Лекция 13. Экстремум функции нескольких переменных

Пусть функция нескольких переменных $u = f(x_1, ..., x_n)$ определена в области *D,* и точка $x^0 = (x_1^0, ..., x_n^0)$ принадлежит данной области.

Функция $u = f(x_1, ..., x_n)$ имеет *максимум* в точке x^0, если существует окрестность этой точки такая, что для всех точек $x = (x_1, ..., x_n)$ из этой окрестности выполняется неравенство $f(x^0) \geq f(x)$ и имеет *минимум*, если $f(x^0) \leq f(x)$. Максимум и минимум называют *экстремумами* функции.

Из данного определения следует, что в окрестности точки максимума приращение функции $\Delta u = f(x) - f(x^0) \leq 0$, а в окрестности точки минимума $\Delta u = f(x) - f(x^0) \geq 0$.

Теорема 1. (Необходимое условие экстремума). Пусть функция $u = f(x_1, ..., x_n)$ имеет экстремум в точке $x^0 = (x_1^0, ..., x_n^0)$. Тогда, если существуют частные производные первого порядка $\frac{\partial f}{\partial x_k}$ $(k = 1, ..., n)$ в точке x^0, то все они равны нулю в этой точке, то есть $\frac{\partial f(x^0)}{\partial x_k} = 0 \ \ (k = 1, ..., n)$.

Доказательство. Предположим, что у функции $u = f(x_1, ..., x_n)$ все переменные зафиксированы, кроме первой, то есть

$$x_2 = x_2^0, x_3 = x_3^0, ..., x_n = x_n^0.$$

Тогда функция $u = f(x_1, x_2^0, ..., x_n^0)$ станет функцией от одной переменной x_1, причем по условию теоремы эта функция имеет экстремум в точке $x_1 = x_1^0$. Известно, что если функция одной переменной имеет экстремум в какой-либо точке, то производная данной функции в этой точке равна нулю.

Таким образом, производная от функции $u = f(x_1, x_2^0, ..., x_n^0)$ по переменной x_1 в точке $x_1 = x_1^0$ равна нулю. Но эта производная совпадает с частной производной от функции $u = f(x_1, ..., x_n)$ по переменной x_1 в точке $x^0 = (x_1^0, ..., x_n^0)$. Следовательно, $\frac{\partial f(x^0)}{\partial x_1} = 0$.

Аналогично можно показать, что остальные частные производные равны нулю. Теорема доказана.

Из этой теоремы следует, что «подозрительными» на экстремум являются точки, в которых все частные производные первого порядка обращаются в нуль. Такие точки называют *стационарными*.

Отметим, что встречаются функции, в отдельных точках которых некоторые частные производные первого порядка имеют бесконечные значения или не существуют. Такие точки тоже надо исследовать на экстремум.

Итак, *для функции* $u = f(x_1, ..., x_n)$ *точкой, «подозрительной» на экстремум, является точка* x^0*, в которой выполняется одно из условий: 1) все частные производные первого порядка данной функции равны нулю, 2) некоторые частные производные данной функции равны бесконечности или не существуют.*

Но не обязательно, что точки «подозрительные» на экстремум, являются таковыми. Например, у функции двух переменных $u = x^2 y$ частные производные $u'_x = 2xy, u'_y = x^2$ равны нулю в точке (0, 0). Но в любой окрестности этой точки приращение функции $\Delta u = x^2 y - 0 = x^2 y$ может принимать и положительные и отрицательные значения. Поэтому точка (0, 0) не является точкой экстремума.

Чтобы проверить, является ли точка x^0 точкой экстремума надо проверить достаточные условия экстремума.

Напомним сначала некоторые определения из курса алгебры.

Квадратичная форма $k(x_1,...,x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \quad a_{ij} = a_{ji} \ (i, j = 1,...,n)$ называется *положительно* (*отрицательно*) *определенной*, если $k(x_1,...,x_n) > 0$ ($k(x_1,...,x_n) < 0$) для любых отличных от нуля $x_1,...,x_n$.

Квадратичная форма, являющаяся положительно или отрицательно определенной, называется также просто *определенной* квадратичной формой.

Квадратичная форма, принимающая как положительные, так и отрицательные значения, называется *неопределенной*.

Справедлива следующая теорема.

Теорема 2. (Достаточное условие экстремума). Пусть функция $u = f(x_1, ..., x_n)$ определена и имеет непрерывные производные второго порядка в некоторой окрестности точки $x^0 = (x_1^0, ..., x_n^0)$. И пусть x^0 является стационарной точкой, т.е. $\frac{\partial f(x^0)}{\partial x_k} = 0 \;\; (k = 1, ..., n)$. Тогда если квадратичная форма

$$k(dx_1, ..., dx_n) = \sum_{i,j=1}^{n} \frac{\partial^2 f(x^0)}{\partial x_i \partial x_j} dx_i dx_j \tag{1}$$

т.е. второй дифференциал $d^2 f(x^0)$, является положительно определенной (отрицательно определенной) квадратичной формой, то точка x^0 является точкой минимума (максимума). Если же квадратичная форма (1) является неопределенной, то в точке x^0 нет экстремума.

При практическом применении этой теоремы возникает вопрос: как установить, будет ли квадратичная форма (1) положительно или отрицательно определенной. Для этой цели может служить, например, *критерий Сильвестра*, доказываемый в курсе алгебры. Он состоит в следующем.

Для того, что бы квадратичная форма $k(x_1, ..., x_n) = \sum\limits_{i,j=1}^{n} a_{ij} x_i x_j, \quad a_{ij} = a_{ji}$, была положительно определенной, необходимо и достаточно, чтобы

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \quad ..., \quad \begin{vmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{2n} & ... & a_{nn} \end{vmatrix} > 0.$$

Для того, что бы квадратичная форма $k(x_1, ..., x_n) = \sum\limits_{i,j=1}^{n} a_{ij} x_i x_j, \quad a_{ij} = a_{ji}$, была отрицательно определенной, необходимо и достаточно, чтобы

$$a_{11} < 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} > 0, \ldots .$$

В нашем случае коэффициенты a_{ij} это есть частные производные второго порядка функции $u = f(x_1, ..., x_n)$ в стационарной точке $x^0 = (x_1^0, ..., x_n^0)$. Т.е. $a_{ij} = \dfrac{\partial^2 f(x^0)}{\partial x_i \partial x_j}$, $i, j = 1, ..., n$.

Сформулируем теперь теорему 2 для функции двух переменных $z = f(x, y)$, выразив условия, накладываемые на квадратичную форму (1), в явном виде через частные производные второго порядка.

Теорема 3. Пусть функция $z = f(x, y)$ определена и имеет непрерывные частные производные первого и второго порядков в окрестности точки (x_0, y_0), которая является стационарной, то есть удовлетворяет условиям $f'_x(x_0, y_0) = 0$, $f'_y(x_0, y_0) = 0$. Введем обозначения:

$$a_{11} = f''_{xx}(x_0, y_0), \quad a_{12} = f''_{xy}(x_0, y_0), \quad a_{22} = f''_{yy}(x_0, y_0).$$

Тогда в точке (x_0, y_0) функция $z = f(x, y)$

1) имеет максимум, если $a_{11}a_{22} - (a_{12})^2 > 0$, $a_{11} < 0$,

2) имеет минимум, если $a_{11}a_{22} - (a_{12})^2 > 0$, $a_{11} > 0$,

3) не имеет экстремума, если $a_{11}a_{22} - (a_{12})^2 < 0$,

4) может иметь экстремум, а может и не иметь, то есть требуется дополнительное исследование, если $a_{11}a_{22} - (a_{12})^2 = 0$.

Доказательство. Очевидно, пункты 1) и 2) следуют из критерия Сильвестра. Пусть выполнено условие $a_{11}a_{22} - (a_{12})^2 < 0$. Рассмотрим квадратичную форму

$$k(dx, dy) = d^2 f(x_0, y_0) = f''_{xx}(x_0, y_0)dx^2 + 2f''_{xy}(x_0, y_0)dxdy + f''_{yy}(x_0, y_0)dy^2 =$$
$$= a_{11}dx^2 + 2a_{12}dxdy + a_{22}dy^2.$$

При $a_{11} \neq 0$ мы можем преобразовать ее к виду:

$$k(dx, dy) = a_{11}dx^2 + 2a_{12}dxdy + a_{22}dy^2 = \frac{1}{a_{11}}\left[(a_{11}dx + a_{12}dy)^2 + (a_{11}a_{22} - a_{12}^2)dy^2\right].$$

Можно заметить, что $k(dx,dy)$ будет неопределенна, так как знак $k(dx,0)$ совпадает со знаком a_{11}, а $k(a_{12},-a_{11})$ имеет знак противоположный знаку a_{11}.

Если $a_{11}=0, a_{22}\neq 0$, то аналогичными преобразованиями можно показать неопределенность квадратичной формы.

Если $a_{11}=0, a_{22}=0$, то $k(dx,dy)=2a_{12}dxdy$. Очевидно, эта квадратичная форма тоже неопределенна, так как знак $k(dx,dy)$ будет противоположен знаку $k(dx,-dy)$. Таким образом, пункт 3) доказан.

Чтобы показать пункт 4), приведем два примера. У функции $z=(x+y)^2$ точка (0,0) является стационарной и точкой экстремума, так как $z\geq 0$ для любых x и y. Далее $z''_{xx}=z''_{xy}=z''_{yy}=2$, поэтому выполняется условие $a_{11}a_{22}-(a_{12})^2=0$. У другой функции $z=xy^3$ точка (0,0) также является стационарной, но не является точкой экстремума, поскольку в любой окрестности нуля эта функция меняет знак. Но все частные производные второго порядка в этой точке равны нулю, поэтому условие $a_{11}a_{22}-(a_{12})^2=0$ снова выполнено. Таким образом, теорема полностью доказана.

Пример. Исследовать на экстремум функцию $z=x^3+y^3-3xy$.

Найдем стационарные точки:

$$z'_x=3x^2-3y=0,\ z'_y=3y^2-3x=0 \quad\Rightarrow\quad (x_1,y_1)=(0,0),(x_2,y_2)=(1,1).$$

Тогда для первой точки $(x_1,y_1)=(0,0)$ получим

$$a_{11}=z''_{xx}(0,0)=0,\ \ a_{12}=z''_{xy}(0,0)=-3,\ \ a_{22}=z''_{yy}(0,0)=0,$$

следовательно, $a_{11}a_{22}-(a_{12})^2=-9<0$, то есть точка $(x_1,y_1)=(0,0)$ не является экстремумом.

Для второй точки $(x_2,y_2)=(1,1)$ найдем

$$a_{11}=z''_{xx}(1,1)=6,\ \ a_{12}=z''_{xy}(1,1)=-3,\ \ a_{22}=z''_{yy}(1,1)=6,$$

отсюда $a_{11}a_{22} - (a_{12})^2 = 27 > 0$, то есть точка $(x_2, y_2) = (1,1)$ является точкой минимума, так как $a_{11} > 0$.

Условный экстремум функции нескольких переменных

Многие задачи на нахождение экстремума сводятся к нахождению экстремумов функции от нескольких переменных, которые связаны между собой дополнительными условиями. В этом случае говорят о задаче нахождения *условного экстремума*.

Пусть требуется найти точки экстремума функции $u = f(x_1, ..., x_n)$ при условии, что справедливы

$$\varphi_1(x_1, ..., x_n) = 0, \ \dots, \ \varphi_m(x_1, ..., x_n) = 0, \ \ m < n. \tag{2}$$

Уравнения (2) называются *уравнениями связи*.

Точка x^0 называется точкой *условного максимума*, если существует окрестность этой точки такая, что для всех точек $x = (x_1, ..., x_n)$ из этой окрестности, удовлетворяющих уравнениям связи, выполняется неравенство $f(x^0) \geq f(x)$ и называется точкой *условного минимума*, если $f(x^0) \leq f(x)$.

Пусть функции $u = f(x_1, ..., x_n)$, $\varphi_1(x_1, ..., x_n)$, $\dots$, $\varphi_m(x_1, ..., x_n)$ непрерывно дифференцируемы в окрестности точки x^0 и в этой точке ранг *матрицы Якоби* для уравнений связи

$$\begin{pmatrix} \partial\varphi_1 / \partial x_1 & \dots & \partial\varphi_1 / \partial x_n \\ \dots & \dots & \dots \\ \partial\varphi_m / \partial x_1 & \dots & \partial\varphi_m / \partial x_n \end{pmatrix}$$

равен *m*. Напомним, что в этом случае система функций $\varphi_1(x_1, ..., x_n)$, $\dots$, $\varphi_m(x_1, ..., x_n)$ независима в окрестности точки x^0. Из определения ранга следует, что в точке x^0 хотя бы один из определителей вида

$$\begin{vmatrix} \partial\varphi_1/\partial x_{i_1} & \dots & \partial\varphi_1/\partial x_{i_m} \\ \dots & \dots & \dots \\ \partial\varphi_m/\partial x_{i_1} & \dots & \partial\varphi_m/\partial x_{i_m} \end{vmatrix}$$

отличен от нуля. Пусть для определенности это будет

$$\begin{vmatrix} \partial\varphi_1/\partial x_{n-m+1} & \dots & \partial\varphi_1/\partial x_n \\ \dots & \dots & \dots \\ \partial\varphi_m/\partial x_{n-m+1} & \dots & \partial\varphi_m/\partial x_n \end{vmatrix} \neq 0 \qquad (3)$$

Тогда по теореме о неявных функциях систему уравнений (2) можно разрешить относительно переменных $x_{n-m+1},...,x_n$:

$$x_{n-m+1} = \psi_1(x_1, ..., x_{n-m}), ..., x_n = \psi_m(x_1, ..., x_{n-m})$$

Подставляя эти равенства в функцию $u = f(x_1, ..., x_n)$, получим функцию

$$u = f(x_1,...,x_{n-m},\psi_1,...,\psi_m) = g(x_1,...,x_{n-m}) \qquad (4)$$

от $n - m$ переменных. Тогда задача о нахождении условного экстремума функции n переменных $u = f(x_1, ..., x_n)$ сводится к задаче о нахождении обычного экстремума функции $n-m$ переменных $u = g(x_1, ..., x_{n-m})$. Это один из способов решения данной задачи.

Рассмотрим другой метод, который называется *методом Лагранжа*.

Теорема 4. Пусть точка x^0 является точкой условного экстремума функции $u = f(x_1, ..., x_n)$ при выполнении уравнений связи (2). Тогда существуют такие числа $\lambda_1, \lambda_2, ..., \lambda_m$, что в точке x^0 выполняются условия

$$\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial \varphi_1}{\partial x_i} + \ldots + \lambda_m \frac{\partial \varphi_m}{\partial x_i} = 0, \quad i = 1, 2, ..., n. \qquad (5)$$

Следствие. (Необходимые условия существования условного экстремума) Введем функцию

$$L(x) = f(x_1, ..., x_n) + \sum_{i=1}^{m} \lambda_i \varphi_i(x_1, ..., x_n),$$

которая называется *функцией Лагранжа*, при этом параметры $\lambda_1, \lambda_2, ..., \lambda_m$ называются *множителями Лагранжа*. Тогда если точка x^0 является точкой условного экстремума для функции $u = f(x_1, ..., x_n)$, то она является стационарной точкой для функции Лагранжа, т.е. в этой точке выполняются равенства:

$$\begin{cases} \dfrac{\partial L(x^0)}{\partial x_k} = 0, \ k = 1, ..., n, \\ \varphi_i(x^0) = 0, \ i = 1, ..., m. \end{cases}$$

Доказательство теоремы 4. Пусть точка $x^0 = (x_1^0, ..., x_n^0)$ является точкой условного экстремума для функции f и пусть в этой точке выполняется условие (3). Тогда точка $(x_1^0, ..., x_{n-m}^0)$ является точкой обычного экстремума для функции g (см. (4)), поэтому в силу теоремы 1 все частные производные функции g равны нулю в этой точке, а, следовательно, в этой точке

$$dg(x_1, ..., x_{n-m}) = 0$$

или

$$df(x_1, ..., x_{n-m}, \psi_1, ..., \psi_m) = 0.$$

Раскрыв дифференциал, получим

$$\frac{\partial f}{\partial x_1} dx_1 + \ldots + \frac{\partial f}{\partial x_n} dx_n = 0. \qquad (6)$$

Далее, вычисляя дифференциалы от левой и правой частей каждого из равенств (2), получим

$$\frac{\partial \varphi_i}{\partial x_1} dx_1 + \ldots + \frac{\partial \varphi_i}{\partial x_n} dx_n = 0, \quad i = 1, ..., m. \tag{7}$$

В формулах (6) и (7) дифференциалы $dx_1, ..., dx_{n-m}$ суть дифференциалы независимых переменных, а дифференциалы $dx_{n-m+1}, ..., dx_n$ суть дифференциалы функций $\psi_1, ..., \psi_m$ в точке x^0. Умножив каждое равенство (7) для функции φ_i на некоторое число λ_i и сложив все эти равенства между собой и с равенством (6), получим

$$\sum_{j=1}^{n} \left(\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial x_j} \right) dx_j = 0.$$

Выберем числа $\lambda_1, \lambda_2, ..., \lambda_m$ так, чтобы в точке x^0 выполнялись равенства

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial x_j} = 0, \quad j = n - m + 1, ..., n. \tag{8}$$

Это можно сделать, так как равенства (8) представляют собой систему m линейных уравнений с неизвестными $\lambda_1, \lambda_2, ..., \lambda_m$, причем определитель матрицы системы есть определитель (3) и он отличен от нуля. При таком выборе чисел $\lambda_1, \lambda_2, ..., \lambda_m$ имеем

$$\sum_{j=1}^{n-m} \left(\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial x_j} \right) dx_j = 0.$$

Здесь уже все дифференциалы независимы и могут принимать любые значения. Беря каждый раз один из дифференциалов равным единице, а остальные равными нулю, получим равенства

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial \varphi_i}{\partial x_j} = 0, \quad j = 1, ..., n - m. \tag{9}$$

Равенства (9) совместно с (8) дают (5). Теорема доказана.

Следствие к теореме 4 позволяет найти точки подозрительные на экстремум, найдя стационарные точки функции Лагранжа. Следующая теорема, которую мы приводим без доказательства, позволяет определить, является ли указанная точка точкой условного экстремума рассматриваемой задачи.

Теорема 5. (Достаточные условия существования условного экстремума). Пусть функции $u = f(x_1, ..., x_n)$, $\varphi_1(x_1, ..., x_n)$, …, $\varphi_m(x_1, ..., x_n)$ дважды непрерывно дифференцируемы в окрестности точки x^0, эта точка удовлетворяет уравнениям связи (2) и в этой точке выполняются необходимые условия условного экстремума, т.е. она является стационарной точкой для функции Лагранжа:

$$L(x) = f(x_1, ..., x_n) + \sum_{i=1}^{m} \lambda_i \varphi_i(x_1, ..., x_n)$$

Тогда функция $u = f(x_1, ..., x_n)$ имеет условный максимум в точке x^0, если второй дифференциал функции Лагранжа в этой точке является отрицательно определенной квадратичной формой переменных $dx_1, ..., dx_n$, т.е. $d^2L(x^0) < 0$ и имеет условный минимум, если $d^2L(x^0) > 0$. При этом должны выполняться условия (7). Если знак второго дифференциала функции Лагранжа в точке x^0 является неопределенным, то в данной точке условного экстремума нет.

Пример. Исследовать функцию $z = x + 2y$ при условии $x^2 + y^2 = 5$ на условный экстремум.

Уравнение связи имеет вид: $\varphi(x, y) = x^2 + y^2 - 5 = 0$. Ранг матрицы $\begin{pmatrix}\partial\varphi/\partial x & \partial\varphi/\partial y\end{pmatrix} = \begin{pmatrix}2x & 2y\end{pmatrix}$ равен единице, поэтому можем применить метод Лагранжа для нахождения условного экстремума.

Составим функцию Лагранжа: $L(x, y, \lambda) = x + 2y + \lambda(x^2 + y^2 - 5)$. Из системы уравнений

$$\frac{\partial L}{\partial x} = 1 + 2\lambda x = 0, \quad \frac{\partial L}{\partial y} = 2 + 2\lambda y = 0, \quad x^2 + y^2 - 5 = 0$$

найдем $(x_1, y_1) = (-1, -2)$ при $\lambda_1 = 1/2$, $(x_2, y_2) = (1, 2)$ при $\lambda_2 = -1/2$.

Так как $L''_{xx} = 2\lambda,\ L''_{xy} = 0,\ L''_{yy} = 2\lambda$, получим $d^2L(x, y, \lambda) = 2\lambda(dx^2 + dy^2)$.

Тогда в точке $(x_1, y_1) = (-1, -2)$ при $\lambda_1 = 1/2$: $d^2L(-1, -2,\ 1/2) = dx^2 + dy^2 > 0$, то есть точка $(x_1, y_1) = (-1, -2)$ является точкой минимума. Аналогично, во второй точке $(x_2, y_2) = (1, 2)$ при $\lambda_2 = -1/2$ получим $d^2L(1, 2, -1/2) = -(dx^2 + dy^2) < 0$, следовательно, точка $(x_2, y_2) = (1, 2)$ является точкой максимума.