ДИФЕРЕНЦІЙОВНІСТЬ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

В цьому розділі розглядатимемо функції багатьох змінних $f : A \to \mathbb{R}$, $A \subset \mathbb{R}^m$, аргументи яких належать простору $(\mathbb{R}^m, \rho)$. Ці аргументи позначатимемо $x = (x_1, x_2, ..., x_m)$.

Введемо поняття похідної такої функції. Оскільки приріст можна додавати до будь-якої змінної, маємо таке означення.

Означення 1. Нехай $f : A \to \mathbb{R}$, x^0 – внутрішня точка множини A. **Частинною похідною** за змінною x_k функції f в точці x^0 називають число

$$\lim_{\triangle x_k \to 0} \frac{f(x_1^0, ..., x_{k-1}^0, x_k^0 + \triangle x_k, x_{k+1}^0, ..., x_m^0) - f(x_1^0, ..., x_{k-1}^0, x_k^0, x_{k+1}^0, ..., x_m^0)}{\triangle x_k},$$

якщо ця границя існує.

Позначення: $\frac{\partial f}{\partial x_k}(x^0) = f'_{x_k}(x^0) = f'_k(x^0)$.

Зауваження. Використовуючи позначення

$$e_k = (0, ..., 0, \underbrace{1}_{k}, 0, ..., 0), \ k = \overline{1, m},$$

означення похідної можна написати коротше:

$$\frac{\partial f}{\partial x_k}(x^0) := \lim_{\triangle x_k \to 0} \frac{f(x^0 + \triangle x_k e_k) - f(x^0)}{\triangle x_k}.$$

Означення 2. Якщо в точці x^0 існують всі частинні похідні, то кажуть, що функція має в цій точці **похідну,** рівну вектору

$$\left(\frac{\partial f}{\partial x_1}(x^0), \frac{\partial f}{\partial x_2}(x^0), ..., \frac{\partial f}{\partial x_m}(x^0)\right).$$

Цей вектор називають також **вектором-градієнтом** функції f в точці x^0.

Позначення: $f'(x^0) = gr\vec{a}df(x^0)$.

Зауваження. Обчислення частинних похідних легко зводиться до звичайних. Дійсно, якщо потрібно обчислити частинну похідну за змінною x_k, то при цьому, за означенням, треба всі змінні, крім x_k, вважати сталими і брати похідну по цій змінній так само, як для функції одної змінної.

Звідси, зокрема, випливає правильність теорем про арифметичні дії та про похідну складної функції для частинних похідних.

Приклади. $f(x_1, x_2, x_3) = x_1 + x_2^2 + x_1x_2x_3 + x_3^{x_2}$. Маємо

$$\frac{\partial f}{\partial x_1} = 1 + x_2x_3; \ \frac{\partial f}{\partial x_2} = 2x_2 + x_1x_3 + x_3^{x_2}\ln x_3; \ \frac{\partial f}{\partial x_3} = x_1x_2 + x_2x_3^{x_2-1}.$$

Далі введемо поняття диференціалу функції багатьох змінних. Відразу зауважимо, що на відміну від функції однієї змінної, для функції багатьох змінних поняття існування похідної і диференційовності відрізняються.

Для цього нам знадобляться такі поняття.

Означення 3. **Нормою** вектора $x \in \mathbb{R}^m$ називають величину $||x|| = \left(\sum\limits_{k=1}^{m} x_k^2 \right)^{\frac{1}{2}}$, що рівна відстані від точки x до початку координат.

З означення легко побачити такі властивості норми:

1) $||x - y|| = \rho(x, y)$;

2) $\forall \alpha \in \mathbb{R}\ \forall x \in \mathbb{R}^m\ :\ ||\alpha x|| = |\alpha| \cdot ||x||$;

3) $\forall x, y \in \mathbb{R}^m\ :\ ||x + y|| \leq ||x|| + ||y||$.

Остання нерівність випливає з нерівності $\rho(x, -y) \leq \rho(x, 0) + \rho(0, -y)$.

Означення 4. Функцію $L : \mathbb{R}^m \to \mathbb{R}$ називають **лінійною,** якщо:

1) $\forall \alpha \in \mathbb{R}\ \forall x \in \mathbb{R}^m\ :\ L(\alpha x) = \alpha L(x)$;

2) $\forall x, y \in \mathbb{R}^m\ :\ L(x + y) = L(x) + L(y)$.

Теорема 1. **(Про характеризацію лінійної функції).** Кожна лінійна функція $L : \mathbb{R}^m \to \mathbb{R}$ має вигляд $L(x) = \sum\limits_{k=1}^{m} L_k x_k$, де $L_k,\ k = \overline{1, m}$ – деякі сталі.

Доведення. $L(x) = L(x_1 e_1 + x_2 e_2 + ... + x_m e_m) = L(x_1 e_1) + ... + L(x_m e_m) = x_1 L(e_1) + ... + x_m L(e_m) = L_1 x_1 + ... + L_m x_m$, де $x_k = L(e_k),\ k = \overline{1, m}$.

Приклади. Функції $f(x_1, x_2) = 3x_1 + 5x_2,\ f(x) = 4x$ – лінійні, функція $f(x) = 3x + 2$ не є лінійною!

Означення 4. Нехай $f : A \to \mathbb{R},\ A \subset \mathbb{R}^m,\ x^0$ – внутрішня точка множини A. Функцію f називають **диференційовною** в точці x^0, якщо існує лінійна функція $L(a) = \sum\limits_{k=1}^{m} L_k a_k,\ a \in \mathbb{R}^m$, така, що

$$f(x^0 + a) - f(x^0) - L(a) = o(||a||),\ ||a|| \to 0,\ \Leftrightarrow$$

$$\lim_{a \to 0} \frac{f(x^0 + a) - f(x^0) - L(a)}{||a||} = 0.$$

Функцію L називають **диференціалом** функції f в точці x^0, а її значення $L(a)$ називають **диференціалом** функції f в точці x^0 і в напрямку a і позначають $df(x^0, a)$.

Зауваження. 1. Часто при записі диференціала замість вектора $a = (a_1, ..., a_m)$ використовують вектор $dx = (dx_1, dx_2, ..., dx_m)$. Тоді $df(x^0, dx) = \sum\limits_{k=1}^{m} L_k dx_k$.

2. Диференціал дозволяє з великою точністю наблизити задану функцію лінійною в околі заданої точки. Це використовують в наближених обчисленнях, а також в геометрії, де відповідна лінійна функція дозволяє записати дотичну для графіка функції однієї змінної та дотичну площину для поверхні, яка є графіком функції двох змінних.

Спорідненим з диференціалом поняттям є поняття похідної за напрямком, що дозволяє визначити швидкість зміни функції при русі від заданої точки в заданому напрямку.

Означення 5. Нехай $f : A \to \mathbb{R},\ A \subset \mathbb{R}^m,\ x^0$ – внутрішня точка множини A. Похідною функції f в точці x^0 в напрямку a називають число $f'_a(x^0) = \lim\limits_{t \to 0} \frac{f(x^0 + ta) - f(x^0)}{t}$, якщо границя існує.

Зауваження. Якщо $a = e_k$, то $f'_a = f'_k$.

Теорема 2. **(Про вигляд диференціала та похідної за напрямком).** Нехай функція f диференційовна в точці x^0. Тоді ця функція має всі частинні похідні і похідні в усіх напрямках. При цьому

$$df(x^0, a) = f'_a(x^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x^0) a_k.$$

Доведення. Нехай $a \neq 0$. Поклавши в означенні диференціала $a = te_k$, отримаємо

$$\lim_{t \to 0} \frac{f(x^0 + te_k) - f(x^0) - L_k t}{|t|} = 0,$$

звідки $L_k = f'_k(x^0)$. Отже, формулу для диференціала доведено. Крім того, якщо в означення диференціала замість a підставити tb, $b \in \mathbb{R}^m$, $||b|| \neq 0$, - фіксоване, отримаємо

$$\lim_{t \to 0} \frac{f(x^0 + tb) - f(x^0) - L(b)t}{|t| \cdot ||b||} = 0,$$

звідки $f'_b(x^0) = L(b)$.

Зауваження. 1. Похідну за напрямком можна записати у вигляді скалярного добутку: $f'_a(x) = (f'(x), a)$.

2. Ця формула дозволяє з'ясувати сенс вектора похідної (вектора-градієнта). За нерівністю Коші-Буняковського

$$|f'_a(x)| \leq ||f'(x)|| \cdot ||a||,$$

причому рівність досягається, коли $f'(x)$ і a пропорційні. Отже, серед усіх напрямків одиничної довжини максимальна похідна за напрямком досягається в напрямку вектора-градієнта.

"Якщо треба в темній кімнаті відшукати гарячу батарею, треба весь час йти в напрямку температурного градієнту".

Теорема 3. **(Про зв'язок диференційовності та неперервності).** Якщо функція диференційовна в точці x^0, то вона неперервна в цій точці.

Доведення. $f(x) - f(x_0) = L(x - x^0) + o(||x - x^0||) \to 0,\ x \to x^0.$

Зауваження. Якщо функція має в точці x всі похідні за напрямками, то звідси не випливає, що вона диференційовна чи навіть неперервна. Наприклад, $f(x_1, x_2) = \begin{cases} 1, & x_2 = x_1^2,\ x_1 > 0; \\ 0, & \text{в інших випадках.} \end{cases}$ Тоді $f'_a(0) = 0$ для довільного напрямку, але функція розривна в т. 0.

Теорема 5. **(Достатня умова диференційовності).** Нехай виконуються умови:

1) $\exists \varepsilon > 0\ \forall z \in B(x^0, \varepsilon)\ \forall k = \overline{1, m}\ \exists \frac{\partial f}{\partial x_k}(z)$;

2) $\forall k = \overline{1, m}\ :\ \frac{\partial f}{\partial x_k} \in C(\{x^0\})$.

Тоді функція f диференційовна в точці x^0.

Доведення. Нехай $||a|| < \varepsilon$. Застосувавши до функції

$$\varphi_1(t) = f(x_1^0 + t, x_2^0 + a_2, ..., x_m^0 + a_m)$$

теорему Лагранжа на відрізку $[0, a_1]$, отримаємо:

$$f(x_1^0 + a_1, x_2^0 + a_2, ..., x_m^0 + a_m) - f(x_1^0, x_2^0 + a_2, ..., x_m^0 + a_m) =$$
$$= f_1'(x_1^0 + \xi_1, x_2^0 + a_2, ..., x_m^0 + a_m)a_1, \text{де} |\xi_1| \le |a_1|.$$

Аналогічно отримуємо рівності

$$f(x_1^0, x_2^0 + a_2, x_3^0 + a_3 ..., x_m^0 + a_m) - f(x_1^0, x_2^0, x_3^0 + a_3, ..., x_m^0 + a_m) =$$
$$= f_2'(x_1^0, x_2^0 + \xi_2, ..., x_m^0 + a_m)a_2,\ |\xi_2| \le |a_2|;$$
$$...$$
$$f(x_1^0, x_2^0, ..., x_{m-1}^0, x_m^0 + a_m) - f(x_1^0, x_2^0, x_3^0, ..., x_{m-1}^0, x_m^0) =$$
$$= f_m'(x_1^0, x_2^0, ..., x_{m-1}^0, x_m^0 + \xi_m)a_m,\ |\xi_m| \le |a_m|.$$

Права частина k-ї отриманої рівності може бути подана у вигляді

$$f_k'(x^0)a_k + (f_k'(x_1^0, ..., x_{k-1}^0, x_k^0 + \xi_k, x_{k+1}^0 + a_{k+1}, ..., x_m^0 + a_m) - f_k'(x^0))a_k =$$
$$= f_k'(x^0)a_k + o(|a_k|),\ a \to \vec{0},$$

бо дужка завдяки неперервності похідної прямує до нуля.

Додавши всі отримані рівності, отримаємо:

$$f(x^0 + a) - f(x^0) = \sum_{k=1}^{m} f_k'(x^0)a_k + o(||a||),\ a \to \vec{0}.$$

Наслідок. Нехай $f : A \to \mathbb{R},\ A \subset \mathbb{R}^m$ – відкрита множина. Якщо $\forall k = \overline{1, m}\ :\ \frac{\partial f}{\partial x_k} \in C(A)$, то f диференційовна в усіх точках множини A.

Теорема 6. **(Про арифметичні дії).** Нехай $f, g : A \to \mathbb{R}$ – диференційовні в точці $x^0 \in A$. Тоді функції $cf, f+g, fg$ диференційовні в точці x^0; якщо $g(x^0) \neq 0$, то функція $\frac{f}{g}$ диференційовна в точці x^0.

Доведення. (для частки)

$$\frac{f(x+a)}{g(x+a)} - \frac{f(x)}{g(x)} = \frac{(f(x+a) - f(x))g(x) - f(x)(g(x+a) - g(x))}{g(x+a)g(x)} =$$

$$= \frac{df(x,a) + o(||a||)}{g(x+a)} - \frac{f(x)(dg(x,a) + o(||a||))}{g(x)g(x+a)} =$$

$$= \frac{df(x,a)}{g(x)} - \frac{f(x)dg(x,a))}{g^2(x)} + o(||a||), \ a \to 0,$$

де враховано неперервність функцій f, g.

Теорема 7. **(Про диференційовність складної функції).** Нехай $f : A \to \mathbb{R}$ диференційовна в точці $x^0 \in A$. Нехай функції $g_1, ..., g_m : (t_0 - \varepsilon, t_0 + \varepsilon) \to A$ такі, що $g_1(t_0) = x_1^0, ..., g_m(t_0) = x_m^0$, і існують похідні $g_k'(t_0)$, де $t_0 \in \mathbb{R},\ \varepsilon > 0$.

Тоді складна функція $h(t) := f(g_1(t), ..., g_m(t)),\ t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ диференційовна в точці t_0, причому

$$h'(t_0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x^0) g_k'(t_0).$$

Доведення. За означенням диференційовності функції f в точці x^0 маємо

$$f(x) - f(x^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x^0)(x_k - x_k^0) + o(||x - x^0||),\ x \to x^0.$$

Покладемо тут $x_k = g_k(t_0 + \triangle t),\ x_k^0 = g_k(t_0),\ \triangle t \neq 0$. Розділивши на $\triangle t$, отримаємо

$$\frac{h(t_0 + \triangle t) - h(t_0)}{\triangle t} = \frac{f(g_1(t_0 + \triangle t), ..., g_m(t_0 + \triangle t)) - f(g_1(t_0), ..., g_m(t_0))}{\triangle t} =$$

$$= \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x^0) \frac{g_k(t_0 + \triangle t) - g_k(t_0)}{\triangle t} + \frac{o(||x - x^0||)}{\triangle t},\ \triangle t \to 0.$$

Спрямувавши $\triangle t$ до нуля, отримаємо шукане, враховуючи, що

$$|x_k - x_k^0| = |g_k(t_0 + \triangle t) - g_k(t_0)| =$$

$$= |g_k'(t_0)\triangle t + o(\triangle t)| = O(\triangle t),\ \triangle t \to 0, 1 \leq k \leq m, \Rightarrow$$

$$\Rightarrow\ ||x - x^0|| = O(\triangle t),\ \triangle t \to 0.$$

Наслідок. Якщо за умов теореми внутрішні функції залежать від кількох змінних, тобто $h(x_1, ..., x_n) := f(g_1(x_1, ..., x_n), ..., g_m(x_1, ..., x_n))$, то формула набуває вигляду

$$\frac{\partial h}{\partial x_j}(x^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial y_k}(y^0) \frac{\partial g_k}{\partial x_j}(x^0).$$