Pochodne wyższych rzędów. (1)

Do tej pory rozważaliśmy funkcje (odwzorowania) $f : \mathbb{R}^k \supset O \longrightarrow \mathbb{R}^l$. Gdy f jest różniczkowalna w $a \in O$ to $f(a+h) = f(a) + f'(a) \cdot h + r(a,h)$ gdzie $f'(a) \in L(\mathbb{R}^k, \mathbb{R}^l)$ oraz $r(a,h) \in \mathbb{R}^l$ spełnia $\lim\limits_{\|h\| \to 0} \frac{\|r(a,h)\|}{\|h\|} = 0$.

Symbol $L(\mathbb{R}^k, \mathbb{R}^l)$ oznacza zbiór odwzorowań liniowych z p-ni wektorowej $\mathbb{R}^k$ do p-ni wektorowej $\mathbb{R}^l$.

Przykład 1 Niech $f : \mathbb{R}^3 \to \mathbb{R}$ będzie dana wzorem $f(x,y,z) = x^2 + x\cos(yz) + z\exp(x^2)$.

Obliczmy pochodne cząstkowe f

$$\frac{\partial f}{\partial x}(x,y,z) = 2x + \cos(yz) + 2xz\exp(x^2)$$

$$\frac{\partial f}{\partial y}(x,y,z) = -xz\sin(yz)$$

$$\frac{\partial f}{\partial z}(x,y,z) = -xy\sin(yz) + \exp(x^2)$$

Widać, że funkcje (pochodne cząstkowe) $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} : \mathbb{R}^3 \longrightarrow \mathbb{R}$ są ciągłe dla wszystkich $a \in \mathbb{R}^3$. Korzystając z twierdzenia udowodnionego na poprzednim wykładzie wnioskujemy, że f jest różniczkowalna we wszystkich punktach $a \in \mathbb{R}^3$ oraz

$f'(a) \in L(\mathbb{R}^3, \mathbb{R}) \simeq M_{1\times 3}(\mathbb{R})$ jest macierzą (2)
o $l=1$ wierszu i $k=3$ kolumnach:

$$f'(a) = \left[\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a), \frac{\partial f}{\partial z}(a)\right].$$

Możemy zacząć myśleć o pochodnej drugiej f. Pochodna pierwsza f' jest odwzorowaniem $f' : \mathcal{O} \to L(\mathbb{R}^3, \mathbb{R})$ nie ma więc postaci $g : \mathcal{O} \to \mathbb{R}^l$ ale jest "prawie" tej postaci gdyż mamy utożsamienie $L(\mathbb{R}^3, \mathbb{R}) \simeq M_{1\times 3}(\mathbb{R}) \simeq \mathbb{R}^3$.

Warto trochę uogólnić samo pojęcie pochodnej do sytuacji w której mamy $f : \mathbb{R}^k \supset \mathcal{O} \to V$ gdzie V jest przestrzenią unormowaną. W naszym przypadku V będzie postaci $V = L(\mathbb{R}^k, \mathbb{R}^l)$ gdzie na przestrzeni V jest dana norma $\|\cdot\|_V$ wzorem $\|A\|_V = \sup_{\substack{x \in \mathbb{R}^k \\ \|x\| \leq 1}} \|Ax\|$, a $A : \mathbb{R}^k \to \mathbb{R}^l$ jest liniowe tzn $A \in L(\mathbb{R}^k, \mathbb{R}^l)$.

Powiemy, że $f : \mathbb{R}^k \supset \mathcal{O} \to V$ jest różniczkowalna w $a \in \mathcal{O}$ jeśli istnieje odwzorowanie liniowe $A \in L(\mathbb{R}^k, V)$ t. że Ah jest resztą:
$r(a,h) = f(a+h) - f(a) - Ah$ jest resztą:

$$\lim_{\|h\|_{\mathbb{R}^k} \to 0} \frac{\|r(a,h)\|_V}{\|h\|_{\mathbb{R}^k}} = 0.$$

Wróćmy do naszego przykładu 1.

$f'(x,y,z) \in L(\mathbb{R}^3, \mathbb{R})$ jest postaci

$$f'(x,y,z) = \left[\frac{\partial f}{\partial x}(x,y,z), \frac{\partial f}{\partial y}(x,y,z), \frac{\partial f}{\partial z}(x,y,z)\right].$$

Czy $f' : \mathbb{R}^3 \longrightarrow L(\mathbb{R}^3, \mathbb{R})$ jest różniczkowalna w $a \in \mathbb{R}^3$ i jak myśleć o $(f')'(a)$.

Po pierwsze zamiast pisać $(f')'(a)$ piszemy $f''(a)$.

Po drugie pochodne cząstkowe pochodnych cząstkowych oznaczamy odpowiednio $\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j} f\right)(a) \overset{\text{ozn}}{=} \frac{\partial^2}{\partial x_i \partial x_j} f(a) \overset{\text{ozn}}{=}$

$\overset{\text{ozn}}{=} f_{x_j x_i}(a)$

Zauważmy, że w naszym przykładzie $f : \mathbb{R}^3 \to \mathbb{R}$ drugie pochodne cząstkowe $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ zależą w ciągły sposób od a. Na przykład

$$\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial}{\partial x}\left(-xy\sin(yz) + \exp(x^2)\right) = -y\sin(yz) + 2x\exp(x^2)$$

czyli odwzorowanie $f' : \mathbb{R}^3 \longrightarrow M_{1\times 3}(\mathbb{R})$,

$$f'(x,y,z) = \left[\frac{\partial f}{\partial x}(x,y,z), \frac{\partial f}{\partial y}(x,y,z), \frac{\partial f}{\partial z}(x,y,z)\right]$$ jest różniczkowalne. Pytanie: jak myśleć o $f''(a) \in L(\mathbb{R}^3, M_{1\times 3}(\mathbb{R}))$?

Zamiast „zawracać sobie głowę" tym konkretnym przykładem zajmijmy się ogólną sytuacją. Niech więc $O \subset \mathbb{R}^k$ będzie otwartym podzbiorem i $f: O \to \mathbb{R}^l$ będzie różniczkowalną funkcją we wszystkich punktach O. Funkcja $f': O \to L(\mathbb{R}^k, \mathbb{R}^l)$ jest różniczkowalna w punkcie $a \in O$ jeśli istnieje odwzorowanie liniowe $A \in L(\mathbb{R}^k, L(\mathbb{R}^k, \mathbb{R}^l))$ takie, że $f'(a+h) - f'(a) - Ah$ jest resztą, to znaczy $\frac{\|f'(a+h) - f'(a) - Ah\|}{\|h\|} \xrightarrow[\|h\| \to 0]{} 0$

Uwaga w liczniku mamy normę odwzorowania liniowego $f'(a+h) - f'(a) - Ah \in L(\mathbb{R}^k, \mathbb{R}^l)$ a w mianowniku mamy normę wektora $h \in \mathbb{R}^k$.

Jak myśleć o $A \in L(\mathbb{R}^k, L(\mathbb{R}^k, \mathbb{R}^l))$?

Ustalmy $h \in \mathbb{R}^k$. Wówczas $A(h) \in L(\mathbb{R}^k, \mathbb{R}^l)$.

Ustalmy $\tilde{h} \in \mathbb{R}^k$. Wówczas $(A(h))(\tilde{h}) \in \mathbb{R}^l$.

Od tej pory $(A(h))(\tilde{h})$ oznaczać będziemy symbolem $A(h, \tilde{h})$. Jest jasne, że

① $A(\alpha_1 h_1 + \alpha_2 h_2, \tilde{h}) = \alpha_1 A(h_1, \tilde{h}) + \alpha_2 A(h_2, \tilde{h})$ oraz

② $A(h, \alpha_1 \tilde{h}_1 + \alpha_2 \tilde{h}_2) = \alpha_1 A(h, \tilde{h}_1) + \alpha_2 A(h, \tilde{h}_2)$ dla

$\alpha_1, \alpha_2 \in \mathbb{R}$, $h, \tilde{h}, h_1, h_2, \tilde{h}_1, \tilde{h}_2 \in \mathbb{R}^k$

Podsumowując: druga pochodna $f''(a)$ jest odwzorowaniem 2-liniowym z $\mathbb{R}^k \times \mathbb{R}^k$ w $\mathbb{R}^l$, czyli odwzorowaniem z $\mathbb{R}^k \times \mathbb{R}^k$ do $\mathbb{R}^l$ spełniającym ① i ② powyżej. Notacja $L(\mathbb{R}^k, \mathbb{R}^k; \mathbb{R}^l)$

Wróćmy do przykładu 1.

$f: \mathbb{R}^3 \to \mathbb{R}$, $\quad f': \mathbb{R}^3 \longrightarrow L(\mathbb{R}^3, \mathbb{R})$, $\quad f'': \mathbb{R} \to L(\mathbb{R}^3; \mathbb{R}^3; \mathbb{R})$

$f(a) \in \mathbb{R}$ $\quad f'(a) \in L(\mathbb{R}^3, \mathbb{R})$, $\quad f''(a) \in L(\mathbb{R}^3, \mathbb{R}^3; \mathbb{R})$.

Ustalmy $h = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$ oraz $\tilde{h} = \begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \\ \tilde{h}_3 \end{bmatrix}$

$$\left(f''(a)\right)(h, \tilde{h}) \simeq \left(f'(a+h) - f'(a)\right)(\tilde{h}) =$$

$$\left[\frac{\partial f}{\partial x}(a+h) - \frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a+h) - \frac{\partial f}{\partial y}(a), \frac{\partial f}{\partial z}(a+h) - \frac{\partial f}{\partial z}\right] \begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \\ \tilde{h}_3 \end{bmatrix}$$

$$\simeq \left[\frac{\partial^2 f}{\partial x^2}(a) h_1 + \frac{\partial^2 f}{\partial y \partial x}(a) h_2 + \frac{\partial^2 f(a)}{\partial z \partial x} h_3, \frac{\partial^2 f(a)}{\partial x \partial y} h_1 + \frac{\partial^2 f(a)}{\partial y^2} h_2 + \frac{\partial^2 f}{\partial z \partial y} h_3, \right.$$

$$\left. \frac{\partial^2 f(a)}{\partial x \partial z} h_1 + \frac{\partial^2 f}{\partial y \partial z}(a) h_2 + \frac{\partial^2 f}{\partial z^2}(a) h_3\right] \begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \\ \tilde{h}_3 \end{bmatrix} =$$

$$[h_1, h_2, h_3] \begin{bmatrix} \frac{\partial^2 f(a)}{\partial x^2}, & \frac{\partial^2 f}{\partial x \partial y}(a), & \frac{\partial^2 f(a)}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x}(a), & \frac{\partial^2 f}{\partial y^2}(a), & \frac{\partial^2 f}{\partial y \partial z}(a) \\ \frac{\partial^2 f}{\partial z \partial x}(a), & \frac{\partial^2 f}{\partial z \partial y}(a), & \frac{\partial^2 f}{\partial z^2}(a) \end{bmatrix} \begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \\ \tilde{h}_2 \end{bmatrix} \quad (*)$$

Czyli w przykładzie 1 2-ga pochodna[10] $f''(a)$ w działaniu na parę wektorów $h, \tilde{h} \in \mathbb{R}^3$ daje skalar dany przez (*).

Ogólniej, jeśli $f : \mathbb{R}^k \supset \mathcal{O} \to \mathbb{R}$ ma drugą pochodną w punkcie $a \in \mathcal{O}$ to

$$(f''(a))(h, \tilde{h}) = \sum_{i,j=1}^{k} h_i \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \tilde{h}_j =$$

$$[h_1, \ldots, h_k] \left[\frac{\partial^2 f(a)}{\partial x_i \partial x_j} \right] \begin{bmatrix} \tilde{h}_1 \\ \vdots \\ \tilde{h}_k \end{bmatrix}$$

Macierz $\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right]_{i,j=1,\ldots,k}$ reprezentuje $f''(a)$ - więcej na algebrze.

Obserwacja dotycząca przykładu 1

$$\frac{\partial^2 f}{\partial x \partial z} = -y \sin(yz) + 2x \exp(x^2)$$

$$\frac{\partial^2 f}{\partial z \partial x} = \frac{\partial}{\partial z}\left(2x + \cos yz + 2xz \exp(x^2)\right) = -y \sin(yz) + 2x \exp(x^2)$$

Czyli pochodne mieszane są w tym przypadku równe. Można sprawdzić analogiczne równości pochodnych mieszanych drugiego rzędu funkcji f.

Czy zatem zawsze $\frac{\partial f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j \partial x_i}(a)$.

Nie zawsze; przykład verte Peano.

Przykład (Peano)

Niech $f(x,y) = \begin{cases} xy\,\dfrac{x^2-y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = 0. \end{cases}$

Dla $(x,y) \neq (0,0)$ funkcja f jest różniczkowalna gdyż ma ciągłe pochodne cząstkowe.

Ponadto $|f(h_1,h_2) - f(0,0)| = \left| h_1 h_2 \dfrac{h_1^2 - h_2^2}{h_1^2+h_2^2} \right| \leq |h_1 h_2| \leq$
$\frac{1}{2}(h_1^2+h_2^2)$ a więc

$$\frac{\left| f(h_1,h_2) - f(0,0) - [0,0]\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \right|}{\sqrt{h_1^2+h_2^2}} \leq \frac{1}{2}\sqrt{h_1^2+h_2^2} \xrightarrow[\|h\| \to 0]{} 0$$

Wniosek f różniczkowalna w $(0,0)$ i pochodna $f'(0,0) = [0,0]$. W szczególności $\frac{\partial f}{\partial x}(0,0) = 0 = \frac{\partial f}{\partial y}(0,0)$

Obliczmy 2-pochodne mieszane f w $(0,0)$

$$\frac{\partial^2}{\partial y \partial x} f(0,0) = \lim_{h\to 0} \frac{\frac{\partial f}{\partial x}(0,h) - \overbrace{\frac{\partial f}{\partial x}(0,0)}^{=0}}{h} = \lim_{h \to 0} \frac{\frac{\partial f}{\partial x}(0,h)}{h} =$$

$$\lim_{h\to 0} \frac{\left[y\,\frac{x^2-y^2}{x^2+y^2} + x\cdot y \cdot \left(\frac{2x}{x^2-y^2} - \frac{x^2-y^2}{(x^2+y^2)^2} \cdot 2x \right) \right] \Big|_{\substack{x=0 \\ y=h}}}{h} = \lim_{h\to 0} \frac{-h}{h} = -1$$

$\#$

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{h \to 0} \frac{\frac{\partial f}{\partial y}(h,0) - \frac{\partial f}{\partial y}(0,0)}{h} = \lim_{h\to 0} \frac{\frac{\partial f}{\partial y}(h,0)}{h} =$$

pochodne mieszane nie są równe.

$$\lim_{h\to 0} \frac{1}{h}\left[\frac{x(x^2-y^2)}{x^2+y^2} + xy\left(\frac{-2y}{x^2+y^2} - \frac{x^2-y^2}{x^2+y^2}\, 2y \right) \right] \Big|_{\substack{x=h \\ y=0}} = \lim_{h\to 0} \frac{h}{h} = +1$$

Twierdzenie (Schwarza o równości pochodnych mieszanych)

Jeśli funkcja $f: \mathbb{R}^k \supset \mathcal{O} \to \mathbb{R}^l$ jest klasy C^1 na $\mathcal{O}$ 8 (to znaczy pochodne $\frac{\partial f}{\partial x_j}$ $i=1,\dots,k$ istnieją i są ciągłe) oraz dla pewnego i oraz j istnieje na $\mathcal{O}$ $\frac{\partial^2 f}{\partial x_i \partial x_j}$ i jest ciągła to pochodna $\frac{\partial^2 f}{\partial x_j \partial x_i}$ istnieje na $\mathcal{O}$ oraz $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Dowód: Bez straty ogólności zał. że $\boxed{k=2,\ l=1 \\ i=1,\ j=2}$

Ustalmy $\begin{bmatrix} x \\ y \end{bmatrix}$ oraz $\begin{bmatrix} x \\ y_0 \end{bmatrix} \in \mathcal{O}$. Wówczas $\underbrace{}_{\text{podcałkowa}}$

$f(x,y) = f(x,y_0) + \int_{y_0}^{y} \frac{\partial}{\partial y} f(x,t)\,dt$ funkcja $\frac{\partial}{\partial y} f(x,y)$ jest ciągła

całka z parametrem x.

oraz pochodna po parametrze $\frac{\partial^2}{\partial xy} f(x,y)$ - ciągła

Korzystając z tw. o różniczkowaniu całki z parametrem widzimy, że

$$\frac{\partial}{\partial x} f(x,y) - \frac{\partial f}{\partial x}(x,y_0) = \int_{y_0}^{y} \frac{\partial^2}{\partial x \partial y} f(x,t)\,dt.$$ Tak więc

$$\frac{\frac{\partial}{\partial x} f(x,y) - \frac{\partial f}{\partial x}(x,y_0)}{y-y_0} = \frac{1}{y-y_0} \int_{y_0}^{y} \left(\frac{\partial^2}{\partial x \partial y} f\right)(x,t)\,dt$$ Biorąc

$y \to y_0$ i korzystając z zasadniczego tw. rach. całkowego widzimy, że lewa strona ma granicę $\frac{\partial}{\partial y}\frac{\partial}{\partial x} f(x,y_0)$ równą $\frac{\partial^2 f}{\partial x \partial y}(x,y_0)$. $\blacksquare$