ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 26 № 4 2020

УДК 517.51

О ГЕОМЕТРИЧЕСКОМ ПОДХОДЕ К НАХОЖДЕНИЮ УСЛОВНЫХ ЭКСТРЕМУМОВ

Д. С. Теляковский, С. А. Теляковский

Даны геометрическая интерпретация и геометрическое доказательство необходимого условия существования условного экстремума. Изложенный подход применим к нахождению условных экстремумов недифференцируемых функций (т.е. когда метод неопределенных множителей Лагранжа в “классической” форме не может быть применен). В качестве примеров рассмотрены неравенство о среднем арифметическом и среднем геометрическом, неравенства Юнга и Йенсена.

Ключевые слова: условный экстремум, поверхность уровня, множители Лагранжа.

D. S. Telyakovskii, S. A. Telyakovskii. Geometric approach to finding constrained extrema.

In this paper, we give a geometric interpretation and a geometric proof of the necessary condition for the existence of a constrained extremum. The presented approach can be applied to finding constrained extrema of nondifferentiable functions (i.e., when Lagrange's method of undetermined multipliers is not applicable in the “classical” form). The following examples are considered: the inequality of arithmetic and geometric means, Young's inequality for products, and Jensen's inequality.

Keywords: constrained extremum, level surface, Lagrange multipliers.

MSC: 26B10

DOI: 10.21538/0134-4889-2020-26-4-244-254

Введение

Настоящий выпуск журнала “Труды Института математики и механики УрО РАН” посвящен столетию Сергея Борисовича Стечкина. Один из авторов этой статьи слушал лекции Сергея Борисовича по математическому анализу. Изложение его было наглядным, а не только формально правильным, четко выделялась основная идея, и становилось ясным, *почему* то или иное утверждение выполняется или не выполняется. В то же время, несмотря на несложное доказательство, алгоритм нахождения условных экстремумов, который дает теорема о множителях Лагранжа, выглядит очень формально. Мы рассматриваем задачу нахождения условных экстремумов с геометрической точки зрения и стремимся сделать изложение наглядным.

Пусть функция $f(\boldsymbol{x})$ определена в некоторой области $G \subset \mathbb{R}^n$ и множество $\Gamma \subset G$. Задача об условном экстремуме функции $f(\boldsymbol{x})$ на Γ состоит в нахождении точек множества Γ, в которых $f(\boldsymbol{x})$, рассматриваемая как функция на Γ, принимает экстремальные значения. Функцию $f(\boldsymbol{x})$ будем называть целевой функцией.

Задача отыскания условных экстремумов и связянные с этой задачей вопросы изучаются в таких активно развивающихся разделах современной математики как вариационное исчисление [1–4] и выпуклый анализ [5].

В разд. 1 настоящей работы на основании геометрических соображений о взаимном расположении поверхностей, касающихся этих поверхностей кривых и плоскостей, и прямых, касательных к этим поверхностям и кривым, получено необходимое условие условного экстремума функции $f(\boldsymbol{x})$ в точках Γ. Несмотря на геометрическую очевидность, полное и аккуратное доказательство этих соображений в рамках традиционных “продвинутых” курсов математического анализа является довольно длинным. Поэтому здесь будут доказаны только те геометрические факты, которые непосредственно используются для обоснования полученного в разд. 1 необходимого условия. Такое обоснование сделано в разд. 2. Наконец, в разд. 3 в качестве примеров доказаны неравенство о среднем арифметическом и среднем геометрическом, неравенства Юнга и Йенсена.

Под производной функции нескольких переменных понимается ее матрица Якоби. В частности, под производной скалярной функции векторного аргумента понимается ее градиент. Для записи векторов используется жирный курсив, нулевой вектор обозначается $\boldsymbol{\theta}$.

1. Необходимое условие условного экстремума с геометрической точки зрения

Если точка $\boldsymbol{x}_0 \in G$ является точкой безусловного экстремума функции $f(\boldsymbol{x})$, то $f(\boldsymbol{x}_0)$ является максимальным или минимальным значением функции $y = f(\boldsymbol{x})$ в некоторой окрестности точки $\boldsymbol{x}_0$. Геометрически это означает, что график функции $y = f(\boldsymbol{x})$ лежит с одной стороны от горизонтальной гиперплоскости $y = f(\boldsymbol{x}_0)$, не выше этой плоскости, если $\boldsymbol{x}_0$ — точка максимума $f(\boldsymbol{x})$, и не ниже, если это точка минимума. Это возможно, только если производная $f'(\boldsymbol{x}) = \nabla f(\boldsymbol{x})$ в точке $\boldsymbol{x}_0$ равна $\boldsymbol{\theta}$ или не существует. В настоящей работе дается аналогичная геометрическая интерпретация необходимого условия существования условного экстремума.

Предположим, что в области G функция $f(\boldsymbol{x})$ непрерывно дифференцируема и ее производная $f'(\boldsymbol{x}) \neq \boldsymbol{\theta}$. Тогда для каждой точки $\boldsymbol{x}_0 \in G$ поверхность уровня $S_{f(\boldsymbol{x}_0)}$ функции $f(\boldsymbol{x})$, проходящая через $\boldsymbol{x}_0$, является гладкой гиперповерхностью, причем с одной стороны от $S_{f(\boldsymbol{x}_0)}$ значения функции $f(\boldsymbol{x})$ больше, а с другой — меньше $f(\boldsymbol{x}_0)$ (рис. 1*а*). Поэтому в точке $\boldsymbol{x}_0 \in \Gamma$ функция $f(\boldsymbol{x})$ принимает экстремальное значение относительно Γ только если в некоторой окрестности точки $\boldsymbol{x}_0$ множество Γ лежит по одну сторону от поверхности уровня $S_{f(\boldsymbol{x}_0)}$ (рис. 1*б* и 1*в*). Заметим, что при наложенных условиях функция $f(\boldsymbol{x})$ не может иметь экстремумов относительно множества Γ во внутренних точках Γ (рис. 1*б*), поэтому будем предполагать, что Γ кривая.

Пусть множество Γ задано системой k уравнений

$$g_1(\boldsymbol{x}) = 0, \ldots, g_k(\boldsymbol{x}) = 0, \ k < n, \quad \text{или в векторном виде} \ \ \boldsymbol{g}(\boldsymbol{x}) = \begin{Vmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) \end{Vmatrix} = \boldsymbol{\theta}.$$

Будем считать, что в G все функции $g_j(\boldsymbol{x})$ непрерывно дифференцируемы, ранг матрицы производной $\boldsymbol{g}'(\boldsymbol{x})$ всюду в G равен k и $\boldsymbol{g}(\boldsymbol{x}_0) = \boldsymbol{\theta}$ в некоторой точке $\boldsymbol{x}_0 \in G$. При этих условиях

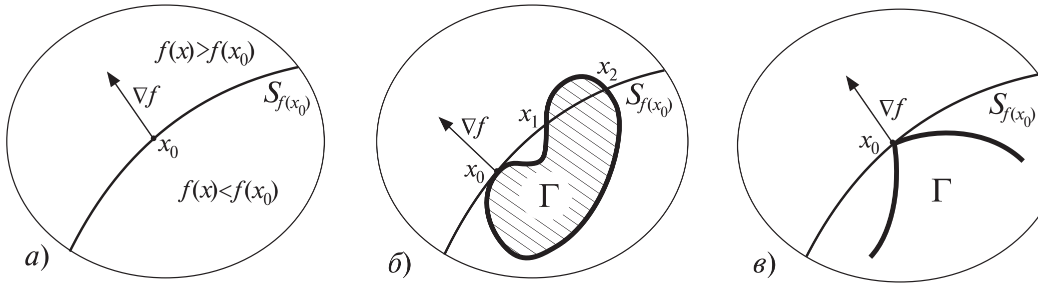

Рис. 1. *а*) Значения функции $f(\boldsymbol{x})$ с той стороны от гиперповерхности $S_{f(\boldsymbol{x}_0)}$, куда направлен градиент $\nabla f\big|_{\boldsymbol{x}_0}$, больше $f(\boldsymbol{x}_0)$, а с противоположной стороны от $S_{f(\boldsymbol{x}_0)}$ меньше $f(\boldsymbol{x}_0)$.
б) В точке $\boldsymbol{x}_0$ условный экстремум у функции $f(\boldsymbol{x})$ относительно Γ есть, в точках $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$ — нет.
в) Гладкость Γ в точках экстремума предполагать необязательно.

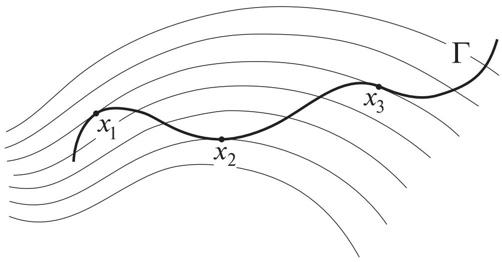

Рис. 2. Тонкими линиями обозначены линии уровня функции $f(\boldsymbol{x})$. В точках $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$ функция $f(\boldsymbol{x})$ имеет условные экстремумы относительно Γ, в точке $\boldsymbol{x}_3$ экстремума нет.

уравнение $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\theta}$ имеет в области G решения, и по теореме о неявно заданной функции через точку $\boldsymbol{x}_0$ проходит гладкое $(n-k)$-мерное многообразием Γ, на котором $\boldsymbol{g}(\boldsymbol{x}_0) \equiv \boldsymbol{\theta}$. Многообразие Γ будем называть кривой условия, а функцию $\boldsymbol{g}(\boldsymbol{x})$ — функцией условия.

Заметим, что условие $\operatorname{rank} \boldsymbol{g}'(\boldsymbol{x}) = k$ нужно по существу. Без него множество решений уравнения $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\theta}$ может быть не гладким многообразием переменной размерности, как можно было бы подумать, а вообще не образовывать гладкой поверхности. Это поазывает уравнение конуса $g(x,y,z) = x^2 + y^2 - z^2 = 0$ с вершиной в начале координат O. Множество решений уравнения $g(x,y,z) = 0$ не содержит никакой гладкой поверхности, проходящей через O, а в точке O имеем $\dfrac{\partial g}{\partial x} = \dfrac{\partial g}{\partial y} = \dfrac{\partial g}{\partial z} = 0$.

Матрица $\boldsymbol{g}'(\boldsymbol{x})$ состоит из k строк, которые являются градиентами ∇g_j компонент функции условия. Поскольку всюду в G ранг матрицы $\boldsymbol{g}'(\boldsymbol{x})$ равен количеству строк этой матрицы, то в каждой точке G строки $\boldsymbol{g}'(\boldsymbol{x})$ (а это градиенты ∇g_j) линейно независимы, в частности, невырождены. Поэтому каждое множество $S_j := \{\boldsymbol{x} \in G : g_j(\boldsymbol{x}) = 0\}$ является гладкой гиперповерхностью. Кривая Γ получается как пересечение $\bigcap_{j=1}^{k} S_j$, и поэтому Γ в каждой своей точке ортогональна нормальным векторам к поверхностям S_j, т. е. k линейно независимым векторам ∇g_j.

Рассмотрим поверхность уровня $S_{f(\boldsymbol{x}_0)}$ функции $f(\boldsymbol{x})$, проходящую через произвольную точку $\boldsymbol{x}_0 \in \Gamma$. Так как Γ и гиперповерхность $S_{f(\boldsymbol{x}_0)}$ являются гладкими, то в достаточно малой окрестности точки $\boldsymbol{x}_0$ кривая Γ может лежать по одну сторону от $S_{f(\boldsymbol{x}_0)}$, только если Γ касается $S_{f(\boldsymbol{x}_0)}$ в точке $\boldsymbol{x}_0$. Это необходимое условие экстремума функции $f(\boldsymbol{x})$ относительно кривой Γ является геометрически очевидным, позже оно будет доказано. Разумеется, условие касания кривой Γ поверхности уровня $S_{f(\boldsymbol{x}_0)}$ достаточным условием экстремума не является (рис. 2).

Касание кривой Γ гиперповерхности $S_{f(\boldsymbol{x}_0)}$ в точке $\boldsymbol{x}_0$ означает, что нормальный к $S_{f(\boldsymbol{x}_0)}$ вектор $\nabla f\big|_{\boldsymbol{x}_0}$ ортогонален Γ в точке $\boldsymbol{x}_0$. Покажем, что это выполнено тогда и только тогда, когда градиент целевой функции $\nabla f\big|_{\boldsymbol{x}_0}$ является линейной комбинацией градиентов функций условий $\nabla g_j\big|_{\boldsymbol{x}_0}$, $j = 1, \dots, k$.

Действительно, пусть кривая Γ касается $S_{f(\boldsymbol{x}_0)}$ в точке $\boldsymbol{x}_0$. Поскольку Γ является гладким $(n-k)$-мерным многообразием, то через каждую точку Γ проходит $(n-k)$-мерная плоскость, касательная к Γ. Пусть $\ell_{\boldsymbol{x}_0}$ — такая плоскость, которая касается Γ в точке $\boldsymbol{x}_0$ и $v_{\boldsymbol{x}_0}$ — ее направляющее $(n-k)$-мерное подпространство (рис. 3, *а*) и *б*)). Все векторы, ортогональные Γ в точке $\boldsymbol{x}_0$, ортогональны $v_{\boldsymbol{x}_0}$. Получены $k+1$ векторы $\nabla f\big|_{\boldsymbol{x}_0}$, $j = 1, \dots, k$, $\nabla g_j\big|_{\boldsymbol{x}_0}$, которые ортогональны $(n-k)$-мерному пространству $v_{\boldsymbol{x}_0}$. Следовательно, эти векторы линейно зависимы. Так как градиенты $\nabla g_j\big|_{\boldsymbol{x}_0}$, $j = 1, \dots, k$, линейно независимы, то вектор $\nabla f\big|_{\boldsymbol{x}_0}$ является их линейной комбинацией, и при некоторых значениях чисел λ_j выполняется равенство

$$\nabla f\big|_{\boldsymbol{x}_0} = \lambda_1 \nabla g_1\big|_{\boldsymbol{x}_0} + \dots + \lambda_k \nabla g_k\big|_{\boldsymbol{x}_0}. \qquad (1)$$

Наоборот, пусть при некоторых значениях λ_j справедливо представление (1). Поскольку кривая Γ ортогональна каждому вектору $\nabla g_j\big|_{\boldsymbol{x}_0}$, $j = 1, \ldots, k$, то Γ ортогональна и их линейной комбинации — градиенту $\nabla f\big|_{\boldsymbol{x}_0}$, который не равен $\boldsymbol{\theta}$. Это значит, что Γ ортогональна нормали к поверхности $S_{f(\boldsymbol{x}_0)}$, следовательно, кривая Γ в точке $\boldsymbol{x}_0$ касается поверхности $S_{f(\boldsymbol{x}_0)}$.

Прежде чем сформулировать полученное необходимое условие условного экстремума, напомним условия, наложенные на функции $f(\boldsymbol{x})$ и $\boldsymbol{g}(\boldsymbol{x})$ и на многообразие Γ. Функции $f(\boldsymbol{x})$ и $g_j(\boldsymbol{x})$, $j = 1, \ldots, k$, предполагались непрерывно дифференцируемыми в области $G \subset \mathbb{R}^n$, $n > k$, в каждой точке области G производная $f'(\boldsymbol{x})$ невырождена, $\operatorname{rank} \boldsymbol{g}'(\boldsymbol{x}) = k$, и гладкое многообразие $\Gamma \subset G$ задано уравнением $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\theta}$. Тогда справедливо следующее необходимое условие условного экстремума.

Теорема. *Если в некоторой точке $\boldsymbol{x}_0 \in \Gamma$ функция $f(\boldsymbol{x})$ имеет экстремум относительно кривой Γ, то в $\boldsymbol{x}_0$ кривая Γ касается поверхности уровня функции $f(\boldsymbol{x})$, проходящей через точку $\boldsymbol{x}_0$. Необходимым и достаточным условием касания является существование набора чисел $\{\lambda_j\}$, $j = 1, \ldots, k$, для которого*

$$\nabla f\big|_{\boldsymbol{x}_0} = \lambda_1 \nabla g_1\big|_{\boldsymbol{x}_0} + \cdots + \lambda_k \nabla g_k\big|_{\boldsymbol{x}_0}.$$

Для нахождения условных экстремумов обычно используют теорему о множителях Лагранжа, согласно которой задача нахождения экстремумов функции $f(\boldsymbol{x})$ при условии $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{\theta}$ эквивалентна задаче нахождения безусловных экстремумов функции Лагранжа

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) := f(\boldsymbol{x}) - \sum_{j=1}^{k} \lambda_j g_j(\boldsymbol{x}) \quad \text{при} \quad \boldsymbol{x} \in G.$$

Согласно предположениям о $f(\boldsymbol{x})$ и $\boldsymbol{g}(\boldsymbol{x})$ функция Лагранжа дифференцируема при $\boldsymbol{x} \in G$ и $\boldsymbol{\lambda} \in \mathbb{R}^k$. В точках экстремума $L(\boldsymbol{x}, \boldsymbol{\lambda})$ имеем

$$\begin{aligned} dL(\boldsymbol{x}, \boldsymbol{\lambda}) &= f'(\boldsymbol{x}) d\boldsymbol{x} - \sum_{j=1}^{n} \lambda_j g_j'(\boldsymbol{x}) d\boldsymbol{x} - \sum_{j=1}^{n} g_j(\boldsymbol{x}) d\lambda_j \\ &= \Big(f'(\boldsymbol{x}) - \sum_{j=1}^{n} \lambda_j g_j'(\boldsymbol{x}) \Big) d\boldsymbol{x} - \sum_{j=1}^{n} g_j(\boldsymbol{x}) d\lambda_j = \boldsymbol{\theta}. \end{aligned}$$

Равенство нулю векторного коэффициента при $d\boldsymbol{x}$ означает выполнение представления (1) в точках экстремума функции Лагранжа, а равенство нулю скалярных коэффициентов при $d\lambda_j$ — принадлежность этих точек кривой условия Γ. Заметим, что коэффициенты $\lambda_1, \ldots, \lambda_k$ разложения в формуле (1) являются множителями Лагранжа из теоремы об условном экстремуме.

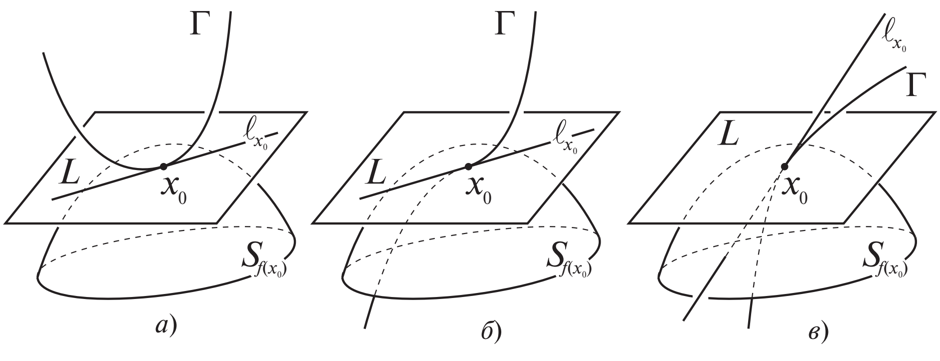

Рис. 3. Кривая Γ касается поверхности $S_{f(\boldsymbol{x}_0)}$ в точке $\boldsymbol{x}_0$ тогда и только тогда, когда касательная к Γ плоскость $\ell_{\boldsymbol{x}_0}$ лежит в плоскости L, касательной к $S_{f(\boldsymbol{x}_0)}$ в $\boldsymbol{x}_0$ (рис. *а*) и *б*)). Если $\ell_{\boldsymbol{x}_0}$ не лежит в L, то на Γ есть точки, лежащие по обе стороны от поверхности $S_{f(\boldsymbol{x}_0)}$ (рис. *в*)).

Поэтому для нахождения точек кривой $\boldsymbol{x} \in \Gamma$, в которых функция $f(\boldsymbol{x})$ может принимать экстремальные относительно Γ значения, достаточно найти точки Γ, в которых градиент ∇f является линейной комбинацией градиентов ∇g_j, $j = 1, \dots, k$. Функцию Лагранжа можно при этом не использовать. Как будет видно из приведенных ниже примеров, исследование найденных точек на наличие экстремума также иногда можно провести без использования функции Лагранжа.

При получении необходимого условия условного экстремума используются следующие геометрические соображения:

а) Если гладкая кривая Γ в достаточно малой окрестности своей точки $\boldsymbol{x}_0$ расположена по одну сторону от проходящей через $\boldsymbol{x}_0$ гладкой гиперповерхности $S_{f(\boldsymbol{x}_0)}$, то Γ в точке $\boldsymbol{x}_0$ касается $S_{f(\boldsymbol{x}_0)}$.

б) Касание Γ гиперповерхности $S_{f(\boldsymbol{x}_0)}$ эквивалентно тому, что касательная к Γ в $\boldsymbol{x}_0$ плоскость $\ell_{\boldsymbol{x}_0}$ лежит в гиперплоскости $L_{\boldsymbol{x}_0}$, касательной к $S_{f(\boldsymbol{x}_0)}$ в точке $\boldsymbol{x}_0$.

Эти соображения очевидны геометрически, но их аккуратное доказательство выходит за рамки традиционного курса математического анализа. Поэтому в следующем разделе работы будут доказаны некоторые геометрические факты, достаточные для обоснования полученного необходимого условия условного экстремума.

То, что в точке условного экстремума кривая условия Γ касается гиперповерхности уровня $S_{f(\boldsymbol{x}_0)}$, использовалось в качестве иллюстрации необходимого условия условного экстремума, например, Р. Курантом [6] и Л. Д. Кудрявцевым [7]. По воспоминаниям слушателей, Г. Е. Шилов в своих лекциях по математическому анализу на механико-математическом факультете МГУ доказывал теорему о необходимом условии условного экстремума с геометрической точки зрения. В его учебнике [8] эта теорема рассмотрена с точки зрения функционального анализа. Доказательства в [8] и в настоящей работе по существу реализуют одну и ту же идею.

2. Доказательство теоремы

Введем обозначения.

Для большей наглядности аргумент функций f и $\boldsymbol{g}$ представим в виде $(\boldsymbol{x}, \boldsymbol{y})$, где $\boldsymbol{x} = (x_1, \dots, x_m)$, $\boldsymbol{y} = (y_1, \dots, y_k)$, $m + k = n$. Это делается для того, чтобы функцию, заданную уравнением $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\theta}$, можно было записать в привычном виде $\boldsymbol{y}(\boldsymbol{x})$.

Пусть функция $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ дифференцируема. Введем обозначения для подматриц ее производной

$$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} := \begin{Vmatrix} \dfrac{\partial g_1}{\partial x_1} & \dots & \dfrac{\partial g_1}{\partial x_m} \\ \dots & \dots & \dots \\ \dfrac{\partial g_k}{\partial x_1} & \dots & \dfrac{\partial g_k}{\partial x_m} \end{Vmatrix} \quad \text{и} \quad \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}} := \begin{Vmatrix} \dfrac{\partial g_1}{\partial y_1} & \dots & \dfrac{\partial g_1}{\partial y_k} \\ \dots & \dots & \dots \\ \dfrac{\partial g_k}{\partial y_1} & \dots & \dfrac{\partial g_k}{\partial y_k} \end{Vmatrix}.$$

Используя эти обозначения, матрицу производной функции $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ можно записать в виде $\left\| \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}, \ \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}} \right\|$.

Приведем поэтапное доказательство теоремы.

1°. Получим явное уравнение прямой ℓ_0, касательной к кривой Γ в точке $(\boldsymbol{x}_0, \boldsymbol{y}_0)$.

Пусть кривая Γ задана уравнением $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\theta}$ и выполнены условия теоремы о существовании неявно заданной функции: функция $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ непрерывно дифференцируема, $\det \dfrac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}} \neq 0$ и точка $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ является решением уравнения $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\theta}$. При этих условиях уравнение $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\theta}$ можно разрешить относительно $\boldsymbol{y}$, т. е. найдутся такая окрестность $U(\boldsymbol{x}_0)$ и такая дифференцируемая функция $\boldsymbol{y}(\boldsymbol{x})$, определенная на этой окрестности, для которых на $U(\boldsymbol{x}_0)$

тождественно $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))\equiv\boldsymbol{\theta}$ и $\boldsymbol{y}(\boldsymbol{x}_0)=\boldsymbol{y}_0$, причем в некоторой окрестности точки $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ уравнение $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{\theta}$ других решений, кроме $(\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))$, $\boldsymbol{x}\in U(\boldsymbol{x}_0)$, не имеет.

Так как функция $\boldsymbol{y}(\boldsymbol{x})$ дифференцируема в точке $\boldsymbol{x}_0$, то

$$\boldsymbol{y}(\boldsymbol{x})=\boldsymbol{y}(\boldsymbol{x}_0)+\boldsymbol{y}'(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)+\boldsymbol{o}(\boldsymbol{x}-\boldsymbol{x}_0),\quad \boldsymbol{x}\to\boldsymbol{x}_0, \tag{2}$$

и уравнение

$$\boldsymbol{y}_{\text{кас}}(\boldsymbol{x})=\boldsymbol{y}(\boldsymbol{x}_0)+\boldsymbol{y}'(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0),\quad \boldsymbol{x}\in\mathbb{R}^m, \tag{3}$$

задает m-мерное линейное многообразие, или, другими словами, m-мерную плоскость ℓ_0, проходящую через точку $(\boldsymbol{x}_0,\boldsymbol{y}_0)$. Условие дифференцируемости (2) означает, что в точке $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ плоскость ℓ_0 и кривая Γ касаются.

2°. По условию система градиентов $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$, $j=1,\dots,k$, линейно независима. Покажем, что она является базисом ортогонального дополнения $(n-k)$-мерного направляющего пространства прямой ℓ_0. Для этого достаточно проверить, что каждый из градиентов $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$, $j=1,\dots,k$, ортогонален прямой ℓ_0.

Кривая Γ определена как множество решений системы уравнений $g_j(\boldsymbol{x},\boldsymbol{y})=0$, $j=1,\dots,k$. Как уже отмечалось, каждое из этих уравнений определяет гладкую гиперповерхность $S_j:=\{(\boldsymbol{x},\boldsymbol{y})\in G\colon g_j(\boldsymbol{x},\boldsymbol{y})=0\}$ и $\Gamma=\bigcap_{j=1}^k S_j$. Так как градиенты $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ являются нормальными векторами гиперплоскостей L_j, касательных к поверхностям S_j в точке $(\boldsymbol{x}_0,\boldsymbol{y}_0)$, то для доказательства ортогональности прямой ℓ_0 градиентам $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ достаточно проверить, что $\ell_0=\bigcap_{j=1}^k L_j$. Поскольку каждая из плоскостей L_j проходит через точку $(\boldsymbol{x}_0,\boldsymbol{y}_0)\in\ell_0$, то равенство $\ell_0=\bigcap_{j=1}^k L_j$ будет доказано, если будет доказано совпадение направляющих подпространств линейных многообразий ℓ_0 и $\bigcap_{j=1}^k L_j$.

Каждая гиперплоскость L_j ортогональна градиенту $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$, поэтому направляющее подпространство линейного многообразия $\bigcap_{j=1}^k L_j$ задается системой однородных линейных уравнений, у матрицы которой градиенты $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ являются строками, т. е. системой с матрицей $\Big\|\dfrac{\partial\boldsymbol{g}}{\partial\boldsymbol{x}},\ \dfrac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\Big\|$. Здесь и далее значения производных взяты в точке $(\boldsymbol{x}_0,\boldsymbol{y}_0)$.

Теперь получим систему, определяющую направляющее подпространство прямой ℓ_0. В некоторой окрестности точки $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ функция $\boldsymbol{y}(\boldsymbol{x})$, $\boldsymbol{x}\in U(\boldsymbol{x}_0)$, является решением уравнения $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{\theta}$. Поэтому $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))\equiv\boldsymbol{\theta}$ в $U(\boldsymbol{x}_0)$. Дифференцируя это тождество, получим, что в $U(\boldsymbol{x}_0)$

$$\frac{d}{d\boldsymbol{x}}\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))=\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{x}}+\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\frac{d\boldsymbol{y}}{d\boldsymbol{x}}\equiv\mathcal{O},$$

где $\mathcal{O}$ — нулевая матрица соответствующего размера. Отсюда, поскольку матрица $\dfrac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}$ невырождена, следует, что

$$\frac{d\boldsymbol{y}}{d\boldsymbol{x}}=-\Big(\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\Big)^{-1}\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{x}}. \tag{4}$$

Прямая ℓ_0 задана уравнением

$$\boldsymbol{y}(\boldsymbol{x})=\boldsymbol{y}(\boldsymbol{x}_0)+\boldsymbol{y}'(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0),\quad \boldsymbol{x}\in\mathbb{R}^m.$$

Записав это уравнение в виде

$$\boldsymbol{y}'(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)-(\boldsymbol{y}(\boldsymbol{x})-\boldsymbol{y}(\boldsymbol{x}_0))=\boldsymbol{\theta},$$

видим, что направляющее подпространство прямой ℓ_0 задано системой с матрицей $\|\boldsymbol{y}'(\boldsymbol{x}_0),-E\|$. Подставив сюда выражение для $\boldsymbol{y}'(\boldsymbol{x}_0)$ из (4), получим

$$\Big\|\frac{d\boldsymbol{y}}{d\boldsymbol{x}},-E\Big\|=\Big\|-\Big(\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\Big)^{-1}\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{x}},-E\Big\|=-\Big(\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\Big)^{-1}\Big\|\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{x}},\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\Big\|.$$

Эквивалентность систем, задающих направляющие подпространства линейных многообразий ℓ_0 и $\bigcap_{j=1}^k L_j$, доказана, поэтому $\ell_0 = \bigcap_{j=1}^k L_j$. Поскольку нормальным вектором плоскости L_j является градиент $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$, то каждый градиент $\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ ортогонален ℓ_0, и, следовательно, система градиентов является базисом ортогонального дополнения направляющего пространства прямой ℓ_0.

Отметим, в частности, что формула (4) для производной функции $\boldsymbol{y}(\boldsymbol{x})$, заданной уравнением $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{\theta}$, аналогична формуле $\dfrac{dy}{dx} = -\dfrac{\partial g}{\partial x}\Big/\dfrac{\partial g}{\partial y}$ производной функции $y(x)$, заданной уравнением $g(x,y)=0$.

3°. Теперь проверим, что если плоскость ℓ_0 не лежит в гиперплоскости L_0, касательной в точке $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ к поверхности уровня $S_{f(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ функции $f(\boldsymbol{x},\boldsymbol{y})$, то точка $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ не может быть точкой экстремума $f(\boldsymbol{x},\boldsymbol{y})$ относительно Γ.

Предположим, что ℓ_0 не содержится в L_0. Тогда найдется одномерная прямая $l_0 \subset \ell_0$, не лежащая в L_0. Прямая l_0 проходит через точку $(\boldsymbol{x}_0,\boldsymbol{y}_0)$, и ее точки имеют координаты $(\boldsymbol{x}_0+\boldsymbol{h}t, \boldsymbol{y}_0+\boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}t)$, где $\boldsymbol{h}\in\mathbb{R}^m$ — некоторый фиксированный вектор, а $t\in\mathbb{R}$. Из уравнения (3) касательного многообразия ℓ_0 видно, что так выражаются координаты точек любой одномерной прямой, содержащейся в ℓ_0. По предположению $l_0 \not\subset L_0$, поэтому направляющий вектор $(\boldsymbol{h},\boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h})$ прямой l_0 не ортогонален нормальному вектору плоскости L_0, т. е. градиенту $\nabla f\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$, и скалярное произведение этих векторов не равно нулю

$$\Big((\boldsymbol{h},\boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}), \nabla f\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}\Big) = \left\|\frac{\partial f}{\partial \boldsymbol{x}} \quad \frac{\partial f}{\partial \boldsymbol{y}}\right\| \left\|\begin{matrix}\boldsymbol{h}\\ \boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}\end{matrix}\right\| \ne 0. \tag{5}$$

Одномерной прямой $l_0\subset\ell_0$ соответствует одномерная кривая $\gamma_0\subset\Gamma$, точки которой имеют координаты $(\boldsymbol{x}_0+\boldsymbol{h}t, \boldsymbol{y}(\boldsymbol{x}_0+\boldsymbol{h}t))$, где величина t настолько мала, что точка $(\boldsymbol{x}_0+\boldsymbol{h}t, \boldsymbol{y}(\boldsymbol{x}_0+\boldsymbol{h}t))$ лежит в той окрестности точки $(\boldsymbol{x}_0,\boldsymbol{y}_0)$, в которой функция $\boldsymbol{y}(\boldsymbol{x})$ является решением уравнения $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{\theta}$. Поскольку функция $\boldsymbol{y}(\boldsymbol{x})$ дифференцируема в точке $\boldsymbol{x}_0$, то координаты точек кривой γ_0 имеют вид

$$\big(\boldsymbol{x}_0+\boldsymbol{h}t, \boldsymbol{y}(\boldsymbol{x}_0)+\boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}t+\boldsymbol{o}(t)\big) \quad \text{при} \quad t\to 0.$$

Это значит, что одномерная кривая γ_0 касается одномерной прямой l_0 в точке $(\boldsymbol{x}_0,\boldsymbol{y}_0)$.

Поэтому для значений функции $f(\boldsymbol{x},\boldsymbol{y})$ на кривой γ_0 имеем

$$f(\boldsymbol{x}_0+\boldsymbol{h}t, \boldsymbol{y}_0+\boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}t+\boldsymbol{o}(t))$$

$$= f(\boldsymbol{x}_0,\boldsymbol{y}_0) + \frac{\partial f}{\partial \boldsymbol{x}}\boldsymbol{h}t + \frac{\partial f}{\partial \boldsymbol{y}}\boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}t + o(t) = f(\boldsymbol{x}_0,\boldsymbol{y}_0) + \left\|\frac{\partial f}{\partial \boldsymbol{x}} \quad \frac{\partial f}{\partial \boldsymbol{y}}\right\| \left\|\begin{matrix}\boldsymbol{h}\\ \boldsymbol{y}'(\boldsymbol{x}_0)\boldsymbol{h}\end{matrix}\right\| t + o(t). \tag{6}$$

Условие (5) означает, что коэффициент при t в линейной относительно t части выражения (6) не равен нулю, и, следовательно, в любой близости от точки $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ есть точки из Γ, в которых функция $f(\boldsymbol{x},\boldsymbol{y})$ принимает значения как большие, так и меньшие $f(\boldsymbol{x}_0,\boldsymbol{y}_0)$.

Значит, если $\ell_0 \not\subset L_0$, то точка $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ не может быть точкой экстремума функции $f(\boldsymbol{x},\boldsymbol{y})$ относительно Γ. Таким образом, если $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ — точка экстремума $f(\boldsymbol{x},\boldsymbol{y})$ относительно Γ, то плоскость ℓ_0, касательная в точке $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ к кривой условия Γ, лежит в гиперплоскости L_0, касательной к поверхности уровня $S_{f(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ функции $f(\boldsymbol{x},\boldsymbol{y})$. Значит, вектор $\nabla f\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ ортогонален прямой ℓ_0. Как было показано в п. 2°, система градиентов $\big\{\nabla g_j\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}\big\}$ является базисом ортогонального дополнения к ℓ_0. Поэтому вектор $\nabla f\big|_{(\boldsymbol{x}_0,\boldsymbol{y}_0)}$ раскладывается по этой системе, т. е. выполнено соотношение (1). Это завершает доказательство теоремы. $\square$

3. Примеры

В качестве примеров докажем неравенство, сравнивающее среднее арифметическое и среднее геометрическое, и неравенства Юнга и Йенсена, рассматривая эти неравенства как задачи нахождения условных экстремумов. Как будет видно на примере неравенства Йенсена, геометрический подход иногда позволяет проводить исследование на экстремум и в случае, когда целевая функция недифференцируема, и поэтому применить функцию Лагранжа обычным образом нельзя. Идея рассмотреть неравенство, связывающее среднее арифметическое и среднее геометрическое, и неравенство Юнга как задачи нахождения условных экстремумов подчерпнута из книги Р. Куранта [6].

Среднее арифметическое и среднее геометрическое. Покажем, что для произвольного набора n положительных чисел $x_1, \ldots, x_n$ выполнено неравенство

$$\frac{x_1 + \ldots + x_n}{n} \geqslant \sqrt[n]{x_1 \ldots x_n}.$$

Будем рассматривать эту задачу как задачу нахождения минимального значения функции $f(\boldsymbol{x}) := x_1 + \ldots + x_n$ на поверхности Γ, заданной уравнением $g(\boldsymbol{x}) := x_1 \ldots x_n = C^n$, где все x_j положительны и C — некоторое число. Таким образом, решается задача нахождения минимума функции $f(\boldsymbol{x})$ при одном условии $g(\boldsymbol{x}) = C^n$.

Поверхность Γ не ограничена, поэтому сначала покажем, что $f(\boldsymbol{x})$ принимает на Γ минимальное значение. Обозначим Γ_0 часть поверхности Γ, на которой $f(\boldsymbol{x}) \leqslant 2nC$. Множество Γ_0 замкнуто и не пусто, так как Γ_0 содержит, например, точку $\boldsymbol{x}_0 := (C, \ldots, C)$. Поэтому непрерывная функция $f(\boldsymbol{x})$ принимает на Γ_0 минимальное положительное значение. Так как согласно определению Γ_0 на $\Gamma \setminus \Gamma_0$ имеем $f(\boldsymbol{x}) > 2nC > f(\boldsymbol{x}_0) = nC$, то минимум, который достигает $f(\boldsymbol{x})$ на Γ_0, является минимумом $f(\boldsymbol{x})$ на всей поверхности Γ.

В точках условного экстремума градиент целевой функции является линейной комбинацией градиентов функций условий. Здесь имеется только одно уравнение связи $g(\boldsymbol{x}) = C^n$, поэтому необходимое условие экстремума $f(\boldsymbol{x})$ на Γ означает коллинеарность $\nabla f \parallel \nabla g$ градиентов $f(\boldsymbol{x})$ и $g(\boldsymbol{x})$ в точках экстремума. Поскольку $\nabla f = (1, \ldots, 1)$ и

$$\begin{aligned}\nabla g &= (x_2 x_3 \ldots x_n,\ x_1 x_3 \ldots x_n,\ \ldots,\ x_1 \ldots x_{n-2} x_{n-1}) \\ &= \Big(\frac{\prod_1^n x_j}{x_1}, \ldots, \frac{\prod_1^n x_j}{x_n}\Big) = \Big(\frac{C^n}{x_1}, \ldots, \frac{C^n}{x_n}\Big),\end{aligned}$$

то вектор ∇g коллинеарен вектору ∇f с равными компонентами только в тех точках поверхности Γ, в которых $x_1 = x_2 = \ldots = x_n$, т. е. только в точке $x_0 := (C, \ldots, C)$. Следовательно, на поверхности Γ функция $f(\boldsymbol{x})$ может иметь экстремум только в точке $\boldsymbol{x}_0$. Поскольку на поверхности Γ у функции $f(\boldsymbol{x})$ есть минимум, то этот минимум достигается в точке $\boldsymbol{x}_0$ и равен $nC = n\sqrt[n]{x_1 \ldots x_n}$.

Таким образом, для произвольного набора n положительных чисел $x_1, \ldots, x_n$ имеем

$$\frac{x_1 + \ldots + x_n}{n} \geqslant \frac{n\sqrt[n]{x_1 \ldots x_n}}{n} = \sqrt[n]{x_1 \ldots x_n},$$

причем равенство выполняется только при $x_1 = x_2 = \cdots = x_n$.

Неравенство Юнга. Докажем, что для любых положительных x и y и любых положительных α и β, для которых $\dfrac{1}{\alpha} + \dfrac{1}{\beta} = 1$ (т. е. $\alpha\beta = \alpha + \beta$), выполнено неравенство

$$\frac{1}{\alpha} x^{\alpha} + \frac{1}{\beta} y^{\beta} \geqslant xy.$$

Аналогично предыдущему примеру будем искать минимум функции $f(x, y) = \dfrac{1}{\alpha} x^{\alpha} + \dfrac{1}{\beta} y^{\beta}$ на кривой Γ, определенной уравнением $g(x, y) = xy = C > 0$. Как и при доказательстве предыдущего неравенства, проверяется, что $f(x,y)$ достигает на Γ положительного минимума. В точках экстремума $f(x,y)$ на кривой Γ должно выполняться условие $\nabla f \parallel \nabla g$. Поскольку $\nabla f = (x^{\alpha-1}, y^{\beta-1})$ и $\nabla g = (y,x)$, то условие коллинеарности градиентов означает, что

$$\frac{y}{x^{\alpha-1}} = \frac{x}{y^{\beta-1}}, \quad \text{т. е.} \quad x^\alpha = y^\beta.$$

Так как точка (x,y) с координатами, удовлетворяющими этому условию, принадлежит кривой Γ с уравнением $xy = C$, то условие $\nabla f \parallel \nabla g$ выполнено только в точке $(x_0,y_0) = (C^{\frac{\beta}{\alpha+\beta}}, C^{\frac{\alpha}{\alpha+\beta}})$. Таким образом, функция $f(x,y)$ на кривой Γ может иметь экстремум только в точке (x_0,y_0). Поскольку у $f(x,y)$ на Γ есть минимум, то найденная точка (x_0,y_0) является точкой минимума. Таким образом, учитывая равенство $\dfrac{1}{\alpha}+\dfrac{1}{\beta}=1$, т. е. $\alpha\beta = \alpha+\beta$, получаем, что для произвольных положительных чисел x и y

$$\frac{1}{\alpha}x^\alpha + \frac{1}{\alpha}y^\beta \geqslant \frac{1}{\alpha}\Big[(xy)^{\frac{\beta}{\alpha+\beta}}\Big]^\alpha + \frac{1}{\beta}\Big[(xy)^{\frac{\alpha}{\alpha+\beta}}\Big]^\beta = \Big(\frac{1}{\alpha}+\frac{1}{\beta}\Big)(xy)^{\frac{\alpha\beta}{\alpha+\beta}} = xy,$$

причем равенство имеет место только при $x^\alpha = y^\beta$. Таким образом, неравенство Юнга доказано.

Неравенство Йенсена. Неравенство Йенсена также можно рассматривать как задачу нахождения условного экстремума. Действительно, пусть на интервале (a,b) определена функция $\varphi(x)$ и задан некоторый набор точек $\{x_j\}$, $j = 1,\dots,n$, этого интервала. Для набора из n неотрицательных чисел t_j при условии $\sum t_j = 1$ надо оценить сумму $\sum t_j\varphi(x_j)$. При этих условиях точки $\big(\sum t_j x_j, \sum t_j\varphi(x_j)\big)$ заполняют минимальную выпуклую оболочку Ω набора точек $\{(x_j,\varphi(x_j))\}$. Поэтому величину $\sum t_j\varphi(x_j)$ можно оценить сверху и снизу функциями, задающими верхнюю и нижнюю границы множества Ω. Неравенство Йенсена дает такую оценку, если функция $\varphi(x)$ выпукла или вогнута.

Покажем, что если функция $\varphi(x)$ выпукла на интервале (a,b), то для любых наборов точек $\{x_j\}$ и чисел $\{t_j\}$ указанного вида выполнено неравенство

$$\varphi\Big(\sum_{j=1}^n t_j x_j\Big) \leqslant \sum_{j=1}^n t_j\varphi(x_j).$$

В отличие от предыдущих примеров гладкость функции $\varphi(x)$ здесь не предполагается.

Будем считать, что $a < x_1 < x_2 < \dots < x_n < b$ и числа t_j положительны. Положим $x^* := t_1x_1 + \dots + t_nx_n$. Ясно, что точка x^* лежит в интервале (a,b). Если функция $\varphi(x)$ выпукла,

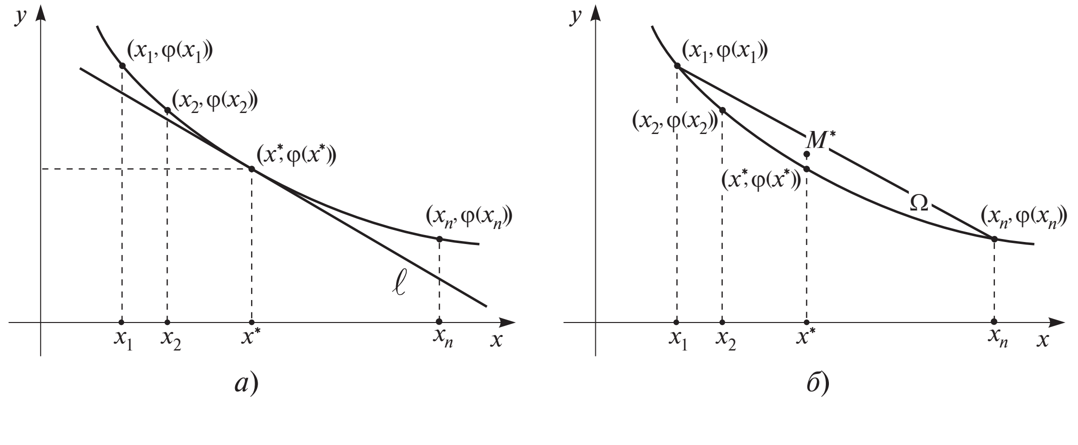

Рис. 4.

то через точку $(x^*, \varphi(x^*))$ ее графика проходит опорная прямая ℓ к графику $y = \varphi(x)$ (рис. 4a). Пусть $y = kx + b$ — уравнение этой прямой. Тогда

$$\sum_{j=1}^{n} t_j(kx_j + b) = k\sum_{j=1}^{n} t_j x_j + b\sum_{j=1}^{n} t_j = kx^* + b = \varphi(x^*).$$

Так как опорная прямая ℓ лежит не выше графика функции $y = \varphi(x)$, то в каждой точке x_j выполнено неравенство $kx_j + b \leqslant \varphi(x_j)$, $j = 1, \ldots, n$, и поэтому

$$\varphi\Big(\sum_{j=1}^{n} t_j x_j\Big) = \varphi(x^*) = \sum_{j=1}^{n} t_j(kx_j + b) \leqslant \sum_{j=1}^{n} t_j \varphi(x_j).$$

Из доказательства видно, что равенство достигается, только если функция $\varphi(x)$ является линейной на отрезке $[x_1, x_n]$.

Заметим, что приведенное доказательство неравенства Йенсена основано на той же идее, что и доказательство этого неравенства путем исследования расположения центра тяжести системы масс t_j, расположенных в точках $\big(x_j, \varphi(x_j)\big)$ графика выпуклой функции. А. Н. Лоуэн [9], пользуясь таким подходом, получил несколько неравенств, которые являются частными случаями неравенства Йенсена для конкретных выпуклых функций. К доказательству самого неравенства Йенсена эта идея была применена М. Б. Балком [10]. Приведем это изящное доказательство, хотя в нем и не используются соображения нахождения экстремумов.

Если Ω — выпуклое множество и масса произвольным образом распределена внутри Ω, то центр масс точка M^* лежит внутри Ω или на $\partial\Omega$. Выпуклость функции $\varphi(x)$ на интервале (a, b) означает, что каждое множество, ограниченное снизу графиком $y = \varphi(x)$, а сверху — отрезком с концами в двух произвольных точках графика $y = \varphi(x)$, является выпуклым. Возьмем n произвольных точек x_j интервала (a, b) и n произвольных положительных чисел t_j, сумма которых равна единице: $\sum t_j = 1$. Будем считать, что $x_1 < x_2 < \cdots < x_n$. Центр масс M^* системы масс t_j, помещенных в точки $(x_j, \varphi(x_j))$, находится в точке $\big(\sum t_j x_j, \sum t_j \varphi(x_j)\big)$. Точка M^* лежит внутри или на границе множества Ω, ограниченного снизу графиком $y = \varphi(x)$ на отрезке $[x_1, x_n]$, а сверху — отрезком с концами в точках $(x_1, \varphi(x_1))$ и $(x_n, \varphi(x_n))$. Поэтому при $x^* := \sum t_j x_j \in [x_1, x_n]$ точка $\big(\sum t_j x_j, \varphi\big(\sum t_j x_j\big)\big)$ графика $\varphi(x)$ лежит не выше центра масс M^* (рис. 4$б$). Это означает выполнение неравенства Йенсена $\varphi\big(\sum t_j x_j\big) \leqslant \sum t_j \varphi(x_j)$, при этом равенство здесь достигается тогда и только тогда, когда функция $\varphi(x)$ на отрезке $[x_1, x_n]$ является линейной.

СПИСОК ЛИТЕРАТУРЫ

1. **Mordukhovich Boris S.** Variational analysis and generalized differentiation, I: Basic theory. Berlin: Springer-Verlag, 2006. 579 p. (Ser. Fundamental Principles of Math. Sci.; vol. 330).
2. **Vinter Richard.** Optimal control. Boston: Birkhäuser, 2010. 507 p.
3. **Santambrogio Filippo.** Optimal transport for applied mathematicians: Calculus of variations, PDEs, and modeling. Base: Birkhäuserl, 2015. 353 p. (Ser. Progress in Nonlinear Differential Equations and Their Applications; vol. 87).
4. **Clarke F.H., Ledyaev Yu.S., Stern R.J., Wolenski R.R.** Nonsmooth analysis and control theory. N Y: Springer-Verlag, 1998. 278 p. (Ser. Graduate Texts in Math.).
5. **Половинкин Е.С., Балашов М.В.** Элементы выпуклого и сильно выпуклого анализа. М.: Физматлит, 2004. 416 с.
6. **Курант Р.** Курс дифференциального и интегрального исчисления. Т. 2. М.: Наука, 1970. 672 с.
7. **Кудрявцев Л.Д.** Математический анализ. Т. 2. М.: Дрофа, 2004. 720 с.
8. **Шилов Г.Е.** Математический анализ. Функции нескольких вещественных переменных. М.: Наука, 1972. 624 с.
9. **Lowan A.N.** Note on an elementary method for generating inequalities. Scripta Mathematica. 1955. Vol. 21, no. 2–3. P. 218–220.

10. **Балк М.Б.** Геометрические приложения понятия о центре тяжести. М.: Физматлит, 1959. 230 с. (Сер. Библиотека математического кружка; вып. 9).

Поступила 9.01.2020
После доработки 7.10.2020
Принята к публикации 26.10.2020

Теляковский Дмитрий Сергеевич
канд. физ.-мат. наук, доцент
Национальный исследовательский ядерный университет (МИФИ)
г. Москва
e-mail: dtelyakov@mail.ru

Теляковский Сергей Александрович
др. физ.-мат. наук, профессор
ведущий науч. сотрудник
Математический институт им. В.А. Стеклова РАН
г. Москва
e-mail: sergeyaltel@yandex.ru

REFERENCES

1. Mordukhovich Boris S. *Variational analysis and generalized differentiation*, I: *Basic theory.* Ser. Grundlehren der mathematischen Wissenschaften, vol. 330, Berlin: Springer-Verlag, 2006, 579 p.
2. Vinter Richard. *Optimal control.* Boston: Birkhauser, 2010, 507 p.
3. Santambrogio Filippo. *Optimal transport for applied mathematicians: Calculus of variations, PDEs, and modeling*, Ser. Progress in Nonlinear Differential Equations and Their Applications, vol. 87, Basel: Birkhäuser, 2015, 353 p.
4. Clarke F.H., Ledyaev Yu.S., Stern R.J., Wolenski R.R. *Nonsmooth analysis and control theory.* Ser. Graduate Texts in Math., N Y: Springer-Verlag, 1998, 278 p.
5. Polovinkin E.S., Balashov M.V. *Elementy vypuklogo i sil'no vypuklogo analiza* [Elements are convex and strongly convex analysis], Moscow: Fizmatlit Publ., 2004, 416 p.
6. Courant R. *Differential and integral calculus*, vol. 2. Ishi Press, 2010, 692 p. ISBN: 978-4871878357. Translated to Russian under the title *Kurs differencial'nogo i integral'nogo ischisleniya*, vol. 2. Moscow: Nauka Publ., 1970, 672 p.
7. Kudryavtsev L.D. *Matematicheskii analiz* [Mathematical analysis], vol. 2., Moscow: Drofa Publ., 2004, 720 p.
8. Shilov G.E. *Matematicheskii analiz. Funktsii neskol'kikh veshchestvennykh peremennykh* [Mathematical analysis. Functions of several real variables]. Moscow: Nauka Publ., 1972, 624 p.
9. Lowan A.N. Note on an elementary method for generating inequalities. *Scripta Mathematica*, 1955, vol. 21, no. 2–3, pp. 218–220.
10. Balk M.B. *Geometricheskie prilozheniya ponyatiya o tsentre tyazhesti* [Geometricheskie prilozheniya ponyatiya o centre tyazhesti]. Moscow: Fizmatlit, 1959, 230 p.

Received January 9, 2020
Revised October 7, 2020
Accepted October 26, 2020

Dmitry Sergeevich Telyakovskii, Cand. Sci. (Phys.-Math.), Prof., National Research Nuclear University (MEPhI), Moscow, 115409 Russia, e-mail: dtelyakov@mail.ru .

Sergey Alexandrovich Telyakovskii, Dr. Phys.-Math. Sci., Prof., Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, 119991 Russia, e-mail: sergeyaltel@yandex.ru .

Cite this article as: D. S. Telyakovskii, S. A. Telyakovskii. Geometric approach to finding the conditional extrema, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2020, vol. 26, no. 4, pp. 244–254 .