!Здравствуйте

Лекция №19

Ряд Тейлора функции многих переменных

Опять-таки, для простоты вывода, рассмотрим функцию двух переменных $f(x,y)$ и получим для нее ряд Тейлора относительно точки (x_0, y_0).

Напомним, что разложение в ряд Тейлора функции $f(x)$ одной переменной имеет вид:

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots$$

А теперь рассмотрим $f(x,y)$ и **зафиксируем аргумент *y*.** Тогда разложение $f(x,y)$ около точки x_0 будет иметь вид

$$f(x,y)=f(x_0,y)+\frac{f'_x(x_0,y)}{1!}(x-x_0)+\frac{f''_{xx}(x_0,y)}{2!}(x-x_0)^2+\ldots \quad (1)$$

А теперь, когда в правой части x_0 фиксировано, разложим оставшиеся функции в ряд Тейлора по y относительно точки y_0:

$$f(x_0,y)=f(x_0,y_0)+\frac{f'_y(x_0,y_0)}{1!}(y-y_0)+\frac{f''_{yy}(x_0,y_0)}{2!}(y-y_0)^2+\ldots,$$

$$f'_x(x_0,y)=f'_x(x_0,y_0)+\frac{f''_{xy}(x_0,y_0)}{1!}(y-y_0)+\ldots,$$

$$f''_{xx}(x_0,y)=f''_{xx}(x_0,y_0)+\ldots .$$

Подставим все это в (1) и соберем слагаемые с одинаковыми степенями $(x-x_0)$ и $(y-y_0)$. Тогда получим

$$f(x,y)=f(x_0,y_0)+\frac{1}{1!}\left[f'_x(x_0,y_0)(x-x_0)+f'_y(x_0,y_0)(y-y_0)\right]+$$

$$+\frac{1}{2!}\left[f''_{xx}(x_0,y_0)(x-x_0)^2+2f''_{xy}(x_0,y_0)(x-x_0)(y-y_0)+f''_{yy}(x_0,y_0)(y-y_0)^2\right]$$

В общем случае функции $f(x)$ от n переменных $x=(x_1, x_2, \dots, x_n)$ ряд Тейлора имеет вид

$$f(x) = f(x_0) + \frac{1}{1!}\sum_{i=1}^{n}\frac{\partial f(x_0)}{\partial x_i}(x_i - x_i^{(0)}) +$$

$$+\frac{1}{2!}\sum_{i,j=1}^{n}\frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}(x_i - x_i^{(0)})(x_j - x_j^{(0)}) + \dots$$

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

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

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

Необходимое условие экстремума.

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

$$f(x_1, x_2^0, \ldots, x_n^0) < f(x_1^0, x_2^0, \ldots, x_n^0),$$

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

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

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

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

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

Таким образом, в точке локального экстремума должно выполняться условие

$$\frac{\partial f(x_0)}{\partial x_1}=0;\ \frac{\partial f(x_0)}{\partial x_2}=0;\ \ldots;\ \frac{\partial f(x_0)}{\partial x_n}=0. \qquad (2)$$

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 (2) можно коротко записать так: в точке локального экстремума должно выполняться условие

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

Достаточные условия локального экстремума.

Пусть выполнены условия (2). 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 $A = \|a_{ij}\|$, в которой n строк и n столбцов. Пусть, далее, $\xi_1, \xi_2, \ldots, \xi_n$ 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

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

Рассмотрим ее миноры

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

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

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

$$a_{ij} = \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}, \tag{3}$$

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.

Доказательство. Мы докажем эту теорему не слишком строго.

Разложим $f(x)$ в ряд Тейлора в окрестности точки x_0. Обозначая $x_i - x_i^0 = \Delta x_i$ и вспоминая, что в точке x_0 имеет место $f'_{x_i}(x_0) = 0$, получим

$$f(x) = f(x_0) + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\Delta x_i \Delta x_j + \ldots,$$

где a_{ij} определены в (3).

Пусть теперь матрица $A=\left\|a_{ij}\right\|$ положительно определенная. Тогда $\sum_{i,j=1}^{n} a_{ij}\Delta x_i \Delta x_j > 0$, когда хотя бы одно из $\Delta x_i \neq 0$. Но тогда, по крайней мере в некоторой окрестности точки x_0, будет выполнено условие $f(x) > f(x_0)$ и в точке x_0 будет локальный минимум.

Если же матрица A отрицательно определенная, то $\sum_{i,j=1}^{n} a_{ij}\Delta x_i \Delta x_j < 0$, и, по крайней мере в некоторой окрестности точки x_0 будет выполнено условие $f(x) < f(x_0)$ и в точке x_0 будет локальный максимум.

Если же матрица A неопределенная, то $\exists \Delta x_i$ такие, что

$\sum_{i,j=1}^{n} a_{ij}\Delta x_i \Delta x_j > 0$, и тогда $f(x) > f(x_0)$, но и $\exists \Delta x_i$ такие, что

$\sum_{i,j=1}^{n} a_{ij}\Delta x_i \Delta x_j < 0$, и тогда $f(x) < f(x_0)$. Поэтому в данном случае в точке x_0 нет ни максимума, ни минимума.

В заключение еще раз отметим, что мы пренебрегли следующим членом в разложении $f(x)$ в ряд Тейлора. Поэтому все эти рассуждения верны лишь в некоторой окрестности точки x_0 и наши возможные минимум и максимум являются лишь локальными.