Здравствуйте!

Лекция №4

Безусловный экстремум функции многих переменных

Рассмотрим вопрос о нахождении экстремума функции $f(x)$ от n переменных $x=(x_1, x_2, \ldots, x_n)$.

Определение. Говорят, что в точке x_0 функция $f(x)$ имеет локальный максимум (минимум), если существует такой шар $R(x_0, \delta)$ с центром в точке x_0 и радиуса $\delta > 0$, что

$$\forall x \in R(x_0, \delta) \;\; x \neq x_0 \quad f(x) < f(x_0) \quad \left(f(x) > f(x_0) \right).$$

Термины «локальный максимум» и «локальный минимум» объединяют в один термин «локальный экстремум»

Необходимое условие экстремума.

Пусть в точке $x_0 = (x_1^0, x_2^0, ..., x_n^0)$ функция $f(x)$ имеет, скажем, локальный максимум. Будем отходить от точки x_0 меняя лишь координату x_1. Тогда

$$f(x_1, x_2^0, ..., x_n^0) < f(x_1^0, x_2^0, ..., x_n^0),$$

то есть $f(x)$ имеет локальный максимум по координате x_1. Но тогда, как это изучалось для функции одной переменной, в точке x_0 должно выполняться условие

$$f'_{x_1}(x_1^0, x_2^0, ..., x_n^0) = 0.$$

Аналогично можно рассуждать и в отношении всех остальных переменных. Рассматривая переменную x_i, запишем

$$f(x_1^0, ..., x_{i-1}^0, x_i, x_{i+1}^0 ..., x_n^0) < f(x_1^0, ..., x_{i-1}^0, x_i^0, x_{i+1}^0 ..., x_n^0),$$

откуда следует, что $f'_{x_i}(x_1^0, x_2^0, ..., x_n^0) = 0$

Таким образом, в точке локального экстремума должно выполняться условие

$$\frac{\partial f(x_0)}{\partial x_1} = 0;\ \frac{\partial f(x_0)}{\partial x_2} = 0;\ \ldots;\ \frac{\partial f(x_0)}{\partial x_n} = 0. \qquad (1)$$

Эти условия дают систему *n* уравнений с *n* неизвестными $x_1, x_2, ..., x_n$. Решая ее, найдем точки, «подозрительные» на экстремум, то есть точки, где **может быть** локальный максимум или минимум.

Заметим, что необходимые условия локального экстремума (1) можно коротко записать так: в точке локального экстремума должно выполняться условие

$$\operatorname{grad} f(x_0) = 0.$$

Достаточные условия локального экстремума.

Пусть выполнены условия (1). Во-первых, это еще не означает, что в точке x_0 имеет место локальный экстремум. Во-вторых, даже если там и имеется локальный экстремум, то надо установить его тип – максимум это или минимум. Ответить на этот вопрос помогают достаточные условия экстремума.

Для их вывода напомним сначала некоторые сведения из курса алгебры по вопросу о положительной и отрицательной определенности матрицы.

Пусть имеется симметричная квадратная матрица $A = \left\| a_{ij} \right\|$, в которой n строк и n столбцов. Пусть, далее, $\xi_1, \xi_2, ..., \xi_n$ есть произвольные числа. Тогда выражение

$$Q = \sum_{i,j=1}^{n} a_{ij} \xi_i \xi_j$$

называется квадратичной формой, соответствующей матрице A.

Матрица A называется **положительно определенной** матрицей, если $Q > 0$ когда хотя бы одно из чисел $\xi_1, \xi_2, ..., \xi_n$ отлично от нуля и $Q = 0$ может быть тогда и только тогда, когда все $\xi_i = 0$. Если при тех же условиях $Q < 0$, то матрица A называется **отрицательно определенной.** Если $\exists\ \xi_1, \xi_2, ..., \xi_n$ такие, что $Q > 0$ и $\exists\ \overline{\xi}_1, \overline{\xi}_2, ..., \overline{\xi}_n$ такие, что $Q < 0$, то матрица A называется **неопределенной.**

Для определения типа матрицы A существует так называемый **критерий Сильвестра.** Пусть

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{Vmatrix}.$$

Рассмотрим ее миноры

$$A_1 = a_{11}; \;\; A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \;\; A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}; \; \ldots \; ;$$

$$A_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}.$$

Если выполнено условие

$$A_1 > 0; \;\; A_2 > 0; \;\; A_3 > 0; \;\; \ldots; \;\; A_n > 0,$$

то матрица *A* положительно определенная. Если выполнено условие

$$A_1 < 0; \;\; A_2 > 0; \;\; A_3 < 0; \;\; \ldots,$$

то есть все нечетные миноры меньше нуля, а все четные – больше нуля, то матрица *A* отрицательно определенная. Во всех остальных случаях расстановки знаков > и < матрица *A* неопределенная.

Конкретизируем теперь вид матрицы A. Пусть элементы матрицы A имеют вид

$$a_{ij} = \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}, \tag{2}$$

где все производные вычисляются в точке x_0 предполагаемого экстремума.

Теорема. Если матрица A положительно определенная, то в точке x_0 локальный минимум.

Если матрица A отрицательно определенная, то в точке x_0 локальный максимум.

Если матрица A неопределенная, то в точке x_0 нет ни локального максимума, ни локального минимума.

Условный экстремум. Метод Лагранжа

В жизни всегда приходится подчиняться каким-то ограничениям, наложенным на нас природой, обществом, финансами и т.д. Математически эта ситуация формализуется как задача нахождения экстремума некоторой функции n переменных при наличии ограничений. Эти ограничения могут иметь самый разнообразный вид, но ниже будет рассмотрена лишь простейшая ситуация, когда задача выглядит так: найти максимум или минимум функции

$$f(x_1, x_2, ..., x_n) \Rightarrow \text{extr}$$

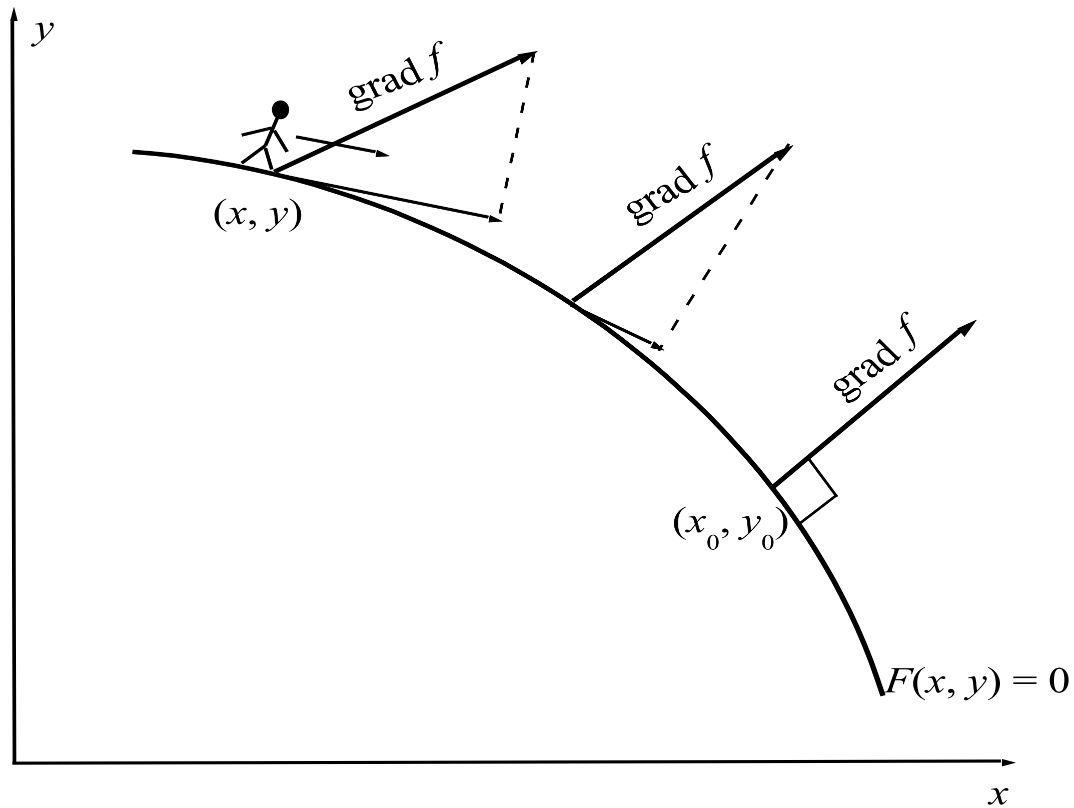
при наличии ограничений на переменные $x_1, x_2, ..., x_n$ вида

$$\begin{cases} F_1(x_1, x_2, ..., x_n) = 0, \\ F_2(x_1, x_2, ..., x_n) = 0, \\ \qquad\qquad ... \\ F_m(x_1, x_2, ..., x_n) = 0. \end{cases}$$

Наличие ограничений (условий) дает и имя этой задачи: она называется задачей на условный экстремум.

Пусть нужно найти максимум функции двух переменных $f(x, y)$ при наличии единственного ограничения $F(x, y) = 0$, то есть решить задачу

$$\begin{cases} f(x, y) \Rightarrow \max, \\ F(x, y) = 0. \end{cases}$$



Если мы желаем увеличить $f(x, y)$, мы должны двигаться по кривой в том направлении, куда указывает проекция градиента.
До каких же пор наше движение будет приводить к увеличению $f(x, y)$? Очевидно, что это увеличение окончится тогда, когда производная по направлению касательной будет равна нулю. Это произойдет тогда, когда $\operatorname{grad} f$ будет перпендикулярен касательной, то есть будет параллелен нормали к кривой.

Но вектор нормали $\vec{N} = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y} \right)$. Условие, что $\operatorname{grad} f$ коллинеарен нормали, приводит к условию

$$\frac{\partial f / \partial x}{\partial F / \partial x} = \frac{\partial f / \partial y}{\partial F / \partial y} = -\lambda .$$

Обозначая это отношение через $-\lambda$, придем к системе уравнений

$$\begin{cases} \dfrac{\partial f}{\partial x} + \lambda \dfrac{\partial F}{\partial x} = 0, \\ \dfrac{\partial f}{\partial y} + \lambda \dfrac{\partial F}{\partial y} = 0, \\ F(x, y) = 0. \end{cases}$$

В этой системе три уравнения, неизвестных также три – это x, y и λ. Так что, решив эту систему, мы получим точку, «подозрительную» на экстремум. Значение λ можно выбросить.

В общем случае рассуждения выглядят следующим образом: пусть надо решить задачу

$$f(x_1, x_2, ..., x_n) \Rightarrow \mathrm{extr},$$

$$\begin{cases} F_1(x_1, x_2, ..., x_n) = 0, \\ F_2(x_1, x_2, ..., x_n) = 0, \\ \quad\quad ... \\ F_m(x_1, x_2, ..., x_n) = 0. \end{cases}$$

Направление возрастания функции $f(x)$ указывает вектор $\mathrm{grad}\, f$. С другой стороны, условие $F_i(x_1, x_2, ..., x_n) = 0$ вырезает из нашего n-мерного пространства некоторую гиперповерхность размерности $n - 1$. Вектор нормали к этой гиперповерхности имеет вид

$$\vec{N}_i = \left(\frac{\partial F_i}{\partial x_1}, \frac{\partial F_i}{\partial x_2}, ..., \frac{\partial F_i}{\partial x_n} \right).$$

Рассмотрим линейное многообразие, образованное этими векторами, то есть совокупность векторов вида

$$\vec{N} = -\lambda_1 \vec{N}_1 - \lambda_2 \vec{N}_2 - \ldots - \lambda_m \vec{N}_m.$$

Тогда точки, «подозрительные на экстремум, должны определяться тем условием, что вектор градиента принадлежит этому многообразию, то есть $\exists \lambda_1, \lambda_2, \ldots, \lambda_m$ такие, что

$$\operatorname{grad} f = -\lambda_1 \vec{N}_1 - \lambda_2 \vec{N}_2 - \ldots - \lambda_m \vec{N}_m.$$

В компонентах это выглядит так:

$$\begin{cases} \dfrac{\partial f}{\partial x_1} + \lambda_1 \dfrac{\partial F_1}{\partial x_1} + \lambda_2 \dfrac{\partial F_2}{\partial x_1} + \ldots + \lambda_m \dfrac{\partial F_m}{\partial x_1} = 0, \\ \dfrac{\partial f}{\partial x_2} + \lambda_1 \dfrac{\partial F_1}{\partial x_2} + \lambda_2 \dfrac{\partial F_2}{\partial x_2} + \ldots + \lambda_m \dfrac{\partial F_m}{\partial x_2} = 0, \\ \ldots \\ \dfrac{\partial f}{\partial x_n} + \lambda_1 \dfrac{\partial F_1}{\partial x_n} + \lambda_2 \dfrac{\partial F_2}{\partial x_n} + \ldots + \lambda_m \dfrac{\partial F_m}{\partial x_n} = 0. \end{cases} \tag{4}$$

Добавляя сюда ограничения задачи

$$\begin{cases} F_1(x_1, x_2, ..., x_n) = 0, \\ F_2(x_1, x_2, ..., x_n) = 0, \\ \qquad ... \\ F_m(x_1, x_2, ..., x_n) = 0, \end{cases} \tag{5}$$

мы получим систему из $n + m$ уравнений относительно $n + m$ неизвестных $x_1, x_2, ..., x_n$, $\lambda_1, \lambda_2, ..., \lambda_m$, решая которую и найдем возможные точки экстремума.

Формально это выглядит так: составляется так называемая функция Лагранжа

$$L = f(x) + \lambda_1 F_1(x) + \lambda_2 F_2(x) + ... + \lambda_m F_m(x).$$

Появляющиеся здесь сомножители $\lambda_1, \lambda_2, ..., \lambda_m$ называются неопределенными множителями Лагранжа. Система уравнений, определяющая возможные точки экстремума, выглядит так

$$\begin{cases} \dfrac{\partial L}{\partial x_i} = 0, \quad i = \overline{1,n}, \\ \dfrac{\partial L}{\partial \lambda_j} = 0, \quad j = \overline{1,m}. \end{cases}$$

Уравнения $\partial L / \partial x_i = 0$ дают систему (4), уравнения $\partial L / \partial \lambda_j = 0$ – систему (5).