14 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 zmiennej. Przyjrzyjmy się temu dokładniej w sytuacji **funkcji dwu zmiennych,** czyli dla $f(x, y)$. Niech $P_0 = (x_0, y_0)$ będzie punktem, w którym liczymy pochodne cząstkowe. Jeśli f jest określona w otoczeniu D punktu P_0, to istnieje liczba $\delta > 0$ taka, że gdy $|x - x_0| < \delta$, to $(x, y_0) \in D$.

Definicja. Pochodną cząstkową względem zmiennej x nazywamy granicę

$$\frac{\partial f}{\partial x}(P_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}. \tag{1.14}$$

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, $\frac{\partial f}{\partial y}(x_0, y_0) = \frac{d}{dy} f(x_0, y)|_{y=y_0}$. Mamy dość jasną sytuację dzięki regułom liczenia pochodnych funkcji jednej zmiennej. Można też spotkać zapis $D_j f(P_0)$, gdzie D_j oznacza operator pochodnej względem zmiennej x_j, lub w przypadku 2 albo 3 zmiennych $D_x f$. Oczywiście, D_x jest to operator liniowy:

$$D_x(f+g)(P_0) = (D_x f)(P_0) + (D_x g)(P_0), \quad D_x(\alpha f)(P_0) = \alpha (D_x f)(P_0) \texttt{ gdy } \alpha \in \mathbb{R}.$$

Wynika to z liniowości operatora różniczkowania funkcji jednej zmiennej. Podobnie, mamy wzór Leibniza np. dla pochodnej $\frac{\partial}{\partial y}$ z iloczynu funkcji:

$$D_y(fg)(P_0) = (D_y f)(P_0) g(P_0) + f(P_0)(D_y g)(P_0).$$

Dotychczas pisaliśmy x_0, y_0, P_0 -by oznaczyć w sposób szczególny punkt, w którym liczymy pochodne, ale równie dobrze możemy nie używać dolnego indeksu "zero" -tak będzie znacznie wygodniej w następnych wyrażeniach.
Przykłady: (1) Dla

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

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

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

(3) Dla $\varphi(x, y) = xy \exp(\frac{x}{y})$ mamy[1]

$$\frac{\partial \varphi}{\partial y} = x \exp(\frac{x}{y}) \left(1 + y \frac{\partial}{\partial} \left(\frac{x}{y}\right)\right) = \frac{x(y - x)}{y} \exp(\frac{x}{y}).$$

(4) Jeżeli $F(x, y) = 0$ dla $xy = 0$, czyli na osiach $0X$, OY oraz $F(x, y) = 1$ w pozostałych punktach $\mathbb{R}^2$, to $\frac{\partial F}{\partial x}(0, 0)$ oraz $\frac{\partial F}{\partial y}(0, 0)$ istnieją i są równe zero. Ale F nie jest ciągła w zerze. Jest więc inaczej, niż w przypadku 1-wymiarowym.

14.1 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**

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

[1] Przyponijmy, że $\exp(z) := e^z$.

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 (jak również Δf) 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 całkowity funkcji x, gdzie $x(s,t) = s$ oznacza funkcję rzutu na oś OX, albo również- jako przyrost zmiennej niezależnej x. W tym przypadku $P = Q + (\Delta x, 0)$ -w zapisie wektorowym w $\mathbb{R}^2$.

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 $\Delta_x \geqslant 0$.

Rozkład przyrostu całkowitego na przyrosty w kierunach osi. Prostym, lecz bardzo ważnym narzędziem jest możliwość wyrażenia przyrostu całkowitego Δf funkcji f jako sumy przyrostów względem poszczególnych zmiennych:

$$f(x_1, y_1) - f(x_0, y_0) = \big(f(x_1, y_1) - f(x_0, y_1)\big) + \big(f(x_0, y_1) - f(x_0, y_0)\big), \quad (2.14)$$

co nieformalnie możemy zapisać w postaci

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

(Dla funkcji n zmiennych Δf jest sumą $\sum_{j=1}^{n} \Delta_{x_j} f$.) Geometryczny sens tej operacji jest taki, że zamiast wzdłuż najkrótszej drogi: po odcinku o końcach $(x_0, y_0), (x_1, y_1)$ -poruszamy się po łamanej o odcinkach równoległych do osi układu współrzędnych. Możemy to zrobić na dwa sposoby (wybierając jako punkt pośredni albo (x_0, y_1), albo (x_1, y_0)).

Uwaga o zapisie punktów pośrednich. Dla punktów t_0, t_1 na osi liczbowej $\mathbb{R}$ dowolny punkt leżący na odcinku pomiędzy nimi wygodnie będzie zapisywać w postaci $t_\alpha := t_0 + \alpha(t_1 - t_0)$, gdzie $\alpha \in [0,1]$. Ten zapis nie wymaga wiedzy o tym, która z liczb t_0, t_1 jest większa- co jest istotne dla przedziału. Ponadto taki zapis nadaje się dobrze do przypadku wektorowego:

$$P_\alpha = P_0 + \alpha(P_1 - P_0) \quad (3.14)$$

jest parametryzacją odcinka o końcach $P_0, P_1 \in \mathbb{R}^d$, gdy α przebiega odcinek $[0,1] \subset \mathbb{R}$. Gdy $\alpha = 0$, to $P_\alpha = P_0$, a dla $\alpha = 1$ otrzymamy "wyjściowe" P_1.

Stosując do funkcji $f(x_0, \cdot)$ zależnej od 1 zmiennej y i różniczkowalnej na odcinku pomiędzy punktami y_0, y_1 twierdzenie Lagrange'a o wartości średniej, otrzymujemy następujący

Lemat o przyrostach częściowych *Istnieje y_β -taki punkt pośredni między y_0, y_1, odpowiadający penej wartości $\beta \in (0,1)$, że*

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

czyli

$$\Delta_y f = (\Delta y) \cdot \frac{\partial}{\partial y} f(x_0, y_\beta). \quad (4.14)$$

Analogicznie zapiszemy przyrost $f(\cdot, y_1)$ względem x, czyli $\Delta_x f = f(x_1, y_1) - f(x_0, y_1)$ w postaci $\Delta_x f = (\Delta x) \cdot \frac{\partial}{\partial x} f(x_\alpha, y_1)$, gdzie $\alpha \in (0,1)$.

Jeśli założymy istnienie obydwu pochodnych cząstkowych (tu oznaczanych jako f'_x, f'_y w większym zbiorze, np. zawierającym prostokąt o wierzchołkach $P_0, P_1, (x_0, y_1), (x_1, y_0)$ to otrzymamy z rozkładu $\Delta f = \Delta_x f + \Delta_y f$ przyrostu całkowitego funkcji na przyrosty względem poszczególnych zmiennych następujący lemat

Lemat o przyrostach całkowitych Przyrost całkowity, $\Delta f := f(x_1, y_1) - f(x_0, y_0)$ jest równy liczbie

$$\Delta f = (x_1 - x_0) f'_x(x_0 + \alpha(x_1 - x_0), y_1) + (y_1 - y_0) f'_y(x_0, y_0 + \beta(y_1 - y_0))$$

dla pewnych $\alpha, \beta \in (0,1)$. W wersji skróconej
$$\Delta f = f'_x(x_\alpha, y_1)\Delta_x + f'_y(x_0, y_\beta)\Delta_y. \tag{5.14}$$

Załóżmy na chwilę, że pochodne cząstkowe f'_x, f'_y są ciągłe, zaś $x(t), y(t)$ są wartościami funkcji $x(\cdot), y(\cdot)$ zmiennej t, różniczkowalnych w otoczeniu odcinka $[t_0, t_1]$, o pochodnych ciągłych. Jeżeli ponadto $x(t_j) = x_j$ dla $j = 1, 2$, to ze wzoru (5.14) po zastąpieniu Δx przez $x'(t_a)\Delta t$ oraz Δy przez $y'(t_b)\Delta t$ dla pewnych $a, b \in (0,1)$ wynika, że dla funkcji złożonej $u(t) := f(x(t), y(t))$ jej iloraz różnicowy można zapisać w postaci:
$$\frac{u(t_1) - u(t_0)}{t_1 - t_0} = f'_x(x_\alpha)x'(t_a) + f'_y(y_\beta)y'(t_b).$$
Gdy $t_1 \to t_0$, to punkty pośrednie t_a, t_b zmierzają do t_0, zaś $x_a \to x_0$, $y_b \to y_0$ Z ciągłości pochodnych cząstkowych f oraz z ciągłości x', y' wynika, że istnieje granica prawej strony ostatniego wzoru. Wówczas dla $x_j = x(t_j), j = 1, 2$ mamy
$$u'(t_0) = f'_x(x_0, y_0)x'(t_0) + f'_y(x_0, y_0)y'(t_0) \qquad (6.14).$$

Ten wzór będzie wykazany w znacznie ogólniejszej sytuacji jako tzw. *Reguła łańcucha*. Za chwilę zobaczymy, że samo istnienie pochodnych cząstkowych nie wystarcza do zachodzenia równości (6.14). Będzie to wskazywać na potrzebę innej definicji różniczkowalności.

14.2 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 płaszczyzny $\mathbb{R}^2$. 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 poniżej podanym wzorem (1), który będzie obowiązywał w „przypadku regularnym". To oznacza, że musimy wprowadzić inną definicję różniczkowalności dla funkcji wielu zmiennych. Jak zauważyliśmy, założenie ciągłości pochodnych cząstkowych funkcji f oraz pochodnych: x', y' względem zmiennej t pozwala na otrzymanie wspomnianego wzoru (6.14):
$$\frac{d}{dt} f(x(t), y(t)) = f'_x(x(t), y(t))x'(t) + f'_y(x(t), y(t))y'(t).$$

Co więcej, gdy zastosujemy "prawidłową" definicję różniczkowalności f w punkcie $(x(t), y(t))$, nie trzeba będzie zakładać dla zachodzenia wzoru (6.14): ani ciągłości pochodnych cząstkowych, ani nawet ich istnienia w otoczeniu tego punktu. Dowód nie będzie też wtedy używał wzoru (2.14), ani (5.14).