ПОХІДНІ СТАРШИХ ПОРЯДКІВ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

Означення 1. Нехай $f : A \to \mathbb{R}$, $A \subset \mathbb{R}^m$ – відкрита множина і $\forall x \in A\ \exists f'_{x_j}(x)$. Тоді **частинною похідною другого порядку** за змінними x_j, x_k в точці $x^0 \in A$ називають частинну похідну за змінною x_k від f'_{x_j} в точці x^0.

Позначення: $f''_{x_j x_k}(x^0) = f''_{jk}(x^0) = \frac{\partial^2 f}{\partial x_j\ \partial x_k}(x^0)$, $\frac{\partial^2 f}{\partial x_j^2}(x^0)$ при $j = k$.

Приклади. Знайти всі похідні другого порядку від функції $f(x_1, x_2, x_3) = x_1^2 + x_1 x_2 - x_3 \sin(x_1 x_2)$.

$$\frac{\partial^2 f}{\partial x_1^2}(x) = 2 + x_3 x_2^2 \sin(x_1 x_2);\ \frac{\partial^2 f}{\partial x_2^2}(x) = x_3 x_1^2 \sin(x_1 x_2);\ \frac{\partial^2 f}{\partial x_3^2}(x) = 0;$$

$$\frac{\partial^2 f}{\partial x_1\ \partial x_2}(x) = \frac{\partial^2 f}{\partial x_2\ \partial x_1}(x) = 1 - x_3 \sin(x_1 x_2) + x_3 x_2 x_1 \sin(x_1 x_2);$$

$$\frac{\partial^2 f}{\partial x_1\ \partial x_3}(x) = \frac{\partial^2 f}{\partial x_3\ \partial x_1}(x) = -x_2 \cos(x_1 x_2);$$

$$\frac{\partial^2 f}{\partial x_2\ \partial x_3}(x) = \frac{\partial^2 f}{\partial x_3\ \partial x_2}(x) = -x_1 \cos(x_1 x_2).$$

В цьому прикладі можна побачити, що змішані другі похідні (похідні по двох різних змінних) не залежать від порядку обчислення. Загальний факт містить така теорема.

Теорема 1. **(Про рівність змішаних похідних).** Нехай $f''_{jk}, f''_{kj} \in C(A)$. Тоді $f''_{jk}(x) = f''_{kj}(x),\ x \in A$.

Доведення. Враховуючи, що всі змінні, крім двох, при обчисленні похідних фіксуються, можна вважати, що f – функція двох змінних. Розглянемо точку $x^0 \in A$ і прирости $\triangle x_1, \triangle x_2$.

Позначимо прирости функції по різних змінних:

$$\varphi(x_1) = f(x_1, x_2^0 + \triangle x_2) - f(x_1, x_2^0)$$

$$\psi(x_2) = f(x_1^0 + \triangle x_1, x_2) - f(x_1^0, x_2).$$

Зауважимо, що

$$\varphi(x_1^0 + \triangle x_1) - \varphi(x_1^0) = \psi(x_2^0 + \triangle x_2) - \psi(x_2^0) =$$

$$= f(x_1^0 + \triangle x_1, x_2^0 + \triangle x_2) - f(x_1^0, x_2^0 + \triangle x_2) - f(x_1^0 + \triangle x_1, x_2^0) + f(x_1^0, x_2^0).$$

За теоремою Лагранжа

$$\varphi(x_1^0 + \triangle x_1) - \varphi(x_1^0) = \varphi'(x_1^0 + \xi_1)\triangle x_1,\ \xi_1 \in [0, \triangle x_1],$$

$$\psi(x_2^0 + \triangle x_2) - \psi(x_2^0) = \psi'(x_2^0 + \xi_2)\triangle x_2,\ \xi_2 \in [0, \triangle x_2].$$

Ще раз застосовуючи теорему Лагранжа, маємо

$$\varphi'(x_1^0 + \xi_1)\triangle x_1 = f'_{x_1}(x_1^0 + \xi_1, x_2^0 + \triangle x_2)\triangle x_1 - f'_{x_1}(x_1^0 + \xi_1, x_2^0)\triangle x_1 =$$
$$= f''_{x_1 x_2}(x_1^0 + \xi_1, x_2^0 + \eta_2)\triangle x_1 \triangle x_2,\ \eta_2 \in [0, \triangle x_2],$$
$$\psi'(x_2^0 + \xi_2)\triangle x_2 = f'_{x_2}(x_1^0 + \triangle x_1, x_2^0 + \xi_2)\triangle x_2 - f'_{x_2}(x_1^0, x_2^0 + \xi_2)\triangle x_2 =$$
$$= f''_{x_2 x_1}(x_1^0 + \eta_1, x_2^0 + \xi_2)\triangle x_1 \triangle x_2,\ \eta_1 \in [0, \triangle x_1].$$

Отже,

$$f''_{x_1 x_2}(x_1^0 + \xi_1, x_2^0 + \eta_2) = f''_{x_2 x_1}(x_1^0 + \eta_1, x_2^0 + \xi_2).$$

Спрямовуючи $\triangle x_1, \triangle x_2 \to 0$, і враховуючи неперервність похідних, отримаємо $f''_{x_1 x_2}(x^0) = f''_{x_2 x_1}(x^0)$.

Зауваження. Аналогічно до похідних другого порядку визначаються похідні вищих порядків. Змішані похідні старших порядків так само не залежать від порядку змінних.

Означення 2. Якщо функція має всі частинні похідні порядку n і вони неперервні на множині A, то функцію f називають n разів неперервно диференційовно.

Позначення: $f \in C^n(A)$.

Означення 3. Похідною другого порядку функції f в точці x є матриця $f''(x) = \left(\frac{\partial^2 f}{\partial x_i\, \partial x_j}(x)\right)_{i,j=1}^{m}$.

Означення 4. Нехай $f \in C^2(A)$. Диференціалом другого порядку від функції f в точці x в напрямку a є диференціал від диференціала першого порядку, тобто $d^2 f(x,a) = f''(x)a^2 = \sum\limits_{j,k=1}^{m} \frac{\partial^2 f}{\partial x_j\, \partial x_k} a_j a_k$.

Зауваження. 1. Якщо від функції $f \in C^2(A)$ порахувати похідну за напрямком a, а від отриманої функції – похідну за напрямком b, то отримаємо другу похідну за напрямками a, b, що матиме вигляд $f''_{ab}(x) = \sum\limits_{j,k=1}^{m} \frac{\partial^2 f}{\partial x_j\, \partial x_k} a_j b_k$.

2. Аналогічно визначається диференціал порядку n – це буде сума всіх частинних похідних порядку n (з врахуванням порядку змінних!), кожна з яких множиться на добуток координат a_j, кожна з яких відповідає одній зі змінних, по яких взята похідна. Позначення: $d^n f(x,a) = f^{(n)}(x)a^n$. Останній вираз можна розуміти, як добуток n-вимірної матриці на n векторів (n-лінійну форму).

Теорема 2. **(Про формулу Тейлора).** Нехай $f \in C^n(A)$, $x \in A$, $a \in \mathbb{R}^m$, $\forall t \in [0,1]\ :\ x + ta \in A$. Тоді існує $\theta \in (0,1)$ таке, що

$$f(x+a) = f(x) + f'(x)a + \frac{1}{2!}f''(x)a^2 + ... +$$

$$+\frac{1}{(n-1)!}f^{(n-1)}(x)a^{n-1} + \frac{1}{n!}f^{(n)}(x+\theta a)a^n.$$

Доведення. Розглянемо функцію

$$\varphi(t) = f(x+ta),\ t \in [0,1].$$

Тоді за теоремою про диференціювання складної функції

$$\varphi^{(k)}(t) = f^{(k)}(x+ta)a^k.$$

По формулі Тейлора для функції φ :

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2!}\varphi''(0) + ... + \frac{1}{(n-1)!}\varphi^{(n-1)}(0) + \frac{1}{n!}\varphi^{(n)}(\theta).$$

Тепер досить підставити сюди похідні функції φ.

Зауваження. Якщо функція $f \in C^\infty(A)$, і існує стала $C > 0$ така, що всі похідні порядку k обмежені величиною C^k, то функцію f можна розкласти в ряд Тейлора, перейшовши до границі при $n \to \infty$ в формулі Тейлора:

$$f(x+a) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x)a^n.$$

ЕКСТРЕМУМИ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

Нехай $f : A \to \mathbb{R},\ A \subset \mathbb{R}^m$.

Означення 1. Точку $x^0 \in A$ таку, що $\forall x \in A\ :\ f(x) \leq f(x^0)$, називають точкою **абсолютного максимума.** Значення $f(x^0)$ називають найбільшим значенням функції f на множині A.

Точку $x^0 \in A$ таку, що $\forall x \in A\ :\ f(x) \geq f(x^0)$, називають точкою **абсолютного мінімума.** Значення $f(x^0)$ називають найменшим значенням функції f на множині A.

Зауваження. Абсолютні максимум і мінімум можуть не існувати. Наприклад, $f(x) = x,\ x \in (0, 1)$. Проте якщо f – неперервна, а A – компакт, то за узагальненням теореми Вейєрштрасса вони існують.

Означення 2. Точку $x^0 \in A$ таку, що

$$\exists r > 0 \ : \ 1) B(x^0, r) \subset A; \ 2) \forall x \in B(x_0, r) \ : \ f(x) \leq f(x^0).$$

називають точкою **локального максимума.**

Точку $x^0 \in A$ таку, що

$$\exists r > 0 \ : \ 1) B(x^0, r) \subset A; \ 2) \forall x \in B(x_0, r) \ : \ f(x) \geq f(x^0).$$

називають точкою **локального мінімума.**

Точку $x^0 \in A$ таку, що

$$\exists r > 0 \ : \ 1) B(x^0, r) \subset A; \ 2) \forall x \in B(x_0, r), \ x \neq x^0 \ : \ f(x) < f(x^0).$$

називають точкою **строгого локального максимума.**

Точку $x^0 \in A$ таку, що

$$\exists r > 0 \ : \ 1) B(x^0, r) \subset A; \ 2) \forall x \in B(x_0, r), \ x \neq x^0 \ : \ f(x) > f(x^0).$$

називають точкою **строгого локального мінімума.**

Ці чотири типи точок називають **точками локального екстремума**.

Теорема 1. **(Необхідна умова локального екстремума).** Нехай x^0 – точка локального екстремума функції f. Тоді кожна частинна похідна $f'_k(x^0),\ k = \overline{1,m}$, або не існує, або рівна нулю.

Доведення. Нехай x^0 – точка локального максимума й k-та частинна похідна в цій точці існує. Тоді $\exists r > 0\ :\ B(x^0, r) \subset A$ і $\forall x \in B(x_0, r)\ :$ $f(x) \leq f(x_0)$. Покладемо

$$g_k(t) := f(x_1^0, ..., x_{k-1}^0, x_k^0 + t, x_{k+1}^0, ..., x_m^0),\ t \in (-r, r).$$

Тоді $g_k(t) \leq g_k(0),\ t \in (-r, r)$ і $g'_k(0) = f'_k(x^0)$. Функція однієї змінної g_k має локальний максимум, отже її похідна рівна нулю.

Означення 3. Внутрішню точку множини A, в якій кожна з частинних похідних функції f або не існує, або рівна нулю, називають **точкою, підозрілою на екстремум.**

Зауваження. Функція не може мати локальних екстремумів в точках, що не є підозрілими на екстремум. Проте не всі підозрілі точки є точками екстремума. Для перевірки їх екстремальності потрібні достатні умови.

Нагадаємо з алгебри такі факти.

ОЗНАЧЕННЯ 4. Нехай D – симетрична матриця $m \times m$ з дійсними елементами. Матрицю D називають **додатно визначеною,** якщо

$$\forall a \in \mathbb{R}^m,\ a \neq 0\ :\ a^t D a > 0.$$

Матрицю D називають **від'ємно визначеною,** якщо

$$\forall a \in \mathbb{R}^m,\ a \neq 0\ :\ a^t D a < 0.$$

Матрицю D називають **невизначеною,** якщо

$$\exists a, b \in \mathbb{R}^m\ :\ a^t D a > 0,\ b^t D b < 0.$$

Критерій Сильвестра. Нехай $d_1, ..., d_m$ – головні мінори матриці D. Матриця D додатно визначена тоді й лише тоді, коли $d_k > 0,\ k = \overline{1, m}$.

З їх допомогою доведемо таку теорему.

Теорема 2. **(Достатня умова локального екстремума).** Нехай $f \in C^2(A),\ A \subset \mathbb{R}^m$ – відкрита множина, $x^0 \in A$ і $\frac{\partial f}{\partial x_k}(x^0) = 0,\ k = \overline{1,m}$. Тоді:

1) якщо матриця $f''(x^0)$ додатно визначена, то x^0 – точка строгого локального мінімума;

2) якщо матриця $f''(x^0)$ від'ємно визначена, то x^0 – точка строгого локального максимума;

3) якщо матриця $f''(x^0)$ невизначена, то x^0 не є точкою екстремума.

Доведення. 1) Нехай $f''(x^0)$ додатно визначена. Тоді за критерієм Сільвестра додатні всі її головні мінори. Якщо тепер розглянути $f''(x)$, то всі елементи цієї матриці неперервні на A, отже головні мінори цієї матриці неперервні на A. Оскільки в точці x^0 всі вони додатні, то вони є додатними в деякому її околі $B(x^0, \delta)$.

Застосуємо в цьому околі формулу Тейлора при $n = 2,\ a = x - x^0$. Враховуючи, що перші похідні функції f в точці x^0 рівні нулю, маємо:

$$f(x) = f(x_0) + \frac{1}{2} f''(x^0 + \theta(x - x^0))(x - x^0)^2.$$

Оскільки $\theta \in (0,1)$, то $x^0 + \theta(x - x^0) \in B(x^0, \delta)$. Дійсно,

$$\rho(x^0 + \theta(x - x^0), x_0) = ||\theta(x - x^0)|| = \theta\rho(x, x_0) < \delta.$$

Отже, друга похідна в цій точці додатно визначена і

$$f(x) > f(x_0),\ x \in B(x_0, \delta),\ x \neq x_0.$$

2) Застосувати доведене до функції $-f$.

3) Оскільки $\exists a, b \in \mathbb{R}^m \ : \ f''(x_0)a^2 > 0, \ f''(x_0)b^2 < 0$, то в деякому околі $B(x_0, \delta)$ ці нерівності зберігаються. Тому за формулою Тейлора при $x - x_0 = \frac{\delta a}{2||a||}$ маємо $x \in B(x_0, \delta), x^0 + \theta(x - x^0) \in B(x_0, \delta)$ і

$$f(x) = f(x_0) + \frac{1}{2} f''(x^0 + \theta(x - x^0))(x - x^0)^2 =$$

$$= f(x_0) + \frac{\delta^2}{8||a||^2} f''(x^0 + \theta(x - x^0))a^2 > f(x_0)$$

і при $x - x_0 = \frac{\delta b}{2||b||}$ маємо $x \in B(x_0, \delta), x^0 + \theta(x - x^0) \in B(x_0, \delta)$ і

$$f(x) = f(x_0) + \frac{1}{2} f''(x^0 + \theta(x - x^0))(x - x^0)^2 =$$

$$= f(x_0) + \frac{\delta^2}{8||b||^2} f''(x^0 + \theta(x - x^0))b^2 < f(x_0).$$

Наслідок. Нехай $f \in C^2(A), \ A \subset \mathbb{R}^2$ – відкрита множина, $x^0 \in A$ і $\frac{\partial f}{\partial x_1}(x^0) = \frac{\partial f}{\partial x_2}(x^0) = 0$. Позначимо $\triangle_1 = f''_{11}(x^0), \ \triangle_2 = f''_{11}(x^0) f''_{22}(x_0) - f''_{12}(x_0)^2$. Тоді:

1) якщо $\triangle_1 > 0, \ \triangle_2 > 0$, то x^0 – точка строгого локального мінімума;

2) якщо $\triangle_1 < 0, \ \triangle_2 > 0$, то x^0 – точка строгого локального максимума;

3) якщо $\triangle_2 < 0$, то x^0 не є точкою екстремума.

Доведення. 3) Матриця другої похідної має два власних числа - додатне і від'ємне. Їм відповідають вектори $a, b \in \mathbb{R}^m$, для яких виконуються умови означення невизначеної матриці.

Приклади. 1. $f(x_1,x_2)=x_1^4-x_1^2+x_2^4,\ x\in\mathbb{R}^2$. Похідні всюди існують. Тому умова підозрілості на екстремум

$$\frac{\partial f}{\partial x_1}=4x_1^3-2x_1=0,\ \frac{\partial f}{\partial x_2}=4x_2^3=0,$$

звідки $(x_1,x_2)=(0,0),(\frac{1}{\sqrt{2}},0),(-\frac{1}{\sqrt{2}},0)$ – точки підозрілі на екстремум.

$$f''=\begin{pmatrix}12x_1^2-2 & 0\\ 0 & 12x_2^2\end{pmatrix}.$$

Звідси

$$d_1=12x_1^2-2,\ d_2=(12x_1^2-2)\cdot 12x_2^2.$$

В точці $(\frac{1}{\sqrt{2}},0)\ :\ d_1=4,\ d_2=24$, отже, це точка строгого локального мінімума; в точці $(-\frac{1}{\sqrt{2}},0)\ :\ d_1=4,\ d_2=24$, отже, це точка строгого локального мінімума; в точці $(0,0)\ :\ d_1=d_2=0$, отже, теорема не працює.

Використаємо в останньому випадку означення.

$$f\left(0,\frac{1}{n}\right)=\frac{1}{n^4}>0=f(0,0);\ f\left(\frac{1}{n},0\right)=\frac{1}{n^4}-\frac{1}{n^2}<0=f(0,0),\ n>1,$$

отже в довільному околі точки $(0,0)$ функція набуває значень як більших, ніж у цій точці, так і менших. Отже це не точка екстремума.

2. $f(x_1,x_2)=|x_1|+x_2^2,\ x\in\mathbb{R}^2$. Похідна по x_1 не існує при $x_1=0$. Похідна по x_2 рівна нулю при $x_2=0$. Тому єдина підозріла на екстремум точка – $(0,0)$. Це точка строгого локального мінімума, бо в інших точках функція набуває додатні значення, а $f(0,0)=0$.