Экстремум функций нескольких переменных

М.Ю.Баландин*

Май 2004

Аннотация

Рассматриваются вопросы, связанные с темой «Экстремум функции нескольких переменных» из второго семестра курса математического анализа. Приводятся примеры решения задач на безусловный экстремум явно и неявно заданных функций, а также на поиск условного экстремума методом множителей Лагранжа и методом исключения части переменных.

Предисловие

Первоначально эти заметки планировались как изложение метода решения задач на условный экстремум с использованием функции Лагранжа, однако по мере того, как к ним дописывалась теоретическая часть, становилось всё более очевидно, что лучше уж добавить сюда примеры и на безусловный экстремум — благо много места они не займут.

Описываемый здесь способ в значительной мере опирается на такие понятия линейной алгебры, как квадратичные формы и знакоопределённость матриц[1]. Это сильно отличается от описываемого в литературе подхода, однако в конечном счёте метод получается хорошо алгоритмизируемым.

Относительно используемых обозначений требуется сделать лишь два замечания. Во-первых, для векторов и матриц используется жирный шрифт (единственным исключением является символ «набла», который иначе как для обозначения векторного оператора Гамильтона вообще не используется). Это же относится к векторам, составленным из дифференциалов (например, $\mathbf{dx}$). Во-вторых, для матрицы вторых производных функции используется обозначение $\mathbf{D}_2(\mathbf{x})$, где под $\mathbf{x}$ понимается вся совокупность аргументов функции (например, $\mathbf{x} = (x, y, z)$).

Все задачи взяты из сборника Б. П. Демидовича [1], а в качестве источника для самостоятельной подготовки всячески рекомендуется книга [2]. Перечень решённых задач (в порядке упоминания по тексту) приведен в таблице 2 на стр. 12.

Подробный ход решения нелинейных систем уравнений приводится не во всех случаях, иногда он описывается лишь в общих чертах. Подобные системы представляют как раз тот случай, когда средства компьютерной алгебры способны оказать значительную помощь (при подготовке заметок использовались Maple и Derive).

*Кафедра прикладной математики НГТУ.

[1] К середине второго семестра, когда курс математического анализа подходит к теме функций нескольких переменных, соответствующие темы как раз оказываются пройденными в курсе линейной алгебры

1. Дифференциалы и линейная алгебра

1.1. Критерий Сильвестра

Сколько-нибудь грамотное решение задач на экстремум функции многих переменных невозможно без представления о квадратичных формах и знакоопределённости квадратных матриц. Напомним сответствующие определения:

> *Квадратная матрица* $\mathbf{A}$ *называется* положительно определённой*, если квадратичная форма* $\omega(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ *на ненулевых векторах* $\mathbf{x}$ *принимает только положительные значения:*
>
> $$\mathbf{A} > 0 \iff \forall \mathbf{x} \neq 0 : \quad \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0.$$
>
> *Аналогично, матрица* $\mathbf{A}$ *называется* отрицательно определённой*, если* $\omega(\mathbf{x}) < 0$ *при* $\mathbf{x} \neq 0$*:*
>
> $$\mathbf{A} < 0 \iff \forall \mathbf{x} \neq 0 : \quad \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} < 0.$$

Сразу из определения следует, что матрица будет положительно (отрицательно) определена тогда и только тогда, когда все её собственные значения положительны (отрицательны). На практике, однако, этим критерием пользоваться неудобно: выкладки слишком громоздки, и нахождение собственных значения для матриц размерности более трех уже весьма проблематично.

Гораздо более простые условия знакоопределённости даёт *критерий Сильвестра*, из которого студенты почему-то помнят лишь половину. Вот он полностью:

> *Для того, чтобы матрица* $\mathbf{A}$ *была положительно определена, необходимо и достаточно, чтобы все её главные миноры* $\Delta_i(\mathbf{A})$ *были положительны:*
>
> $$\mathbf{A} > 0 \iff \forall i : \quad \Delta_i(\mathbf{A}) > 0.$$
>
> *Для того, чтобы матрица* $\mathbf{A}$ *была отрицательно определена, необходимо и достаточно, чтобы её главные миноры имели чередующиеся знаки,* ***начиная с отрицательного****:*
>
> $$\mathbf{A} < 0 \iff \forall i : \quad \operatorname{sgn}(\Delta_i(\mathbf{A})) = (-1)^i.$$

В частности, матрица $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ будет положительно определена при $a > 0$, $ad - bc > 0$ и отрицательно определена при $a < 0$, $ad - bc > 0$.

1.2. Вектор, матрица и дифференциал

В этом параграфе мы установим матрично-векторные формы записи первого и второго дифференциала функции нескольких переменных, что позволит упростить рассуждения в § 3, касающемся условного экстремума.

Как правило, все студенты помнят выражение для градиента функции в виде вектора, составленного из частных производных:

$$\operatorname{grad} f(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots \frac{\partial f}{\partial x_n}\right]^{\mathrm{T}}.$$

Если ввести в рассмотрение оператор Гамильтона «набла», составленный из покомпонентного взятия частных производных (не из самих производных!)

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots \frac{\partial}{\partial x_n}\right)^{\mathrm{T}},$$

то $\operatorname{grad} f = \nabla f$, где операция умножения понимается в смысле умножения вектора ∇ на скаляр f, так что $[\nabla f]_i = \frac{\partial f}{\partial x_i}$. (Порядок сомножителей в таких операциях считается важным, так как операции дифференцирования должны приписываться к функциям спереди).

Если, кроме того составить из дифференциалов независимых переменных вектор $\mathbf{dx} = [dx_1, dx_2, \dots dx_n]^{\mathrm{T}}$, то первый дифференциал функции $f(\mathbf{x})$ можно записывать в векторном виде

$$df = \mathbf{dx} \cdot \operatorname{grad} f = (\mathbf{dx} \cdot \nabla) f.$$

Второй дифференциал, как известно, равен

$$d^2 f = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i\, \partial x_j}\, dx_i\, dx_j.$$

Знакомые с линейной алгеброй студенты без особого труда узна̀ют в этом выражении квадратичную форму $\mathbf{dx}^{\mathrm{T}}\mathbf{D}_2(\mathbf{x})\mathbf{dx}$. Мы будем предполагать, что смешанные производные не зависят от порядка дифференцирования[2] $\left(\frac{\partial^2 f}{\partial x_i\, \partial x_j} = \frac{\partial^2 f}{\partial x_j\, \partial x_i}\right)$, так что матрица $\mathbf{D}_2(\mathbf{x})$ в этих заметках будет считаться симметричной.

Вновь воспользовавшись оператором Гамильтона, матрицу вторых производных можно записать в виде

$$\mathbf{D}_2(\mathbf{x}) = (\nabla \cdot \nabla^{\mathrm{T}}) f,$$

и тогда матрично-векторная форма записи второго дифференциала (которая, главным образом, и понадобится нам в дальнейшем), имеет вид

$$d^2 f = \mathbf{dx}^{\mathrm{T}}\mathbf{D}_2(\mathbf{x})\mathbf{dx} = \mathbf{dx}^{\mathrm{T}}\left[(\nabla \cdot \nabla^{\mathrm{T}}) f\right]\mathbf{dx}.$$

1.3. Условия локального экстремума

Если формулировать условия локального экстремума в терминах дифференциалов (а не в терминах производных, как чаще всего почему-то делают студенты), то для случая нескольких переменных они ничем не будут отличаться от случая одной переменной:

- если дифференцируемая функция $f(\mathbf{x})$ имеет локальный экстремум в точке $\mathbf{x}_0$, то эта точка является для неё стационарной, т.е. $df\big|_{\mathbf{x}=\mathbf{x}_0} = 0$;
- стационарная точка $\mathbf{x}_0$ функции $f(\mathbf{x})$ будет являться локальным минимумом, если $d^2 f\big|_{\mathbf{x}=\mathbf{x}_0} > 0$ и локальным максимумом, если $d^2 f\big|_{\mathbf{x}=\mathbf{x}_0} < 0$.

Зная эти условия, нетрудно переформулировать их и в терминах производных, если учесть рассуждения предыдущего параграфа:

- в точке локального экстремума $\mathbf{x}_0$ дифференцируемой функции $f(\mathbf{x})$ все её частные производные равны нулю: $\frac{\partial f}{\partial x_i}\big|_{\mathbf{x}=\mathbf{x}_0} = 0$;

[2] На вопрос о том, когда это справедливо, студенты стандартно отвечают «а когда функция непрерывна!» Ответ совершенно неправильный, ибо если у функции в некоторой точке существуют вторые производные, то она в ней просто *обязана* быть непрерывной.

- стационарная точка $\mathbf{x}_0$ функции $f(\mathbf{x})$ будет являться локальным минимумом (максимумом), если в ней матрица вторых производных $\mathbf{D}_2(\mathbf{x})$ функции f определена положительно (отрицательно).

Заметим ещё, что в силу сделанных предположений матрица $\mathbf{D}_2(\mathbf{x})$ является симметричной. Это ускоряет её нахождение и делает все её собственные значения вещественными.

2. Безусловный экстремум

2.1. Явно заданная функция

Алгоритм поиска экстремумов явно заданной функции $f(\mathbf{x})$ сам по себе очень прост и состоит всего из двух шагов:

- найти стационарные точки функции, решив систему уравнений $\nabla f(\mathbf{x}) = 0$. Система в общем случае получается нелинейной и её решение может быть связано с чисто вычислительными трудностями;

- исследовать каждую из найденных стационарных точек, выяснив в этих точках знакоопределённость матрицы вторых производных $\mathbf{D}_2(\mathbf{x})$ функции $f(\mathbf{x})$.

Рассмотрим этот алгоритм на примере задачи **№ 3633**, где нужно исследовать на экстремум функцию

$$z(x, y) = e^{x^2-y}(5 - 2x + y).$$

Вычисляя частные производные функции и приравнивая их к нулю, получаем нелинейную систему двух уравнений

$$\begin{cases} -2e^{x^2-y}\big(2x^2 - x(y+5) + 1\big) = 0 \\ e^{x^2-y}(2x - y - 4) = 0 \end{cases}. \tag{1}$$

В данном случае можно воспользоваться тем фактом, что $e^{x^2-y} \neq 0$, а первое уравнение поделить еще и на двойку, так что система (1) упрощается:

$$\begin{cases} 2x^2 - x(y+5) = -1 \\ 2x - y = 4 \end{cases}.$$

Выражая из второго уравнения $y = 2x-4$ и подставляя в первое уравнение, получаем $x = 1$, откуда $y = -2$. Итак, функция имеет одну стационарную точку $\mathbf{x}_0 = (1, -2)^{\mathrm{T}}$.

Находим теперь вторые производные функции и составляем из них матрицу:

$$\mathbf{D}_2(\mathbf{x}) = \mathbf{D}_2(x, y) = e^{x^2-y}\begin{bmatrix} -8x^3 + 4x^2(y+5) - 12x + 2y + 10 & 4x^2 - 2x(y+4) + 2 \\ 4x^2 - 2x(y+4) + 2 & -2x + y + 3 \end{bmatrix}.$$

Подставляя сюда координаты стационарной точки, получаем

$$\mathbf{D}_2(\mathbf{x}_0) = \begin{bmatrix} -2e^3 & 2e^3 \\ 2e^3 & -e^3 \end{bmatrix}.$$

Эта матрица не является знакоопределённой, так как $\Delta_1 = -2e^3 < 0$ и $\Delta_2 = -2e^6 < 0$. Следовательно, стационарная точка не является точкой экстремума (седловая точка), и функция вообще не имеет локальных экстремумов.

В качестве другого примера рассмотрим **№ 3628**. Просят исследовать на локальный экстремум функцию

$$z(x, y) = xy + \frac{50}{x} + \frac{20}{y}$$

при условии что обе переменные положительны (первый квадрант).

Выписываем систему $\nabla z = 0$:

$$\begin{cases} y - \frac{50}{x^2} = 0 \\ x - \frac{20}{y^2} = 0 \end{cases}.$$

Переменные легко выражаются одна через другую, так что решение системы[3] не вызывает никаких проблем: $\mathbf{x}_0 = (5, 2)^{\mathrm{T}}$.

Вычисляя вторые производные, составляем матрицу $\mathbf{D}_2(\mathbf{x}_0)$:

$$\mathbf{D}_2(\mathbf{x}_0) = \begin{bmatrix} \frac{4}{5} & 1 \\ 1 & 5 \end{bmatrix}.$$

Для неё $\Delta_1 = \frac{4}{5} > 0$ и $\Delta_2 = 3 > 0$, так что матрица определена положительно и точка $x = 5$, $y = 2$ является точкой локального минимума (в ней $z(5, 2) = 30$).

2.2. Неявно заданная функция

Решение задачи о безусловном экстремуме неявно заданной функции почти ничем не отличается от случая явной функции, всё различие заключено лишь в способе нахождения производных и дифференциалов. Их вычисление облегчается тем, что второй дифференциал нужен лишь в стационарной точке, а это заметно упрощает выкладки.

В качестве примера рассмотрим **№ 3651**, где нужно найти точки экстремума неявно заданной функции $z(x, y)$, определённой уравнением

$$x^2 + y^2 + z^2 - 2x + 2y - 4z - 10 = 0. \tag{2}$$

Самой большой проблемой для студента является распознать, что функция задана именно *неявно*. Сплошь и рядом приходится видеть, как студенты, напрочь игнорируя запись $z = z(x, y)$, просто-напросто рассматривают левую часть (2) как некоторую *явно заданную функцию трёх переменных* и работают с ней как с $u(x, y, z)$ по правилам предыдущего параграфа. Окончание соотношения $(\ldots = 0)$ при этом тоже игнорируется.

Правильное же решение задачи заключается в том, чтобы взять полный дифференциал от соотношения (2):

$$(2x - 2)dx + (2y + 2)dy + (2z - 4)dz = 0 \tag{3}$$

и выразить из этого уравнения дифференциал рассматриваемой функции dz:

$$dz = \frac{x - 1}{2 - z}\, dx + \frac{y + 1}{2 - z}\, dy.$$

Приравнивая его к нулю, получаем простую систему[4]

$$\begin{cases} x - 1 = 0 \\ y + 1 = 0 \\ 2 - z \neq 0 \end{cases}.$$

[3]Интерес вызывает лишь одно из трёх существующих решений, так как два остальных являются комплексными (система приводится к уравнениям третьей степени).

[4]Разумеется, в общем случае система может получиться и очень сложной.

Итак, $x = 1$ и $y = -1$. Проверяем, будет ли в этой точке выполняться условие $z \neq 2$. Подстановка найденных x и y в (2) даёт

$$z^2 - 4z - 12 = 0,$$

откуда $z_1 = 6$ и $z_2 = -2$. Видим, что существуют *две*[5] функции $z(x, y)$, удовлетворяющие уравнению (2), и обе отвечают условию $z \neq 2$.

Рассмотрим теперь вопрос о характере экстремума. Нужно найти d^2z (для обеих функций); этот дифференциал легко получается повторным дифференцированием (3):

$$2\,dx^2 + \cancelto{0}{(2x-2)d^2x} + 2\,dy^2 + \cancelto{0}{(2y+2)d^2y} + \cancelto{0}{2\,dz^2} + (2z-4)d^2z = 0.$$

Здесь $d^2x = d^2y \equiv 0$, так как x и y полагаются независимыми переменными, а $dz = 0$, так как нас интересует стационарная точка функции $z_i(x, y)$.

Окончательно

$$d^2z = \frac{1}{2-z}\,dx^2 + \frac{1}{2-z}\,dy^2,$$

откуда матрица вторых производных имеет вид

$$\mathbf{D}_2(x, y, z_i) = \begin{bmatrix} \frac{1}{2-z_i} & 0 \\ 0 & \frac{1}{2-z_i} \end{bmatrix}.$$

Нетрудно убедиться, что для функции $z_1(x, y)$ она определена отрицательно (локальный максимум), а для $z_2(x, y)$ — положительно (локальный минимум).

Интересно, что в этой задаче матрица $\mathbf{D}_2$ получилась не зависящей от x и y. Это говорит о том, что функции z_1 и z_2 являются *всюду выпуклыми* и для них любая стационарная точка автоматически является точкой экстремума.

3. Условный экстремум

Способ решения задач на условный экстремум, изложенный в этом параграфе, является авторским «изобретением» в том смысле, что в известных автору книгах он не излагается; фактически же это не что иное, как матрично-векторная форма записи метода, описанного, например, в [2].

Пусть имеется задача

$$\begin{cases} u(\mathbf{x}) = u(x, y, \ldots, z) \to \min / \max \\ \varphi_i(\mathbf{x}) = 0, \quad i = \overline{1, n}, \quad n < \dim \mathbf{x} \end{cases}.$$

Для её решения необходимо составить *функцию Лагранжа*, число переменных которой будет равно $\dim \mathbf{x}$ (числу переменных функции $u(\mathbf{x})$), увеличенному на n:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = u(x, y, \ldots, z) + \sum_{i=1}^{n} \lambda_i \varphi_i(x, y, \ldots, z) = u(\mathbf{x}) + \boldsymbol{\lambda} \cdot \boldsymbol{\varphi}(\mathbf{x}), \tag{4}$$

где $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]^{\mathrm{T}}$ и $\boldsymbol{\varphi}(\mathbf{x}) = [\varphi_1(\mathbf{x}), \ldots, \varphi_n(\mathbf{x})]^{\mathrm{T}}$.

[5] Это следует и из того факта, что уравнение (2) является квадратным относительно z.

Далее нужно найти стационарную точку этой функции, решив систему (в общем случае нелинейную)

$$
\begin{cases}
\dfrac{\partial u}{\partial x} + \displaystyle\sum_{i=1}^{n} \lambda_i \dfrac{\partial \varphi_i}{\partial x} = 0 \\
\dfrac{\partial u}{\partial y} + \displaystyle\sum_{i=1}^{n} \lambda_i \dfrac{\partial \varphi_i}{\partial y} = 0 \\
\dots \\
\dfrac{\partial u}{\partial z} + \displaystyle\sum_{i=1}^{n} \lambda_i \dfrac{\partial \varphi_i}{\partial z} = 0 \\
\varphi_i(x, y, \dots, z) = 0, \quad i = \overline{1, n}
\end{cases}
$$

Пусть решением системы является точка $(\mathbf{x}_0, \boldsymbol{\lambda}_0)$. Необходимо подставить $\boldsymbol{\lambda}_0$ в функцию Лагранжа (4) и для определения характера экстремума исследовать знакоопределённость матрицы вторых производных $\mathbf{D}_2(\mathbf{x})$ получившейся функции $L(\mathbf{x}, \boldsymbol{\lambda}_0)$ в точке $\mathbf{x} = \mathbf{x}_0$.

Если эта матрица является знакоопределённой, то характер экстремума установлен и задача решена. В противном случае нужно попробовать учесть ограничения. Для этого необходимо продифференцировать функции φ_i, подставить в производные координаты точки $\mathbf{x}_0$ и получить для дифференциалов линейную систему

$$
\mathbf{A}\,\mathbf{dx} = 0, \quad \mathbf{A} \,:\, n \times \dim \mathbf{x}, \quad n < \dim \mathbf{x}.
$$

Такая система недоопределена и часть дифференциалов можно выразить через другие: $\mathbf{dx} = \boldsymbol{\Phi}\,\mathbf{d}\tilde{\mathbf{x}}$. С учётом этого, нужно исследовать матрицу $\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{D}_2(\mathbf{x}_0)\boldsymbol{\Phi}$ — её знакоопределённость отвечает на вопрос о характере экстремума уже однозначно.

На практике эти выкладки оказываются гораздо проще, чем в теории. Самым неприятным этапом по-прежнему является решение нелинейной системы при поиске стационарной точки. Рассмотрим, например, задачу **№ 3687**, условие которой сформулировано в геометрических терминах:

> *При каких размерах открытая прямоугольная ванна вместимости V имеет наименьшую[6] боковую поверхность?*

Ответом, очевидно, должны являться три числа $x, y, z > 0$, связанные с объемом V и определяющие три измерения требуемой ванны (рис. 1).

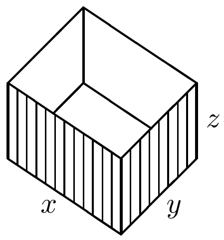

Рис. 1: к задаче № 3687.

Для изображённой на рисунке ванны площадь боковой поверхности состоит из площадей пяти граней и равна $S(x, y, z) = 2xz + 2yz + xy$. Таким образом, получаем задачу на условный

[6] Задача вполне практическая. Чем меньше боковая поверхность ванны, тем медленнее она остывает.

экстремум:

$$\begin{cases} S(x,y,z) = 2xz + 2yz + xy \to \min \\ xyz = V \end{cases}.$$

Составляем функцию Лагранжа, которая в данном случае будет зависеть от четырёх переменных:

$$L(x,y,z,\lambda) = 2xz + 2yz + xy + \lambda(xyz - V). \tag{5}$$

Найдём её частные производные и приравняем их к нулю. Получится нелинейная система

$$\begin{cases} 2z + y + \lambda yz = 0 \\ 2z + x + \lambda xz = 0 \\ 2x + 2y + \lambda xy = 0 \\ xyz = V \end{cases}. \tag{6}$$

Для решения этой системы вычтем из первого уравнения второе; получившееся равенство $(y-x)(1+\lambda z) = 0$ говорит о том, что либо $x = y$, либо $\lambda = -\frac{1}{z}$. Из этих двух вариантов второй не годится: подставляя $\lambda = -\frac{1}{z}$ в первое или второе уравнение, мы придём к $2z = 0$, что невозможно по смыслу задачи (если $z = 0$, то $xyz = 0 \neq V$).

Итак, $x = y$, что неудивительно: поменяв местами x и y, мы получим ту же самую ванну, но повёрнутую другим боком. С учётом этого, первое и второе уравнения системы (6) оказываются одинаковыми и мы приходим к системе меньшей размерности:

$$\begin{cases} 2z + x + \lambda xz = 0 \\ 4x + \lambda x^2 = 0 \\ x^2 z = V \end{cases}. \tag{7}$$

Из второго уравнения (7) получаем $x(4 + \lambda x) = 0$, откуда либо $x = 0$ (что недопустимо по смыслу задачи), либо $\lambda = -\frac{4}{x}$. Размерность системы ещё уменьшается:

$$\begin{cases} x - 2z = 0 \\ x^2 z = V \end{cases}. \tag{8}$$

Такая система уже легко разрешима: подставляя во второе уравнение (8) выраженное из первого её уравнения $x = 2z$, получаем $z = \sqrt[3]{V/4}$ и далее находим все остальные координаты стационарной точки функции Лагранжа (5):

$$\mathbf{x}_0 = (x_0, y_0, z_0)^{\mathrm{T}} = (\sqrt[3]{2V}, \sqrt[3]{2V}, \sqrt[3]{V/4})^{\mathrm{T}}, \quad \lambda_0 = -\sqrt[3]{32/V}.$$

Далее необходимо убедиться, что найденная точка действительно является *максимумом* функции $S(x,y,z)$ с учётом ограничения $xyz = V$. Для этого найдём матрицу вторых производных функции $L(\mathbf{x}, \lambda_0)$:

$$\mathbf{D}_2(\mathbf{x}) = \begin{bmatrix} 0 & 1 - z\sqrt[3]{32/V} & 2 - y\sqrt[3]{32/V} \\ 1 - z\sqrt[3]{32/V} & 0 & 2 - x\sqrt[3]{32/V} \\ 2 - y\sqrt[3]{32/V} & 2 - x\sqrt[3]{32/V} & 0 \end{bmatrix}.$$

В точке $\mathbf{x}_0$ эта матрица равна

$$\mathbf{D}_2(\mathbf{x}_0) = \begin{bmatrix} 0 & -1 & -2 \\ -1 & 0 & -2 \\ -2 & -2 & 0 \end{bmatrix}.$$

Сама по себе она не является знакоопределённой. Поэтому, чтобы решить вопрос о характере экстремума, необходимо учесть ограничение $xyz = V$.

Дифференциал этого выражения даёт $yz\,dx + xz\,dy + xy\,dz = 0$. Подставляя сюда координаты точки $\mathbf{x}_0$, получаем

$$\tfrac{1}{2}\sqrt[3]{4V^2}\,dx + \tfrac{1}{2}\sqrt[3]{4V^2}\,dy + \sqrt[3]{4V^2}\,dz = 0 \;\Rightarrow\; dx + dy + 2dz = 0.$$

Здесь три дифференциала связаны одним линейным уравнением, так что один из них однозначно определяется двумя остальными. Например, можно записать, что $dz = -\frac{1}{2}dx - \frac{1}{2}dy$. В матрично-векторной форме эта запись выглядит следующим образом:

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}. \tag{9}$$

Если ввести обозначения

$$\tilde{\mathbf{x}} = [x, y]^{\mathrm{T}} \quad \text{и} \quad \mathbf{\Phi} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix},$$

то соотношение (9) примет вид $\mathbf{dx} = \mathbf{\Phi}\,\mathbf{d\tilde{x}}$ и второй дифференциал функции $L(\mathbf{x}, \lambda_0)$ с учётом ограничения равен $d^2_{\text{огр}}L(\mathbf{x}_0, \lambda_0) = \mathbf{d\tilde{x}}^{\mathrm{T}}\left[\mathbf{\Phi}^{\mathrm{T}}\mathbf{D}_2(\mathbf{x}_0)\mathbf{\Phi}\right]\mathbf{d\tilde{x}}$. Вопрос о характере условного экстремума, таким образом, сводится к вопросу о знакоопределённости редуцированной матрицы $\tilde{\mathbf{D}}_2(\mathbf{x}_0) = \mathbf{\Phi}^{\mathrm{T}}\mathbf{D}_2(\mathbf{x}_0)\mathbf{\Phi}$, имеющей размерность 2×2.

Находим эту матрицу:

$$\tilde{\mathbf{D}}_2(\mathbf{x}_0) = \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & -1 & -2 \\ -1 & 0 & -2 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Эта матрица определена положительно, так как для неё $\Delta_1 = 2 > 0$ и $\Delta_2 = 3 > 0$. Поэтому $d^2_{\text{огр}}L(\mathbf{x}_0) > 0$ и точка $\mathbf{x}_0$ является точкой условного минимума, что и требовалось условиями задачи.

В качестве другого примера рассмотрим **№ 3663(а)**, где требуется найти условный экстремум функции трёх переменных с двумя ограничениями:

$$\begin{cases} u(x, y, z) = xyz \to \max \,/\, \min \\ x^2 + y^2 + z^2 = 1 \\ x + y + z = 0 \end{cases}.$$

Составляем функцию Лагранжа, которая в данном случае будет зависеть от пяти переменных:

$$L(x, y, z, \lambda_1, \lambda_2) = xyz + \lambda_1(x^2 + y^2 + z^2 - 1) + \lambda_2(x + y + z). \tag{10}$$

Находим её частные производные и приравниваем к нулю, приходя в результате к нелинейной системе:

$$\begin{cases} yz + 2\lambda_1 x + \lambda_2 = 0 \\ xz + 2\lambda_1 y + \lambda_2 = 0 \\ xy + 2\lambda_1 z + \lambda_2 = 0 \\ \quad x^2 + y^2 + z^2 = 1 \\ \qquad x + y + z = 0 \end{cases} \tag{11}$$

n	$\mathbf{x}_0$			$\boldsymbol{\lambda}_0$	
	x	y	z	λ_1	λ_2
1	$\frac{1}{\sqrt{6}}$	$\frac{1}{\sqrt{6}}$	$-\frac{2}{\sqrt{6}}$	$\frac{1}{2\sqrt{6}}$	$\frac{1}{6}$
2	$-\frac{1}{\sqrt{6}}$	$-\frac{1}{\sqrt{6}}$	$\frac{2}{\sqrt{6}}$	$-\frac{1}{2\sqrt{6}}$	$\frac{1}{6}$
3	$\frac{1}{\sqrt{6}}$	$-\frac{2}{\sqrt{6}}$	$\frac{1}{\sqrt{6}}$	$\frac{1}{2\sqrt{6}}$	$\frac{1}{6}$
4	$-\frac{1}{\sqrt{6}}$	$\frac{2}{\sqrt{6}}$	$-\frac{1}{\sqrt{6}}$	$-\frac{1}{2\sqrt{6}}$	$\frac{1}{6}$
5	$\frac{2}{\sqrt{6}}$	$-\frac{1}{\sqrt{6}}$	$-\frac{1}{\sqrt{6}}$	$-\frac{1}{2\sqrt{6}}$	$\frac{1}{6}$
6	$-\frac{2}{\sqrt{6}}$	$\frac{1}{\sqrt{6}}$	$\frac{1}{\sqrt{6}}$	$\frac{1}{2\sqrt{6}}$	$\frac{1}{6}$

Таблица 1: К задаче №3663(а).

Она имеет шесть различных решений[7], которые приведены в таблице 1.

Из этих шести решений рассмотрим, например, первое:

$$\mathbf{x}_0 = (x_0, y_0, z_0)^{\mathrm{T}} = \left(\tfrac{1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, -\tfrac{2}{\sqrt{6}}\right)^{\mathrm{T}}, \quad \boldsymbol{\lambda}_0 = (\lambda_1, \lambda_2)^{\mathrm{T}} = \left(\tfrac{1}{2\sqrt{6}}, \tfrac{1}{6}\right)^{\mathrm{T}}.$$

Подставляя в функцию Лагранжа λ_1 и λ_2, получаем

$$L(\mathbf{x}, \boldsymbol{\lambda}_0) = xyz + \frac{x^2 + y^2 + z^2 - 1}{2\sqrt{6}} + \frac{x + y + z}{6}.$$

Составляем матрицу вторых производных:

$$\mathbf{D}_2(\mathbf{x}) = \begin{bmatrix} \frac{1}{\sqrt{6}} & z & y \\ z & \frac{1}{\sqrt{6}} & x \\ y & x & \frac{1}{\sqrt{6}} \end{bmatrix}.$$

В точке $\mathbf{x}_0$ она равна

$$\mathbf{D}_2(\mathbf{x}_0) = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Эта матрица не является знакоопределённой, так как вычисления дают $\Delta_1 = \frac{1}{\sqrt{6}} > 0$, $\Delta_2 = -\frac{1}{2} < 0$, $\Delta_3 = -\frac{3}{2\sqrt{6}} < 0$.

Чтобы учесть ограничения, возьмём от них полные дифференциалы и подставим координаты $\mathbf{x}_0$. Получим линейную систему

$$\begin{cases} \dfrac{2\,dx}{\sqrt{6}} + \dfrac{2\,dy}{\sqrt{6}} - \dfrac{4\,dz}{\sqrt{6}} = 0 \\ dx + dy + dz = 0 \end{cases}. \tag{12}$$

[7] Поясним вкратце, как они появляются. Вычитая из первого уравнения системы (11) второе, получаем $(y - x)(z - 2\lambda_1) = 0$. С учётом $x = y$ из последнего уравнения имеем $z = -2x$, тогда предпоследнее уравнение даёт $6x^2 = 1$. Отсюда получаем первое и второе решения. Остальные четыре можно найти из других комбинаций первого, второго и третьего уравнений (11).

Как известно из линейной алгебры, в такой системе два из трёх дифференциалов будут однозначно выражаться через один оставшийся, например, dy и dz через dx. Тогда систему (12) можно записать в виде

$$\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} dy \\ dz \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} dx.$$

(Здесь мы умножили первое уравнение на $\sqrt{6}$ и разделили на 2.)

Найдя, что $dy = -dx$ и $dz = 0$, можно записать в матрично-векторном виде

$$\mathbf{dx} = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} dx.$$

Вычислим редуцированную матрицу (в данном случае она имеет размерность 1×1, т.е. является скаляром) $\tilde{\mathbf{D}}_2(\mathbf{x}_0) = \mathbf{\Phi}^{\mathrm{T}} \mathbf{D}_2(\mathbf{x}_0) \mathbf{\Phi}$, где $\mathbf{\Phi} = (1, -1, 0)^{\mathrm{T}}$:

$$\tilde{\mathbf{D}}_2(\mathbf{x}_0) = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \sqrt{6}.$$

Таким образом, $d^2_{\text{огр}} L(\mathbf{x}_0, \boldsymbol{\lambda}_0) = \sqrt{6}\, dx^2 > 0$ и точка $x = y = \frac{1}{\sqrt{6}}$, $z = -\frac{2}{\sqrt{6}}$ является точкой условного минимума. Аналогично нужно рассмотреть остальные пять решений системы (11).

В заключение темы условного экстремума необходимо отметить частую студенческую ошибку, связанную с переусложнением задачи. Так, в **№ 3701** нужно найти кратчайшее расстояние между прямой $x - y - 2 = 0$ и квадратичной параболой $y = x^2$. Типичные рассуждения выглядят следующим образом: пусть точка (x_1, y_1) лежит на прямой, а точка (x_2, y_2) — на параболе (рис. 2). Тогда минимизируемой функцией является расстояние между ними $\varrho = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, а в качестве ограничений выступают сами уравнения линий $x_1 - y_1 - 2 = 0$ и $y_2 - x_2^2 = 0$. Получается следующая задача на условный экстремум:

$$\begin{cases} \varrho(x_1, y_1, x_2, y_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \to \min \\ x_1 - y_1 - 2 = 0 \\ y_2 - x_2^2 = 0 \end{cases}.$$

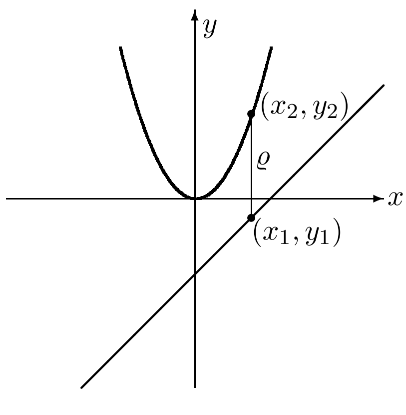

Рис. 2: к задаче № 3701.

В этом рассуждении нет ошибок и задача на условный экстремум сформулирована правильно. Однако функция Лагранжа будет зависеть от шести переменных и поиск её стационарной точки окажется весьма нетривиальной проблемой. Более хитрые студенты догадаются искать минимум не *расстояния*, а *квадрата расстояния* — задача от этого упростится, но ненамного.

Правильным будет рассматривать вместо точек (x_1, y_1) и (x_2, y_2) точки $(x_1, x_1 - 2)$ и (x_2, x_2^2). Тогда получится гораздо более простая задача на *безусловный* экстремум:

$$\varrho^2(x_1, x_2) = (x_1 - x_2)^2 + (x_1 - 2 - x_2^2)^2 \to \min .$$

Стационарная точка этой функции определяется системой

$$\begin{cases} 4x_1 - 2x_2 - 2x_2^2 = 4 \\ -2x_1 + 2x_2 - 4x_2(x_1 - 2 - x_2^2) = 0 \end{cases},$$

имеющей решение[8] $x_1 = \frac{11}{8}$, $x_2 = \frac{1}{2}$. Матрица вторых производных в этой точке равна

$$\mathbf{D}_2(\mathbf{x}_0) = \begin{bmatrix} 4 & -4 \\ -4 & \frac{15}{2} \end{bmatrix}$$

и является положительно определённой. Квадрат расстояния равен $\varrho^2(\frac{11}{8}, \frac{1}{2}) = \frac{49}{32}$ и, соответственно, $\varrho_{\min} = \frac{7}{4\sqrt{2}}$.

В [2] такой метод называется *методом исключения части переменных*; пользоваться им можно в тех случаях, когда ограничения позволяют достаточно легко выразить одни переменные через другие.

Аналогично решается и задача **№ 3654**:

$$\begin{cases} z(x, y) = xy \to \max / \min \\ x + y = 1 \end{cases}.$$

Вместо составления функции Лагранжа (она зависела бы от трёх переменных) гораздо проще записать $y = 1 - x$, и тогда z становится функцией лишь одной переменной: $z = z(x) = x - x^2$. Найти её экстремум можно даже без использования дифференциального исчисления, достаточно выделить полный квадрат:

$$z(x) = x - x^2 = \tfrac{1}{4} - \left(x - \tfrac{1}{2}\right)^2 .$$

Как видно, максимум этой функции равен $\frac{1}{4}$ и достигается при $x = \frac{1}{2}$. Отсюда $y = 1 - x = \frac{1}{2}$.

§ 2.1 (стр. 4)	§ 2.2 (стр. 5)	§ 3 (стр. 6)
3633, 3628	3651	3687, 3663(a), 3701, 3654

Таблица 2: Указатель решённых задач.

[8] Выражая из первого уравнения $x_1 = 1 + \frac{x_2}{2} + \frac{x_2^2}{2}$ и подставляя во второе уравнение, получаем $2x_2^3 - 3x_2^2 + 5x_2 = 2$, откуда $x_2 = \frac{1}{2}$ по формулам Кардано.

Литература

[1] Б. П. Демидович *Сборник задач и упражнений по математическому анализу.* — М.: АСТ, 2002.

[2] В. Ф. Бутузов и др. *Математический анализ в вопросах и задачах.* — М.: Физматлит, 2001.

[3] Л. Д. Кудрявцев *Курс математического анализа в 3 томах.* — М.: Высш. шк., 1988.