационарности). Эти точки удовлетворяют системе

$$\begin{cases} a_k^2 (u_k - t) = \lambda u_k, \quad k = 1, \dots, n \\ \sum_{k=1}^n u_k^2 = 1. \end{cases} \tag{20}$$

Все участвующие переменные (кроме фиксированных параметров $a_1, \dots, a_n$)) зависят от времени t, для сокращения записи мы не отображаем этого в обозначениях. Так как при любом t все координаты вектора u ограничены по модулю единицей, выделение главный частей при $t \to +\infty$ в первом равенстве системы (20) дает асимптотическое равенство $\lambda u_k \sim -a_k^2 t$; его правая часть — бесконечно большая функция, значит такова же и $\lambda = \lambda(t)$. Сокращение равенства на λ дает

$$u_k(t) \sim a_k^2 g(t), \qquad \text{где} \quad g(t) = -\frac{t}{\lambda}$$

Так как функция $g(t)$ не зависит от k, получается, что стационарный вектор u «асимптотически коллинеарен» вектору $A = (a_1^2, \dots, a_n^2)$. Для завершения решения остается увидеть, что $g(t)$ имеет конечный предел в бесконечности. С этой целью совместим уравнение связи (второе уравнение системы (20)) с полученным асимптотическим равенством, точнее, его квадратом: $u_k^2 \sim a_k^4 g^2$. Положительные асимптотические равенства (в отличие от знакопеременных!) можно складывать (если для $k = 1, 2$ верно $\varphi_k \sim \psi_k > 0$, то $\varphi_k = \psi_k(1 + o(1)) \Rightarrow \varphi_1 + \varphi_2 = \psi_1 + \psi_2 + o(1)\psi_1 + o(1)\psi_2 = \psi_1 + \psi_2 + o(1)(\psi_1 + \psi_2) + o(1)(\psi_1 + \psi_2) = \psi_1 + \psi_2 + o(1)(\psi_1 + \psi_2)$), поэтому мы имеем:

$$1 = \sum_{k=1}^n u_k^2(t) \sim \sum_{k=1}^n a_k^4 g^2(t) = g^2(t) ||A||^2 \quad \Rightarrow \quad g^2(t) \to ||A||^{-2} \quad \text{при} \quad t \to +\infty.$$

Предельные значения $\pm||A||^{-1}$ отвечают ближней и дальней точкам соответсвенно, координаты ближней точки, конечно же, должны быть положительными. Таким образом, в координатах u искомое предельное решение есть $u = \frac{A}{||A||}$, переходя к исходным переменным получаем ответ.

Ответ: $\lim_{t\to+\infty} p(t) = (\sum_{k=1}^{n} a_k^4)^{-1/2} \cdot (a_1^3, \ldots, a_n^3)$.

3. Достаточное условие условного экстремума на основе знака второго дифференциала функции Лагранжа

Для классификации подозрительных на экстремум точек широко применяется достаточное условие условного экстремума в терминах знакоопределенности второго дифференциала функции Лагранжа. К недостаткам метода относится довольно большой объем вычислений в общем случае. Мы приводим два варианта этого условия, грубый и общий. Первый требует меньше вычислений, но он применим для более узкого класса задач. Кроме того, ГРУБЫЙ ВАРИАНТ ДОСТАТОЧНОГО УСЛОВИЯ (в отличие от общего варианта) НЕ ПОЗВОЛЯЕТ ДЕЛАТЬ ВЫВОД ОБ ОТСУТСТВИИ УСЛОВНОГО ЭКСТРЕМУМА (не только в случае полуопределенной квадратичной формы, но ДАЖЕ и ДЛЯ НЕОПРЕДЕЛЕННОЙ квадратичной формы!) Часто поступают так: сначала пытаются применить грубый вариант достаточного условия, а если не удается, то переходят к общему утверждению.

Грубое достаточное условие условного экстремума. *Пусть $m, n \in \mathbb{N}$, $m < n$, множество O открыто в $\mathbb{R}^n$, $f \in C^2(O)$, $F \in C^2(O \mapsto \mathbb{R}^m)$. Пусть в точке a выполняется необходимое условие условного экстремума, и λ — отвечающий точке a набор множителей Лагранжа (т.е. (a, λ) есть решение системы (9)), и $L(x)$ — функция Лагранжа, отвечающая данному набору λ.*

Тогда, если второй дифференциал

$$d_a^2 L = \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial x_i \partial x_j}(a) dx_i dx_j$$

как функция переменных $dx_1, \ldots, dx_n$ есть
— положительно определенная квадратичная форма, то a — это точка условного минимума функции f при условии (3);
— отрицательно определенная квадратичная форма, то a — это точка условного максимума функции f при условии (3).

Грубое достаточное условие условного экстремума есть частный случай общего достаточного условия.

Достаточное условие условного экстремума (общий вариант). *Пусть $m, n \in \mathbb{N}$, $m < n$, множество O открыто в $\mathbb{R}^n$, $f \in C^2(O)$, $F \in C^2(O \mapsto \mathbb{R}^m)$, Предположим, что в точке a выполняется необходимое условие условного экстремума, λ — отвечающий точке a набор множителей Лагранжа (т.е. (a, λ) есть решение системы (9)), и $L(x)$ — функция Лагранжа, отвечающая данному набору λ.*

Тогда, если сужение второго дифференциала на касательное пространство T_a к E в точке a

$$d_a^2 L\Big|_{T_a} = \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial x_i \partial x_j}(a) dx_i dx_j\Big|_{\{dx_1,\ldots,dx_n \ : \ d_a F = 0_m\}} \tag{21}$$

как функция переменных $dx_1, \dots, dx_n$ есть
— положительно определенная квадратичная форма, то a — это точка условного минимума функции f при условии (3);
— отрицательно определенная квадратичная форма, то a — это точка условного максимума функции f при условии (3).
— неопределенная (знакопеременная) квадратичная форма, то a есть седловая точка для функции f при условии (3).

Для доказательства достаточного условия условного экстремума удобна формула для второго дифференциала композиции.

Утверждение 2. *Пусть $O \subseteq \mathbb{R}^n$, $D \subseteq \mathbb{R}^k$ $f : O \mapsto \mathbb{R}$, $\Phi : D \mapsto O$, $\Phi(u) = x$, Φ и f дважды дифференцируемы в точках u и x соответственно. Тогда для произвольного $v \in \mathbb{R}^k$ верно равенство*

$$d_u^2(f \circ \Phi)(v) = d_x^2 f(d_u\Phi(v)) + d_x f(d_u^2\Phi(v)). \tag{22}$$

Более кратко утверждение 2 можно выразить следующим образом:

$$d^2(f \circ \Phi) = d^2 f \circ d\Phi + df \circ d^2\Phi.$$

Доказательство утверждения 2. Второй дифференциал композиции вычислим, продифференцировав первый $d_u(f \circ \Phi)(v) = d_x f \circ d_u\Phi(v) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \cdot d_u\Phi_i(v)$, тогда

$$d_u^2(f \circ \Phi)(v) = d_u\left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \cdot d_u\Phi_i(v)\right) =$$
$$\sum_{i,j=1} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \cdot d_u\Phi_i(v) d_u\Phi_j(v) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) \cdot d_u^2\Phi_i(v) = d_x^2 f(d_u\Phi(v)) + d_x f(d_u^2\Phi(v)),$$

(при последнем переходе мы воспользовались тем, что дифференциалы первого и остальных порядков к отображению применяются покоординатно: $d_u\Phi_i(v) = (d_u\Phi(v))_i$ и $d_u^2\Phi_i(v) = (d_u^2\Phi(v))_i$.

Доказательство достаточного условия условного экстремума. Т.к. на множестве E функция $f(x)$ совпадает с функцией Лагранжа $L(x)$, задачу исследования на экстремум $f|_E$ можно заменить задачей об исследовании $L|_E$. По теореме о способах задания гладкой поверхности существует регулярная C^2-гладкая локальная параметризация Φ поверхности E вблизи точки x_*, эту параметризацию, не умаляя общности, можно считать глобальной (иначе перейдем к рассмотрению меньшего множества O). Тогда Φ — это гомеоморфизм некоторого открытого в $\mathbb{R}^k$ множества D на E. В силу замечания 2 наличие экстремума функции $L|_E$ в точке x_* (т.е. условного экстремума функции f) равносильно наличию экстремума функции $g = L \circ \Phi$ в точке $u_* = \Phi^{-1}(x_*)$ — внутренней точке области определения этой функции, а значит, к функции g применимы условия безусловного экстремума: необходимое выполняется, т.к. $d_{u_*} g = d_{x_*} L \circ d_{u_*}\Phi$ и $d_{x_*} L = \mathbf{0}$ по условию (в x_* выполняется необходимое условие условного экстремума для f). Применяя при вычислении второго дифференциала композиции утверждение 2 и снова используя условие $d_{x_*}L = \mathbf{0}$, получаем:

$$d^2_{u_*}(L\circ\Phi)(v) = d^2_{x_*}L(d_{u_*}\Phi(v)) + d_{x_*}L(d^2_{u_*}\Phi(v)) = d^2_{x_*}L(d_{u_*}\Phi(v)).$$

Дифференциал параметризации $d_{u_*}\Phi$ осуществляет изоморфизм между $\mathbb{R}^k$ и касательным пространством T_{x_*} к E, поэтому $d^2_{u_*}(L\circ\Phi)$ является положительной (отрицательной либо неопределенной) квадратичной формой тогда и только тогда, когда $d^2_{x_*}L|_{T_*}$ является положительной (соответственно, отрицательной либо неопределенной) формой. Последнее наблюдение дает утверждение теоремы.

Пример 4. Исследовать на экстремум $f(x,y,z) = x+y+z$ при условиях $\frac{a^2}{x}+\frac{b^2}{y}+\frac{c^2}{z}=1,\ \ a\geq b\geq c>0$.

Решение. Определим функцию Лагранжа $L(x,y,z) = (x+y+z)-\lambda\left(\frac{a^2}{x}+\frac{b^2}{y}+\frac{c^2}{y}\right)$. В силу необходимого условия условного экстремума координаты подозрительной на условный экстремум точки (x,y,z) должны удовлетворять системе уравнений Лагранжа:

$$\begin{cases}\frac{\partial L}{\partial x}=\frac{\partial L}{\partial y}=\frac{\partial L}{\partial z}=0,\\ \frac{a^2}{x}+\frac{b^2}{y}+\frac{c^2}{z}=1.\end{cases} \Leftrightarrow \begin{cases}\frac{\lambda a^2}{x^2}=\frac{\lambda b^2}{y^2}=\frac{\lambda c^2}{z^2}=-1,\\ x\cdot\frac{a^2}{x^2}+y\cdot\frac{b^2}{y^2}+z\cdot\frac{c^2}{z^2}=1.\end{cases} \Leftrightarrow \quad \frac{a^2}{x^2}=\frac{b^2}{y^2}=\frac{c^2}{z^2}=-\frac{1}{\lambda}>0, \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \begin{cases}\lambda=-t^2, \quad t>0,\\ f(x,y,z)=t^2.\\ x=\operatorname{sign}(x)\cdot at,\ y=\operatorname{sign}(y)\cdot bt,\ z=\operatorname{sign}(z)\cdot ct.\end{cases} \tag{23}$$

Данная система разрешима тогда и только тогда, когда знаки переменных x,y,z удовлетворяют неравенству

$$a\cdot\operatorname{sign}(x)+b\cdot\operatorname{sign}(y)+c\cdot\operatorname{sign}(z)>0, \tag{24}$$

если неравенство верно, то

$$t=a\cdot\operatorname{sign}(x)+b\cdot\operatorname{sign}(y)+c\cdot\operatorname{sign}(z). \tag{25}$$

Второй дифференциал функции Лагранжа в точке (x,y,z), удовлетворяющей системе (23) есть

$$d^2_{(x,y,z)}L = 2t^2\left(\frac{a^2}{x^3}dx^2+\frac{b^2}{y^3}dy^2+\frac{c^2}{z^3}dz^2\right) = 2\left(\frac{dx^2}{x}+\frac{dy^2}{y}+\frac{dz^2}{z}\right).$$

В первом октанте $x>0, y>0, z>0$ имеется единственная подозрительная на условный экстремум точка — точка, заданная третьим уравнением системы (23) при $t=a+b+c$. Очевидно, в этой точке d^2L — это положительная квадратичная форма, то есть, уже краткий вариант достаточного условия условного экстремума позволяет сделать вывод о том, что (x,y,z) — это точка минимума.

Также очевидно, что в октанте $x<0, y<0, z<0$ отсутствуют точки, удовлетворяющие уравнению связи.

В остальных октантах ответ на вопрос о наличии точек условного экстремума не столь очевиден, продифференцируем уравнение связи и исследуем знак второго дифференциала функции Лагранжа при условии, что дифференциалы переменных связаны между собой полученным уравнением. В точке, удовлетворяющей системе Лагранжа, дифференцируя уравнение связи, получаем:

$$\frac{a^2}{x^2}dx + \frac{b^2}{y^2}dy + \frac{c^2}{z^2}dz = 0 \quad \Rightarrow \quad dx + dy + dz = 0. \tag{26}$$

Тогда

$$\frac{1}{2}d^2_{(x,y,z)}L\Big|_{dx+dy+dz=0} = \frac{dx^2}{x} + \frac{dy^2}{y} + \frac{(dx+dy)^2}{z} = \left(\frac{1}{x}+\frac{1}{z}\right)dx^2 + 2\frac{dx\,dy}{z} + \left(\frac{1}{y}+\frac{1}{z}\right)dy^2.$$

Матрица полученной квадратичной формы есть

$$\begin{pmatrix} \frac{1}{x}+\frac{1}{z} & \frac{1}{z} \\ \frac{1}{z} & \frac{1}{y}+\frac{1}{z} \end{pmatrix},$$

ее главные миноры $\Delta_1 = \frac{1}{x} + \frac{1}{z}$, $\Delta_2 = \left(\frac{1}{x}+\frac{1}{z}\right)\left(\frac{1}{y}+\frac{1}{z}\right) - \frac{1}{z^2} = \frac{x+y+z}{xyz} = \frac{f(x,y,z)}{xyz}$. В точке условного экстремума главный минор Δ_2 в силу критерия Сильвестра должен быть неотрицательным. В силу (23) числитель дроби $\frac{f(x,y,z)}{xyz}$ положителен, значит, эта дробь положительна если знаменатель положителен. Точки первого октанта, где все координаты положительны, нами уже исследован. Тогда осталось исследовать подозрительные точки, где две координаты из трех отрицательны. Поскольку по условию задачи $a \geq b \geq c > 0$ и знаки переменных x, y, z удовлетворяют неравенству (24), возможен лишь случай $x > 0, y < 0, z < 0$, в котором (24) — условие разрешимости системы уравнений Лагранжа — принимает вид

$$a > b + c. \tag{27}$$

И в таком случае, конечно, $a > c$, и тогда $\Delta_1 = \frac{1}{at} - \frac{1}{ct} = -\frac{a-c}{act} < 0$, то есть, исследуемая точка есть точка условного максимума. Координаты обеих найденных точек определяет последнее равенство системы (23) и (25).

Ответ. Условный минимум в точке $(a+b+c)\cdot(a,b,c)$.
При $a > b + c$ условный максимум в точке $(a-b-c)\cdot(a,-b,-c)$.

4. Симметричные и симметризуемые задачи об экстремуме

Следующий пример рассматривается в [5] (см. §5, **201**), где при решении задачи редуцируется число переменных, что сводит задачу к задаче о безусловном экстремуме, а исследование подозрительной точки проводится по индукции. Ниже мы показываем, что без уменьшения числа переменных — с сохранением симметрии — задача имеет короткое планиметрическое решение

Пример 5. Для произвольного натурального $n \geq 3$ среди всех n-угольников, вписанных в в данный круг радиуса R, найти тот, площадь которого наибольшая.

Решение. Заметим, что площадь вписанного в круг треугольника с двумя фиксированными вершинами максимальна, когда этот треугольник равнобедренный (тогда максимальна высота, проведенная к стороне, соединяющей две зафиксированные вершины). Это означает, что если вписанный в окружность n-угольник имеет максимальную площадь, то всякий треугольник, построенный на любой вершине n-угольника и двух смежных с ней, также должен быть равнобедренным. Тогда n-угольник должен быть правильным.

Ответ: наибольшую площадь имеет правильный n-угольник.

Один из популярных приемов решения задачи об условным экстремуме, особенно в случае малых размерностей, это уменьшение числа переменных — если система уравнений связи позволяет относительно легко выразить какие-то переменные через остальные. Однако в некоторых ситуациях оказывается удобным применить обратное действие, то есть увеличить число переменных, налагая на переменные некоторые ограничения. При этом задача о безусловном экстремуме превращается в задачу об условном экстремуме, но в задаче возникает некая симметризация, упрощающая вычисления. Примеры ниже иллюстрируют эту идею.

Пример 6. (вариация задачи Гюйгенса). Исследовать на экстремум функцию $f(x_1, \dots, x_n) = \frac{x_1 \cdot x_2 \cdot \ldots \cdot x_n}{(a+x_1)(x_1+x_2)\ldots(x_{n-1}+x_n)(x_n+b)}$ на множестве $\{(x_1, \dots, x_n) \mid a < x_1 < x_2 < \dots < x_n < b\}$, где a и b — положительные параметры и $a < b$.

Решение. Более удобный, чем у функции f, вид для исследования имеет функция $\ln \frac{1}{af} = \ln\left(1 + \frac{x_1}{a}\right) + \ln\left(1 + \frac{x_2}{x_1}\right) + \dots + \ln\left(1 + \frac{x_n}{x_{n-1}}\right) + \ln\left(1 + \frac{b}{x_n}\right)$, она имеет общие с f точки экстремума, но сами экстремумы при этом другого знака. Чтобы превратить функцию в симметрично зависящую от своих переменных, положим

$$t_1 = \frac{x_1}{a},\ t_2 = \frac{x_2}{x_1},\ \dots\ , t_n = \frac{x_n}{x_{n-1}},\ t_{n+1} = \frac{b}{x_n}. \tag{28}$$

Так определенные переменные связывает соотношение $t_1 \cdot t_2 \cdot\ \ldots\ \cdot t_{n+1} = \frac{b}{a}$, или

$$\sum_{k=1}^{n+1} \ln t_k = \ln \frac{b}{a}. \tag{29}$$

Условия $a < x_1 < x_2 < \dots < x_n < b$ равносильны условиям

$$t_1 > 1, \dots, t_{n+1} > 1. \tag{30}$$

Таким образом, исходная задача сведена к задаче об исследовании на условный экстремум функции $g(t_1, \dots, t_{n+1}) = \sum_{k=1}^{n+1} \ln(1 + t_k)$ при условиях (29) и (30). Функция связи регулярна всюду в области (30). Функция Лагранжа, отвечающая такой постановке задачи, есть

$$L(t_1, \dots, t_{n+1}) = \sum_{k=1}^{n+1} \ln(1 + t_k) - \lambda \left(\sum_{k=1}^{n+1} \ln t_k - \ln \frac{b}{a} \right).$$

Эта функция симметрично зависит от переменных $t_1, \ldots, t_{n+1}$; для всех номеров координат $k = 1, \ldots, n+1$

$$L'_{t_k}(t_1, \ldots, t_{n+1}) = \frac{1}{1+t_k} - \frac{\lambda}{t_k},$$

то есть все координаты точки подозрительной на условный экстремум удовлетворяют одному и тому же уравнению

$$\frac{t}{t+1} = \lambda \tag{31}$$

относительно переменной t с одним и тем же значением параметра λ. Левая часть этого уравнения строго монотонно зависит от t на луче $\{t > 1\}$, поэтому все координаты подозрительной точки должны совпадать; их общее значение t определяется из уравнения связи: $t = \sqrt[n+1]{\frac{b}{a}}$. Смешанные частные производные второго порядка функции Лагранжа, очевидно, тождественно равны нулю, а чистые в подозрительной точке удовлетворяют соотношениям

$$L''_{t_k^2}(t, \ldots, t) = -\frac{1}{(1+t)^2} + \frac{\lambda}{t^2} = \frac{1}{t^2}\left(\lambda - \left(\frac{t}{1+t}\right)^2\right) = \frac{(\lambda - \lambda^2)}{t^2} = \frac{\lambda(1-\lambda)}{t^2}.$$

В силу (31), очевидно, $\lambda \in (0; 1)$, поэтому все чистые частные производные функции Лагранжа второго порядка в подозрительной точке положительны, поэтому уже применение краткого варианта достаточного условия условного экстремума (см. [2], стр. 5) позволяет сделать вывод о том, что точка $(t, \ldots, t)$ — это точка условного минимума для функции $g(t_1, \ldots, t_n)$ при условиях (29) и (30). Осталось вернуться к переменным $x_1, \ldots, x_n$ используя (28).

Ответ: функция f имеет в заданной области единственную точку максимума $(x_1, \ldots, x_n)$, координаты которой $x_k = a\left(\frac{b}{a}\right)^{\frac{k}{n+1}}$ образуют геометрическую прогрессию со знаменателем $t = \sqrt[n+1]{\frac{b}{a}}$.

Следующая задача может быть решена с помощью тех же идей, что в последнем примере.

Упражнение. Исследовать функцию $f(x_1, \ldots, x_n) = \frac{a^2}{x_1} + \frac{x_1^2}{x_2} + \cdots + \frac{x_{n-1}^2}{x_n} + \frac{x_n^2}{b}$, где $a > 0, b > 0$ в области $x_1 > 0, \ldots, x_n > 0$.

Подсказка. Положив $x_0 = a$, $x_{n+1} = b$, $u_k = \frac{x_{k-1}^2}{x_k}$ для $k = 1, \ldots, n+1$, использовать соотношение $u_1^{2^n} u_2^{2^{n-1}} \cdot \ldots \cdot u_n^2 u_{n+1} = \frac{a^{2^{n+1}}}{b}$.

Список литературы

[1] Демидович Б. П., *Сборник задач и упражнений по математическому анализу: Учебное пособие для вузов.* М.: ООО «Издательство АСТ», 2005.

[2] Дубова Т. П., Семенова О. Л., *Условный экстремум. Отыскание экстремальных значений (методические указания).* СПб.: репозиторий СПбГУ, 2021.

[3] Зорич В. А., *Математический анализ. Часть I.* М: МЦНМО, 2002.

[4] Макаров Б. М., Подкорытов А. Н., *Гладкие функции и отображения.* М.: МЦНМО, 2020.

[5] Фихтенгольц Г. М., *Курс дифференциального и интегрального исчисления. В 3 т. Т.1* / Пред. и прим. А. А. Флоринского. — 8-е изд.— М.: ФИЗМАТЛИТ, 2003.

Навигатор по тексту