УНИВЕРСИТЕТ ИТМО

Методы оптимизации

Лекция 1. Аналитические методы определения экстремума функции

Селина Елена Георгиевна

Ауд. 302

Аналитические методы определения экстремума функции одной переменной

Сформулируем задачу оптимизации функции одной переменной на множестве вещественных чисел. Заданы множество X и функция f(x), определенная на X, требуется найти точки минимума или максимума функции f на X.

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

Функция $f(x)$ называют целевой функцией, множество X – допустимым множеством, элемент $x \in X$ – допустимой точкой задачи.

Дадим некоторые определения.

Определение 1. Точка x^* представляет глобальный минимум функции f(x) на множестве X, если $x^* \in X$ и

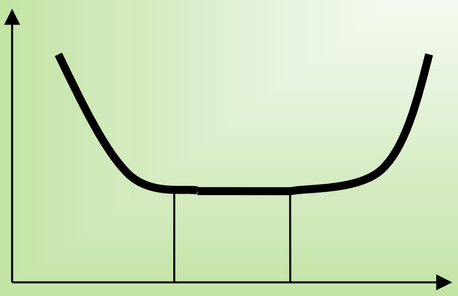
$$f(x^*) \leq f(x) \qquad (2)$$

для всех $x \in X$ (рис. 1, а).

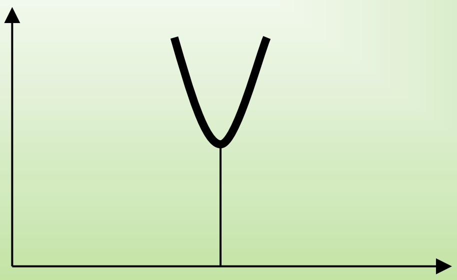
Определение 2. Точка x^* называется точкой строгого глобального минимума функции f(x) на множестве X, если $x^* \in X$ и

$$f(x^*) < f(x) \qquad (3)$$

для всех $x \in X, x \neq x^*$ (рис. 1, б).



а)



б)

Рис. 1

Определение 3. Точка x^* представляет локальный минимум функции $f(x)$на множестве X, если при некотором достаточно малом έ > 0 для всех $x \neq x^*$, $x \in X$, удовлетворяющих условию

$$|x - x^*| \leq \varepsilon$$

выполнено неравенство

$$f(x^*) \leq f(x) \qquad (4)$$

Определение 4. Если неравенство (4) – строгое, то точка x^*является точкой строгого локального минимума функции $f(x)$.

Глобальный минимум одновременно является и локальным, но не наоборот.

Все определения для максимума функции получаются заменой в выражениях (2), (3), (4) знака неравенства на обратный.

Если x^* представляет глобальный минимум функции f(x) на множестве X, будем записывать:

$$f(x^*) = \min_{x \in X} f(x)$$

Для рассмотрения вопроса о существовании решения вспомним теорему, известную из математического анализа.

Теорема 1 (*теорема Вейерштрасса).* Непрерывная функция $f(x)$ достигает на замкнутом ограниченном множестве X своего минимума (во внутренней или граничной точке).

Следующая теорема определяет необходимое условие локального экстремума.

Теорема 2. Для того, чтобы функция $f(x)$, определенная на вещественной оси, имела безусловный локальный экстремум в точке x*, необходимо, чтобы в выполнялось условие

$$\left(\frac{df}{dx}\right)_{x=x^*} = 0 \qquad (5)$$

Определение 5. Точки, в которых выполнено необходимое условие экстремума, называются **стационарными.**

Условие (5) выделяет стационарные точки, но не определяет их характера. Это может быть максимум, минимум или точка перегиба.

Для определения характера стационарной точки нам надо вспомнить ещё одну теорему.

Следующая теорема определяет достаточные условия экстремума.

Теорема 3. Для того, чтобы функция $f(x)$ имела в стационарной точке x* безусловный локальный минимум (максимум), достаточно, чтобы ее вторая производная была положительна (отрицательна):

$$\left(\frac{d^2 f}{dx^2}\right)_{x=x^*} > 0 \qquad \text{(min)}$$

$$\left(\frac{d^2 f}{dx^2}\right)_{x=x^*} < 0 \qquad \text{(max)}$$

Имеет место следующая теорема, которая определяет достаточные условия общего вида.

Теорема 4. Пусть функция одной переменной $f(x)$ в точке x* имеет непрерывные производные до k- го порядка включительно и

$$f'(x^*) = f''(x^*) = \cdots = f^{(k-1)}(x^*) = 0, \qquad f^{(k)}(x^*) \neq 0$$

Тогда, если k – четное число, то функция $f(x)$ имеет в точке x* локальный максимум при $f^{(k)}(x^*) < 0$ и локальный минимум при $f^{(k)}(x^*) > 0$. Если k нечетно, то функция $f(x)$ не имеет в точке x* ни максимума, ни минимума.

Для решения задачи минимизации функции одной переменной $f(x)$ на отрезке [a,b] в явном виде можно использовать следующий алгоритм.

Шаг 1. Решить уравнение

$f'(x) = 0$ (6)

на интервале (a;b), то есть найти все стационарные точки $x_1, x_2, \ldots, x_{1k-1} \in (a; b)$. Присвоить $x_0 = a$. $x_k = b$.

Шаг 2. В стационарных точках x_i вычислить значения $f(x_i)$ функции $f(x), i = 0, \ldots k$.

Шаг 3. Найти $f^* = \min\limits_{0 \le i \le k} f(x_i) = f(x_m)$. Присвоить $x^* = x_m$.

Пример 1. Найти минимум функции $f(x) = 3x^2 - 3x + 1$ на отрезке $x \in$ [-2;2].

Шаг 1. Найдём производную:

$$f'(x) = 3x^2 - 3$$

Решим уравнение

$$3x^2 - 3 = 0$$

на интервале (-2;2). Получим $x_1 = -1, x_2 = 1.$

Присвоим $x_0 = -2, x_3 = 2.$

Шаг 2. В стационарных точках x_i вычислить значения $f(x_i)$ функции $f(x), i = 0, \dots 3.$

$$f(x_0) = -17$$

$$f(x_1) = 3$$

$$f(x_2) = -1$$

$$f(x_3) = 1$$

Шаг 3. Находим

$$f^* = \min_{0 \le i \le 3} f(x_i) = \min(-17, 3, -1, 1) = -1 = f(x_0).$$

Таким образом, получили:

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

Аналитические методы определения экстремума функции многих переменных

Рассмотрим задачу поиска экстремума функции многих переменных

$$f(x) = f(x_1, x_2, \ldots, x_n), \qquad (7)$$

где $x \in R^n$ – n-мерный вектор.

Напомним, что совокупность всех n-мерных векторов, для которых введены операции сложения, вычитания, умножение на число, скалярного произведения называют n-мерным вещественным евклидовым пространством. Например, R^1 – это множество вещественных чисел, R^2 – двухмерное множество на плоскости, R^3 – множество представляет собой трехмерную поверхность.

Скалярное произведение векторов записывают в виде

$$x^T y = (x, y) = \sum_{i=1}^{n} x_i y_i . \qquad (8)$$

Нормой (длиной) вектора называют число

$$\|x\| = \sqrt{x^T x} = \sqrt{\sum_{k=1}^{n} x_k^{\ 2}} . \qquad (9)$$

Расстояние между точками *x* и y вычисляется по формуле

$$\rho(x, y) = \sqrt{\sum (x_k^{\ 2} - y_k^{\ 2})} . \qquad (10)$$

Для функции f(x) многих переменных точка x представляет собой вектор, f ′(x) − вектор первых производных (градиент) функции f(x),

f ′′(x) − симметричную матрицу вторых частных производных (матрицу Гессе − гессиан) функции f(x).

Необходимое условие безусловного экстремума

Рассмотрим задачу безусловной минимизации скалярной функции $f(x)$ векторного аргумента x размерности n.
Необходимое условие безусловного экстремума следует из следующей теоремы.

Теорема 5. Для того чтобы в точке x^* функция $f(x^1, \ldots, x^n)$ имела безусловный локальный экстремум, необходимо, чтобы все ее частные производные в этой точке должны быть равны нулю:

$$\frac{\partial f(x^*)}{\partial x} = 0, \qquad i = 1,2,\ldots,n \qquad (11)$$

Как и в случае функции одной переменной, точки, в которых выполнено необходимое условие экстремума, называются ***стационарными***.

Условие стационарности записывается в одной из следующих эквивалентных форм:

$$\operatorname{grad}(f(x)) = \nabla f(x) = f'(x) = 0,$$

где $\nabla f(x) = f'(x)$ – n-мерный вектор с

компонентами $\dfrac{\partial f}{\partial x_i}(x)$, i=1, ..., n, который называется градиентом функции f (x) в точке x.

Из курса математического анализа известно, что градиент функции, как вектор, определяет направление возрастания функции, а антиградиент $-\nabla f(x)$ – направление убывания функции.

Заметим, что необходимое условие экстремума (11) эквивалентно равенству нулю дифференциала функции f (x): df(x) = 0. В самом деле, если выполнено условие (2.4), то для любых dx_i, i = l, ..., n, имеем

$$df(x^0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^0)dx_i = 0 \qquad (12)$$

Справедливо и обратное утверждение, так как из последнего равенства в силу произвольности независимых приращений dx_i, i = 1, ..., n, следует, что все частные производные в точке x^0 равны нулю: $\frac{\partial f}{\partial x_i}(x^0) = 0$, i = l, ..., n

Условия (11) образуют систему n уравнений для определения n компонент вектора x. Эти уравнения могут иметь различную природу и допускать любое количество решений, в частности, не иметь ни одного.

Условие (11) будем называть необходимым условием экстремума первого порядка.

Достаточные условия экстремума функции многих переменных

После того как решение x^0 системы уравнений (11) будет найдено, необходимо еще определить характер стационарной точки*.*

Характер стационарной точки x^0 связан со знакоопределенностью матрицы Гессе $f''(x^0)$.

Матрицей Гессе (Гессианом) называется матрица вторых производных функции:

$$f''(x^0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1}(x^0) & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n}(x^0) \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1}(x^0) & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^n}(x^0) \end{bmatrix}$$

Для анализа поведения функции в точке потребуются некоторые свойства квадратичных функций.

Характер стационарной точки x^0 функции *f(x)* связан со знакоопределенностью квадратичной формы

$$\sum_{k,i=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^k}(x^0)\xi^i \xi^k = \xi^T f''(x^0)\xi. \qquad (12)$$

Квадратичная форма называется **положительно определенной**, если все ее значения при вещественных значениях переменных, не равных одновременно нулю, положительны.

$$\xi^T f''(x^0)\xi > 0 \qquad (13)$$

для любых векторов $\xi \neq 0$.

Соответственно, симметричная матрица вторых производных $f''(x^0)$ называется ***положительно определенной***, если выполнено (13). ***Отрицательно определенным*** квадратичным формам и матрицам соответствуют противоположный знак в неравенстве (13).

Справедлива следующая теорема:

Теорема 6. Для того чтобы дважды непрерывно дифференцируемая функция n переменных f (x) имела в стационарной точке x^0 безусловный локальный минимум (максимум), достаточно, чтобы матрица ее вторых производных была положительно (отрицательно) определенной.

Условия знакоопределенности квадратичных форм и матриц.

Критерий Сильвестра.

Согласно этому критерию, необходимым и достаточным условием положительной определенности квадратичной формы ($x^{\mathrm{T}} Ax$), где $A = \{a_{ij}\}$ — симметричная $n \times n$ матрица, является выполнение n неравенств:

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0.$$

Необходимым и достаточным условием отрицательной определенности квадратичной формы ($x^{T}Ax$), где $A = \{a_{ij}\}$ — симметричная $n \times n$ матрица, является выполнение n неравенств:

$$(-1)a_{11} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots$$

$$\ldots, (-1)^{n} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0.$$

Аналитические методы определения экстремума функции двух переменных

Необходимые условия экстремума функции f(x1,x2) в точке M(x_1,x_2):

$$\frac{\partial f(x^*)}{\partial x} = 0, \qquad i = 1{,}2 \qquad (14)$$

Функция f(x1,x2) имеет в данной точке максимум, если

$$\Delta = \frac{\partial^2 f(x_1 x_2)}{\partial x_1 \partial x_1} \frac{\partial^2 f(x_1 x_2)}{\partial x_2 \partial x_2} - \left[\frac{\partial^2 f(x_1 x_2)}{\partial x_1 \partial x_2}\right]^2 > 0; \qquad \frac{\partial^2 (x_1, x_2)}{\partial x_1 \partial x_1} < 0$$

и минимум, если $\Delta > 0$ и $\frac{\partial^2 (x_1, x_2)}{\partial x_1 \partial x_1} > 0$

Если $\Delta < 0$, то в точке M нет экстремума.

Если $\Delta = 0$, то функция f(x1,x2) в точке M может иметь экстремум, но может и не иметь его.

Пример. Найти экстремум функции $z = x^2 + xy + y^2 - 3x - 6y$.

Решение. Найдем частные производные первого порядка:

$$\frac{\partial z}{\partial x} = 2x + y - 3; \qquad \frac{\partial z}{\partial y} = x + 2y - 6.$$

Воспользуемся необходимым условием существования экстремума:

$$\begin{cases} \dfrac{\partial z}{\partial x} = 0; \\ \dfrac{\partial z}{\partial y} = 0. \end{cases} \qquad \begin{cases} 2x + y - 3 = 0; \\ x + 2y - 6 = 0. \end{cases}$$

Решая систему уравнений, находим координаты x и y стационарных точек: $x = 0$; $y = 3$, т. е. $M(0; 3)$.

Вычислим частные производные второго порядка и найдем их значения в точке М.

A = $\dfrac{\partial^2 z}{\partial x^2} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial z}{\partial x}\right)$ = 2;   C = $\dfrac{\partial^2 z}{\partial y^2} = \dfrac{\partial}{\partial y}\left(\dfrac{\partial z}{\partial y}\right)$ = 2;

B = $\dfrac{\partial^2 z}{\partial x\,\partial y} = 1$.

Вычислим $\Delta = AC - B^2 = 2 \cdot 2 - 1 > 0$, $A = 2 > 0$.

Следовательно, в точке М(0; 3) заданная функция имеет минимум. Значение функции в этой точке $z_{min} = -9$.

Условный экстремум функции двух переменных

Определение 6. Условным экстремумом функции $z = f(x; y)$ называется экстремум этой функции, достигнутый при условии, что переменные x и y связаны уравнением

$$\varphi(x; y) = 0$$

(уравнение связи).

Отыскание условного экстремума функции $z = f(x; y)$ можно свести к исследованию на обычный экстремум ***функции Лагранжа***

$u = f(x; y) + \lambda\varphi(x; y)$, где λ – неопределенный постоянный множитель.

Необходимые условия экстремума функции Лагранжа имеют вид

$$\begin{cases} \dfrac{\partial u}{\partial x} \equiv \dfrac{\partial f}{\partial x} + \lambda \dfrac{\partial \varphi}{\partial x} = 0, \\ \dfrac{\partial u}{\partial y} \equiv \dfrac{\partial f}{\partial y} + \lambda \dfrac{\partial \varphi}{\partial y} = 0, \\ \varphi(x; y) = 0. \end{cases} \qquad (15)$$

Из этой системы трех уравнений находят x и y – координаты точки, подозрительной на экстремум, и λ.

Достаточным условием, из которого можно выяснить характер экстремума, служит знак дифференциала второго порядка

$$d^2u(x,y) = u''_{xx}dx^2 + 2u''_{xy}dxdy + u''_{yy}dy^2 \qquad (16)$$

Если в стационарной точке d^2u >0, то функция z=f(x,y) имеет в данной точке условный минимум, если же d^2u <0, то условный максимум. Причем, $dx,\ dy$ связанны дифференциалом условия:

$$d\varphi(x,y) = \frac{\partial\varphi}{\partial x}dx + \frac{\partial\varphi}{\partial y}dy = 0$$

Есть и другой способ для определения характера экстремума. Из уравнения связи получаем:

$$\varphi'_x dx + \varphi'_y dy = 0$$

$$dy = -\frac{\varphi'_x}{\varphi'_y} dx,$$

поэтому в любой стационарной точке имеем:

$$d^2u(x,y) = u''_{xx}dx^2 + 2u''_{xy}dxdy + u''_{yy}dy^2 = d^2u(x,y)$$
$$= u''_{xx}dx^2 + 2u''_{xy}dx\left(-\frac{\varphi'_x}{\varphi'_y}dx\right) + u''_{yy}\left(-\frac{\varphi'_x}{\varphi'_y}dx\right)^2$$
$$= -\frac{dx^2}{\left(\varphi'_y\right)^2}\cdot(-\left(\varphi'_y\right)^2 u''_{xx} + 2\varphi'_x\varphi'_y u''_{xy} - (\varphi'_x)^2 u''_{yy})$$

Второй сомножитель (расположенный в скобке) можно представить в такой форме:

$$H = \begin{vmatrix} 0 & \varphi'_x & \varphi'_y \\ \varphi'_x & u''_{xx} & u''_{xy} \\ \varphi'_y & u''_{xy} & u''_{yy} \end{vmatrix} \qquad (17)$$

Если H > 0, то d^2u < 0, что указывает на условный максимум. Аналогично, при H < 0 и d^2u > 0, т.е. имеем условный минимум функции z = f(x,y).

Пример. Найти экстремум функции z = xy при условии, что x и y связаны уравнением 2x + 3y – 5 = 0.

Решение.

Рассмотрим функцию Лагранжа $u = xy + \lambda(2x + 3y - 5)$.
Имеем , $\dfrac{\partial u}{\partial x} = y + 2\lambda$ $\dfrac{\partial u}{\partial y} = x + 3\lambda$.

Из системы уравнений, определяющей необходимые условия экстремума

$$\begin{cases} \dfrac{\partial u}{\partial x} = 0; \\ \dfrac{\partial u}{\partial y} = 0; \\ \varphi(x; y) = 0. \end{cases} \qquad \begin{cases} y + 2\lambda = 0; \\ x + 3\lambda = 0; \\ 2x + 3y - 5 = 0. \end{cases}$$

находим $\lambda = -\dfrac{5}{12}$, $x = \dfrac{5}{4}$, $y = \dfrac{5}{6}$.

Выясним характер экстремума в каждой стационарной точке.

$$u''_{xx} = 0$$

$$u''_{yy} = 0$$

$$u''_{xy} = 1$$

$$d^2u(x, y) = u''_{xx}dx^2 + 2u''_{xy}dxdy + u''_{yy}dy^2$$
$$= 0 \cdot dx^2 + 2dxdy + 0 \cdot dy^2 == 2dxdy$$

Из уравнения связи 2x + 3y – 5 = 0 имеем:

$$d\varphi(x, y) = \frac{\partial}{\partial x}(2x + 3y - 5)dx + \frac{\partial}{\partial x}(2x + 3y - 5)dy$$
$$= 2dx + 3dy$$

$$2dx + 3dy = 0$$

$$dy = -\frac{2}{3}dx$$

Подставим во второй дифференциал:

$$d^2u(x,y) = 2dxdy = 2dx \cdot \left(-\frac{3}{2}dx\right) = -\frac{4}{3}dx^2 < 0$$

Следовательно, что в точке $\left(\frac{5}{4};\frac{5}{6}\right)$

функция $z = xy$ достигает наибольшего значения

$z_{max} = \frac{25}{24}.$

Спасибо за внимание!