9 Pochodne cząstkowe, różniczki zupełne

9.1 Pochodne cząstkowe

Dla funkcji f wielu zmiennych możemy ustalać wartość wszystkich zmiennych z wyjątkiem jednej, np. x_j -otrzymując funkcję jednej zmiennej. Jej pochodna będzie oznaczana $\frac{\partial f}{\partial x_j}$ lub f'_{x_j} i nazywana pochodną cząstkową f względem tej zmienej. Przyjrzyjmy się sytuacji **dla dwu zmiennych,** czyli dla $f(x, y)$. Niech $P_0 = (x_0, y_0)$ będzie punktem, w którym będziemy liczyć pochodne cząstkowe. Jeśli f jest określona w otoczeniu D punktu P_0, to istnieje $\delta > 0$ taka, że gdy $|x - x_0| < \delta$, to $(x, y_0) \in D$.

Definicja. Pochodną cząstkową względem zmiennej x, to granica

$$\frac{\partial f}{\partial x}(P_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}. \tag{1}$$

Innymi słowy, jest to pochodna funkcji $f(\cdot, y_0)$ w punkcie x_0, czyli $\frac{d}{dx} f(x, y_0)|_{x=x_0}$ -choć nie zawsze ten zapis jest wygodny. Czasami wygodniejszy bywa zapis

$$\frac{\partial}{\partial x} f(x, y)|_{(x,y)=(x_0,y_0)} \quad \text{lub} \quad \frac{\partial}{\partial x} f(x, y)|_{(x_0,y_0)}.$$

Bardziej skrótowa wersja, to oznaczenie $f'_x(x_0, y_0)$, lub nawet $f_x(x_0, y_0)$ zastępujące symbol $\frac{\partial f}{\partial x}(x_0, y_0)$. Analogicznie definiujemy pochodne względem drugiej zmiennej, $\frac{\partial f}{\partial y}(x_0, y_0) = \frac{d}{dy} f(x_0, y)|_{y=y_0}$ Możemy więc stosować znane reguły liczenia pochodnych funkcji jednej zmiennej.

Dotychczas pisaliśmy x_0, y_0 -by oznaczyć w sposób szczególny te ustalone wartości, ale równie dobrze możemy nie używać dolnego indeksu -tak będzie w następnych przykładach, czy w zadaniach. Na **przykład**, dla

$$F(x, y) = x^4 + x^3 y + xy^2 + 6y + 132$$

mamy $\frac{\partial}{\partial x} F(x, y) = 4x^3 + 3x^2 y + y^2$, zaś $\frac{\partial}{\partial y} F(x, y) = x^3 + 2xy + 6$.

Podobnie,

$$\frac{\partial}{\partial x}(x^y) = yx^{y-1}, \quad \frac{\partial}{\partial y}(x^y) = x^y \ln x.$$

Podstawowe dwie zasady: liniowość i wzór Leibniza dla pochodnej iloczynu przenoszą się na pochodne cząstkowe:

$$\frac{\partial}{\partial y}(a \cdot f(x, y) + b \cdot g(x, y)) = a \cdot \frac{\partial}{\partial y} f(x, y) + b \cdot \frac{\partial}{\partial y} g(x, y),$$

$$\frac{\partial}{\partial y}(f(x, y) \cdot g(x, y)) = (\frac{\partial}{\partial y} f(x, y)) \cdot g(x, y) + f(x, y) \cdot \frac{\partial}{\partial y} g(x, y).$$

Iloraz różnicowy z definicji pochodnej cząstkowej - na przykład $\frac{\partial f}{\partial y}$ - możemy też zapisywać używając zamiast współrzędnych- zapisu wektorowego, co będzie szczególnie przydatne w przypadku większej ilości zmiennych. Gdy $e_2 = (0, 1)$ oznacza wektor z bazy kanonicznej odpowiadający zmiennej y, to

$$\frac{\partial f}{\partial y}(P_0) = \lim_{h \to 0} \frac{f(P_0 + he_2) - f(P_0)}{h}.$$

Analogiczny zapis można stosować definiując tzw. pochodną kierunkową w punkcie P_0 w kierunku wektora niezerowego w, oznaczaną $\partial^w f(P_0)$. Definicja pochodnej kierunkowej jest następująca:

$$\partial^w f(P_0) = \lim_{h \to 0} \frac{f(P_0 + hw) - f(P_0)}{h}.$$

Pochodne cząstkowe są więc pochodnymi w kierunku wektorów bazowych z bazy kanonicznej.

9.2 Przyrosty zmiennych i przyrosty wartości funkcji

Przyrostem funkcji (lub przyrostem wartości funkcji) $f : D \to \mathbb{R}$ między punktami $P, Q \in D$ nazwiemy różnicę wartości f w tych punktach, czyli **liczbę**

$$\Delta f := f(P) - f(Q).$$

Gdy punkty P, Q różnią się tylko jedną ze współrzędnych (np. x), to mówimy o przyroście f ze względu na tę zmienną, lub o przyroście częściowym. Możemy przyrost f ze względu na zmienną x oznaczyć $\Delta_x f$. Zapis ten nie jest precyzyjny, (nie zawiera informacji o punktach P, Q), ale jest dość sugestywny. Dla uproszczenia rozpatrujmy tu funkcje dwu zmiennych: (x, y). Na przykład, $\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta_x f}{\Delta x}$. W mianowniku mamy tu również przyrost funkcji: Δx możemy traktować jednak albo: jako przyrost funkcji x, gdzie $x(s,t) = s$ oznacza funkcję rzutu na oś OX, albo również- jako przyrost zmiennej niezależnej x. W tym przypadku $P = Q + (\Delta x, 0)$ lub $P = Q + \Delta x \cdot e_1$ -w zapisie wektorowym.

W przeciwnym (a raczej - w ogólnym) przypadku mówimy o przyroście całkowitym, gdzie $P = Q + (\Delta x, \Delta y)$. Podobnie, jak dla jednej zmiennej, słowo „przyrost" ma znaczenie formalne- może być zarówno $\Delta_x < 0$, jak i $\Delta_x \geqslant 0$.

Prostym, lecz bardzo ważnym narzędziem jest możliwość wyrażenia przyrostu całkowitego Δf funkcji f jako sumy przyrostów względem poszczególnych (tu 2) zmiennych:

$$f(x_1, y_1) - f(x_0, y_0) = (f(x_1, y_1) - f(x_0, y_1)) + (f(x_0, y_1) - f(x_0, y_0)),$$

co nieformalnie możemy zapisać w postaci

$$\Delta f = \Delta_x f + \Delta_y f.$$

(Dla funkcji n zmiennych Δf jest sumą $\sum_{j=1}^{n} \Delta_{x_j} f$.)

Stosując do funkcji $f(x_0, \cdot)$ zależnej od 1 zmiennej y twierdzenie Lagrange'a o wartości średniej, otrzymujemy na przykład następujący

Lemat o przyrostach *Gdy dla ustalonej wartości x_0 funkcja f ma pochodne cząstkowe $\frac{\partial}{\partial y} f(x_0, y)$, $\frac{\partial}{\partial x} f(x, y_1)$ dla wszystkich $x \in [x_0, x_1]$, $y \in [y_0, y_1]$, to istnieje $y_\beta \in (y_0, y_1)$ -taki punkt pośredni między y_0 oraz y_1, że*

$$\Delta_y f := f(x_0, y_1) - f(x_0, y_0) = (y_1 - y_0) \frac{\partial}{\partial y} f(x_0, y_\beta),$$

Przyrost całkowity, $\Delta f := f(x_1, y_1) - f(x_0, y_0)$ jest dany wzorem

$$\Delta f = (x_1 - x_0) \frac{\partial}{\partial x} f(x_\alpha, y_1) + (y_1 - y_0) \frac{\partial}{\partial y} f(x_0, y_\beta) \tag{2}$$

dla pewnych $\alpha, \beta \in (0, 1)$, gdzie $x_\alpha = x_0 + \alpha(x_1 - x_0), y_\beta = y_0 + \beta(y_1 - y_0)$.

Właśnie w takiej postaci: x_α, y_β wygodnie jest zapisywać punkty pośrednie Na przykład, $y_{\frac{1}{2}} = \frac{1}{2}(y_1 + y_0)$ jest punktem na środku odcinka łączącego y_0, y_1. Zaletą tego zapisu jest możliwość jego stosowania niezależnie od tego, czy $y_0 < y_1$, czy też $y_1 < y_0$. (Będziemy go też stosować w przyszłości do opisu punktów $P_\alpha = P_0 + \alpha(P_1 - P_0)$ z odcinka o końcach $P_0, P_1 \in \mathbb{R}^d$ parametrem $\alpha \in [0, 1]$.

Jeśli założymy istnienie obydwu pochodnych cząstkowych w otoczeniu P_0, zaś punkt P_1 zmierza do P_0, to np. z ograniczoności pochodnych cząstkowych w otoczeniu wyniknie, że Δf zmierza do zera (=ciągłość f w punkcie P_0. Gdy założymy więcej: ciągłość pochodnych cząstkowych, to w dość prosty sposób ze wzoru (2) można wywnioskować, że gdy dodatkowo funkcje $x(t), y(t)$ mają pochodne ciągłe w otoczeniu t_0, zaś $P_0 = (x(t_0), y(t_0))$, to pochodna z funkcji złożonej $t \mapsto f(x(t), y(t))$ w punkcie t_0 spełnia równość:

$$\frac{d}{dt} f(x(t), y(t))|_{t=t_0} = \frac{\partial}{\partial x} f(P_0) x'(t_0) + \frac{\partial}{\partial y} f(P_0) y'(t_0). \tag{3}$$

9.3 Dwa przykłady negatywne

Niestety, metoda ustalania zmiennych (i rozkładu przyrostów) nie jest do końca skuteczna -będziemy musieli zbadać, w jakich sytuacjach możemy ją stosować. Gdyby bowiem przyjąć, że "*różniczkowalność oznacza jedynie istnienie wszystkich pochodnych cząstkowych w każdym punkcie*", to okaże się, że złożenie takich funkcji nie musi być różniczkowalne, a nawet ciągłe! Oto przykład:

Niech $h(x, y) = \frac{xy}{x^2+y^2}$ gdy $x \neq 0$ lub $y \neq 0$. Jeśli określimy $h(0, 0) = 0$, to otrzymamy funkcję na płaszczyźnie $\mathbb{R}^2$, która jest ciągła względem każdej ze zmiennych z osobna oraz ma pochodne cząstkowe w każdym punkcie. W punkcie $(0, 0)$ pochodnych cząstkowych nie liczymy ze wzoru na pochodną ilorazu (tak możemy postąpić w innych punktach), lecz z definicji. Ponieważ $\forall_{x,y}\, h(0, y) = 0 = h(x, 0)$, bez problemu zauważamy, że $h'_x(0, 0) = h'_y(0, 0) = 0$.

Jeśli teraz $x, y : \mathbb{R} \to \mathbb{R}$ są funkcjami różniczkowalnymi jednej zmiennej t, to rozważmy funkcję złożoną: $g(t) = h(x(t), y(t))$, $g : \mathbb{R} \to \mathbb{R}$. Można się spodziewać, że złożenie funkcji "różniczkowalnych" w proponowanym powyżej sensie - będzie, jak w przypadku jednej zmiennej, funkcją różniczkowalną. Ale dla „bardzo porządnych funkcji wewnętrznych" $x(t) = t = y(t)$ nasze złożenie nie będzie nawet funkcją ciągłą! $g(t) = h(t, t) = \frac{1}{2}$ dla $t \neq 0$. natomiast dla $t = 0$ będzie $g(0) = 0$.

Gdyby natomiast w liczniku ułamka zamiast xy umieścić x^3, otrzymamy funkcję ciągłą, mającą w każdym punkcie wszystkie pochodne cząstkowe. Tym razem definiujemy $\phi(x, y) = \frac{x^3}{x^2+y^2}$ poza punkem $(0, 0)$ i z nierówności $|\phi(x, y)| \leqslant \sqrt{x^2 + y^2}$ wynika, że granica w zerze ϕ istnieje i jest równa zero. Teraz $\phi'_x(0, 0) = 1, \phi'_y(0, 0) = 0$, dla $x(t) = t = y(t)$ pochodna złożenia: $\phi(t, t) = \frac{1}{2}t$ względem t, równa $\frac{1}{2}$, nie daje się wyrazić powyżej podanym wzorem (3), który będzie obowiązywał w „przypadku regularnym". To oznacza, że musimy wprowadzić inną definicję różniczkowalności dla funkcji wielu zmiennych.

Co więcej, gdy zastosujemy "prawidłową" definicję różniczkowalności f w punkcie $(x(t), y(t))$, nie trzeba będzie zakładać dla zachodzenia wzoru (3): ani ciągłości pochodnych cząstkowych, ani nawet ich istnienia w otoczeniu tego punktu. Dowód nie będzie też wtedy używał wzoru (2).

9.4 Różniczkowalność

Przypomnijmy na wstępie, że każde odwzorowanie liniowe $L : \mathbb{R}^3 \ni (x, y, z) \to L(x, y, z) \in \mathbb{R}^k$ jest postaci $L(x, y, z) = Ax + By + Cz$, gdzie $A, B, C \in \mathbb{R}^k$ są pewnymi wektorami- a mianowicie obrazami przez L wektorów z bazy kanonicznej $\vec{i} = (1, 0, 0), \vec{j} = (0, 1, 0), \vec{k} = (0, 0, 1)$. Takie odwzorowania są ciągłe. W przypadku gdy $k = 1$, A, B, C -są to liczby rzeczywiste, zaś wartość L jest iloczynem skalarnym wektorów o współrzędnych (x, y, z) oraz (A, B, C). Zacznijmy od najprostszego przykładu: Dla funkcji dwu zmiennych o wartościach w $\mathbb{R}$, mamy wektor (A, B) wyznaczający takie odwzorowanie L. Powiemy wówczas, że wektor przyrostu zmiennej niezależnej (czyli argumentu) ma współrzędne (h, k), zaś ten wektor zmierza do zera, gdy do zera dąży jego euklidesowa norma, czyli gdy $\|(h, k)\| = \sqrt{h^2 + k^2} \to 0$.

Definicja (wariant uproszczony) Funkcja $f(x, y)$ określona w otoczeniu punktu $P_0 = (x_0, y_0)$ jest w tym punkcie różniczkowalna, gdy istnieją stałe $A, B \in \mathbb{R}$ takie, że

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - Ah - Bk}{\sqrt{h^2 + k^2}} = 0.$$

Zapisując ten ułamek jako $\alpha(h, k)$ możemy tę definicję zapisać w równoważnej postaci, gdzie $\Delta f := f(x_0 + h, y_0 + k) - f(x_0, y_0)$, zaś $\Delta P := (h, k) \in \mathbb{R}^2$:

f jest różniczkowalna w punkcie P_0 gdy stnieje odwzorowanie liniowe $L : \mathbb{R}^2 \to \mathbb{R}$ oraz funkcja α określona w otoczeniu zera taka, że

$$\Delta f := f(P_0+\Delta P)-f(P_0) = L(\Delta P)+\alpha(\Delta P)\|\Delta P\|, \quad \texttt{gdzie} \quad \lim_{\Delta P\to\vec{0}} \alpha(\Delta P) = 0.$$

Takie odwzorowanie liniowe L nazywamy różniczką (lub różniczką zupełną) odwzorowania f w punkcie P_0 i oznaczamy je symbolem $d_{P_0}f$.

Wyrażenie typu $\alpha(\Delta P)\cdot\|\Delta P\|$, w którym $\alpha(\Delta P)$ zmierza do zera nazywamy wyrażeniem typu o-małe od normy wektora ΔP, zapisując je jako $o(\|\Delta P\|)$ przy $\Delta P \to 0$. Różniczkowalność oznacza więc istnienie takiego odwzorowania liniowego L, które przybliża przyrost całkowity f z dokładnością do o-małego od normy przyrostu argumentu.

Możemy określić teraz, jak wyliczyć współczynniki A, B macierzy tego odwzorowania: Jeśli weźmiemy tylko przyrosty w kierunku osi OX, czyli gdy $\Delta P = (h, 0)$, to $\|(h, 0\| = |h|$ i warunek różniczkowalności implikuje zmierzanie do zera modułu ułamka $\alpha(h, k)$ równego

$$\left|\frac{f(x_0+h, y_0) - f(x_0, y_0) - Ah}{|h|}\right| = \left|\frac{f(x_0+h, y_0) - f(x_0, y_0)}{h} - A\right|.$$

Oznacza to dokładnie, że $A = \lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}$, czyli $A = \frac{\partial f}{\partial x}(P_0)$. Podobnie, zmierzajac do zera w kierunku OY, czyli dla $(h, k) = (0, k), k \to 0$ wnioskujemy, ze musi być $B = \frac{\partial f}{\partial y}(P_0)$. Wykazaliśmy więc, że z różniczkowalności wynika istnienie pochodnych cząstkowych i różniczka jest iloczynem skalarnym przez gradient, czyli przez wektor pochodnych cząstkowych. Definicję różniczkowalności można teraz bez trudu uogólnić na przypadek odwzorowań wielu zmiennych, o wartościach wektorowych w $\mathbb{R}^k$.

Definicja Mówimy, że odwzorowanie $F : D \subset \mathbb{R}^d \to \mathbb{R}^k$ jest różniczkowalne w punkcie $\mathbf{x}$, jeżeli jego dziedzina, D jest otoczeniem punktu $\mathbf{x}$ w $\mathbb{R}^d$ oraz istnieje takie odwzorowanie liniowe $L : \mathbb{R}^d \to \mathbb{R}^k$, że

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x}) - L(\mathbf{h})\|}{\|h\|} = 0. \tag{4}$$

Wówczas odwzorowanie L nazywamy różniczką zupełną F w punkcie $\mathbf{x}$, oznaczając $L = d_{\mathbf{x}}F$.

Przyrostowi argumentu o wektor $\mathbf{h}$ odpowiada przyrost wartości F o wektor $\Delta F := F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x})$. Różniczkowalność oznacza, że ten przyrost można „z dokładnością do $o(\|\mathbf{h}\|)$ przy $\mathbf{h} \to \mathbf{0}$" przybliżyć przez wartość pewnego liniowego odwzorowania na wektorze $\mathbf{h}$. To sformułowanie oznacza, że pozostała w wyniku tego przybliżenia reszta: $r(\mathbf{h}) := \Delta F - L(\mathbf{h})$ (błąd przybliżenia) podzielona przez $\|\mathbf{h}\|$ nadal zmierza do zera.

Następujace (równoważne) sformułowanie powyższej definicji będzie zwłaszcza przydatne przy badaniu różniczki złożenia:

Odwzorowanie F jest różniczkowalne w punkcie $\mathbf{x}$ należącum do wnętrza dziedziny D, jeśli istnieją: takie odwzorowanie liniowe $L : \mathbb{R}^d \to \mathbb{R}^k$, oraz odwzorowanie $\alpha : (\texttt{otoczenie zera w } \mathbb{R}^d) \to \mathbb{R}^k$, że

$$F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + L(\mathbf{h}) + \|h\|\alpha(h), \quad \texttt{gdzie} \quad \lim_{h\to 0}\|\alpha(h)\| = 0. \tag{5}$$

Oczywiście, gdy zarówno iloraz, jak i jego mianownik (tu $\|\mathbf{h}\|$) zmierzają do zera, to licznik też musi zmierzać do zera. Ale odwzorowania liniowe na $\mathbb{R}^d$ są zawsze ciągłe, więc $L(\mathbf{h}) \to 0$, co implikuje zmierzanie do zera samego przyrostu: ΔF, gdy $\mathbf{h} \to \mathbf{0}$ (=ciągłość F).

Wniosek. *W punktach, w których funkcja jest różniczkowalna, jest ona również ciągła.*

Z wcześniejszych przykłądów widzieliśmy, że istnienie wszystkich pochodnych cząstkowych nie gwarantuje ciągłości, nie gwarantuje też w takim razie różniczkowalności. Tak jest zresztą nawet dla f równej 0 w początku układu i równej $\frac{x^3}{x^2+y^2}$ w pozostałych punktach, chociaż w tym przypadku zarówno istnieją pochodne cząstkowe, jak i f jest ciągła. Proszę sprawdzić, z jakiego powodu nie jest ona różniczkowalna.

Przykład 1. Odwzorowanie liniowe $L : \mathbb{R}^d \to \mathbb{R}^k$ jest różniczkowalne i w każdym punkcie $\mathbf{x} \in \mathbb{R}^d$ jego różniczka jest równa L. Wynika to z równości $L(\mathbf{x}+\mathbf{h}) - L(\mathbf{x}) = L(\mathbf{h})$, więc licznik jest tu stale równy zero.