Pochodne wyższego rzędu

Zacznijmy od pochodnych cząstkowych, dla których sprawa jest prosta.
Niech $f: \mathcal{U} \to \mathbb{R}^m$ ma w $\mathcal{U}$ pochodną cząstkową $\frac{\partial f}{\partial x_i}: \mathcal{U} \to \mathbb{R}^m$. Ta, jako funkcja z $\mathcal{U}$ w $\mathbb{R}^m$, też może mieć pochodne cząstkowe – w pewnych punktach $\mathcal{U}$, albo i wszędzie.
Jeżeli zatem funkcja $\frac{\partial f}{\partial x_i}: \mathcal{U} \to \mathbb{R}^m$ ma w $a \in \mathcal{U}$ ~~p~~ j-tą pochodną cząstkową, to nazywamy ją drugą, (j,i)-tą pochodną cząstkową funkcji f w a i oznaczamy

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = D_j D_i f(a) = f_{x_i x_j}(a) = \dots$$

Analogicznie definiujemy pochodne cząstkowe 3-go i wyższych rzędów.
Jak łatwo zauważyć, dla $f: \mathcal{U} \subset \mathbb{R}^3 \to \mathbb{R}^m$ możemy badać
3 pochodne cząstkowe pierwszego rzędu

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$$

9 pochodnych cząstkowych drugiego rzędu

$$\frac{\partial^2 f}{\partial x \partial x} \quad \frac{\partial^2 f}{\partial x \partial y} \quad \frac{\partial^2 f}{\partial x \partial z} \quad \frac{\partial^2 f}{\partial y \partial x} \quad \frac{\partial^2 f}{\partial y^2} \quad \frac{\partial^2 f}{\partial y \partial z}$$

$\parallel$ ozn

$$\frac{\partial^2 f}{\partial x^2} \qquad \frac{\partial^2 f}{\partial z \partial x} \quad \frac{\partial^2 f}{\partial z \partial y} \quad \frac{\partial^2 f}{\partial z^2}$$

27 pochodnych cząstkowych rzędu 3 itd...
Sporo.

A co z wyższymi pochodnymi / różniczkami
$f: \underset{\mathbb{R}^n}{\overset{\cap}{\mathcal{U}}} \to \mathbb{R}^m$?

Jeżeli f jest różniczkowalna w $\mathcal{U}$, to

$Df : \mathcal{U} \longrightarrow \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$

to jest przestrzeń liniowa izomorficzna z $\mathbb{R}^{n \cdot m}$ (gdy wybierzemy bazy w $\mathbb{R}^n$ i $\mathbb{R}^m$, możemy mówić o macierzy homomorfizmu liniowego z $\mathbb{R}^n$ w $\mathbb{R}^m$, macierz to element $\mathbb{R}^{n \cdot m}$)

Możemy zatem zdefiniować $D(Df)$:

$$D(Df) : \mathcal{U} \longrightarrow \mathrm{Hom}(\mathbb{R}^n, \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)).$$

104) Niech $x \in \mathcal{U}$.

$D(Df)(x)$, jest przekształceniem liniowym z $\mathbb{R}^n$ w $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$:

$$D^2 f(x) \overset{ozn}{=} D(Df)(x) \in \mathrm{Hom}(\mathbb{R}^n, \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m))$$

Jeżeli zatem ustalimy $v \in \mathbb{R}^n$, to

$D^2 f(x)(v) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, więc dla $w \in \mathbb{R}^n$

$D^2 f(x)(v) w \in \mathbb{R}^m$.

$D^2 f(x)(v) w$ zależy w sposób liniowy zarówno od v, jak i od w, jest więc przekształceniem DWULINIOWYM z $\mathbb{R}^n$ w $\mathbb{R}^m$

Nie ma powodu też w szczególny sposób w zapisie odróżniać v od w (choć warto x, względem którego $D^2 f$ nie musi być liniowe). Dlatego będziemy pisać

$D^2 f(x)(v, w)$

$$D^2 f(x) : \mathbb{R}^n \times \mathbb{R}^n \xrightarrow{\text{dwulin.}} \mathbb{R}^m$$

102)

W ten sam sposób możemy definiować
$D^3 f$ jako pochodną przekształcenia
$D^2 f : \mathcal{U} \to \mathrm{Hom}(\mathbb{R}^n, \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)) \approx \mathbb{R}^{n \cdot n \cdot m}$

i ogólnie
$D^k f$ jako pochodną $D^{k-1} f : \mathcal{U} \to \mathrm{Hom}(\mathbb{R}^n, \underbrace{\mathrm{Hom}(\mathbb{R}^n, \mathrm{Hom}(\ldots, \mathbb{R}^m)))}_{\cong \ \mathbb{R}^{n^{k-1} m}}$

Wtedy $D^k f(x)$, jeżeli istnieje,
możemy utożsamić z przekształceniem k-liniowym
z $\mathbb{R}^n$ w $\mathbb{R}^m$:

$$D^k f(x)(v_1, v_2, \ldots, v_k) = w \in \mathbb{R}^m$$

$v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ liniowe względem każdego z argumentów $v_1, \ldots, v_k$.

~~Dokład~~ Niech T będzie przekształceniem k-liniowym z $\mathbb{R}^n$ w $\mathbb{R}^m$ $\left(T : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{k \text{ razy}} \longrightarrow \mathbb{R}^m \right)$.

Dokładnie tak samo, jak dla przekształceń ~~...~~ dwuliniowych pokażemy, że T (a dokładniej jego funkcje współrzędne $T_1, T_2, \ldots, T_m$) jest wielomianem stopnia k względem współrzędnych swoich argumentów, jest więc ciągłe.

103)

Dla dowolnych $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$

$$T(v_1, v_2, \ldots, v_k) = \|v_1\| \cdot \|v_2\| \cdot \ldots \cdot \|v_k\| \, T\left(\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \ldots, \frac{v_k}{\|v_k\|}\right)$$

$$\leq \|v_1\| \cdot \|v_2\| \cdot \ldots \cdot \|v_k\| \sup\{|T(w_1, \ldots, w_k)| : \|w_1\| = \|w_2\| = \ldots = \|w_k\| = 1\} =$$

$$= \|v_1\| \cdot \ldots \cdot \|v_k\| \cdot \sup_{S^{n-1} \times S^{n-1} \times \ldots \times S^{n-1}} |T|$$

T jest ciągłe (jako przekształcenie z $\mathbb{R}^n \times \ldots \times \mathbb{R}^n$ w $\mathbb{R}^m$)

$A = S^{n-1} \times \ldots \times S^{n-1}$ jest zwartym podzbiorem $\mathbb{R}^n \times \ldots \times \mathbb{R}^n$, więc $\sup_A |T|$ jest skończone. Wielkość tę nazywamy normą operatorową przekształcenia T i oznaczamy $|||T|||$.

$$T(v_1, \ldots, v_k) \leq |||T||| \cdot \|v_1\| \cdot \|v_2\| \cdot \ldots \cdot \|v_k\|.$$

Wróćmy do $D^k f$.

Uwaga: Jeżeli istnieje, to $D^k f(x)$ jest przekształceniem k-liniowym, więc $D^k f(x) : \mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}^m$ jest ciągłe JAKO FUNKCJA ZMIENNYCH $v_1, \ldots, v_k \in \mathbb{R}^n \times \ldots \times \mathbb{R}^n$, nie jako funkcja zmiennej x!

(~~Zgoda, że jest~~

Pochodnych cząstkowych – nawet drugiego rzędu – mamy całe zoo. Trochę nam w opanowaniu sytuacji pomaga twierdzenie pochodzące od Hermanna Schwarza:

Twierdzenie (Schwarza o równości pochodnych mieszanych)
Załóżmy, że $\mathcal{U} \subset \mathbb{R}^n$ jest otwarty, zaś $f \in C^1(\mathcal{U}, \mathbb{R}^m)$. Załóżmy dodatkowo, że

1) dla pewnych $i, j \in \{1, 2, \dots, n\}$ we wszystkich punktach $x \in \mathcal{U}$ istnieje pochodna cząstkowa drugiego rzędu $\frac{\partial^2 f}{\partial x_i \partial x_j}$ i że jest ona w $\mathcal{U}$ ciągła.

lub

2) funkcja f jest dwukrotnie różniczkowalna w pewnym $a \in \mathcal{U}$.

<u>Teza:</u> 1) $\frac{\partial^2 f}{\partial x_j \partial x_i}$ istnieje we wszystkich punktach $\mathcal{U}$ i $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$

2) dla dowolnych $i, j \in \{1, 2, \dots, n\}$ $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$.

Nim przejdziemy do dowodu punktu 1) Twierdzenia Schwarza, przyjrzyjmy się przykładowi, pochodzącemu od Giuseppe Peano, pokazującemu, że założenie ciągłości drugiej pochodnej mieszanej jest konieczne:

<u>Przykład</u> (Peano)

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

Dla $(x,y)\neq(0,0)$ funkcja f, z twierdzeń o pochodnej iloczynu, złożenia itd., jest różniczkowalna, a jej pochodną możemy łatwo wyznaczyć. Co w $(x,y)=(0,0)$?

Wykażemy, że f jest różniczkowalna w $(0,0)$, a jej pochodną jest przekształcenie zerowe. Oszacujmy w tym celu

$$\frac{\| f(x,y) - f(0,0) - \overbrace{\langle (0,0),(x,y)\rangle}^{0\cdot(x,y)} \|}{\|(x,y)\|} = \left| \frac{xy\,(x^2-y^2)}{(x^2+y^2)^{3/2}} \right| \leq$$

$$\leq \frac{\frac{1}{2}(x^2+y^2)(x^2+y^2)}{(x^2+y^2)\sqrt{x^2+y^2}} = \frac{1}{2}\sqrt{x^2+y^2} \xrightarrow[(x,y)\to(0,0)]{} 0 .$$

bo $|xy| \leq \frac{1}{2}(x^2+y^2)$, $|x^2-y^2| \leq x^2+y^2$

Stąd, jak przypuszczaliśmy, f jest różniczkowalna w $(0,0)$, a jej pochodna jest przekszt. zerowe.

W szczególności $\frac{\partial f}{\partial x}(0,0)=\frac{\partial f}{\partial y}(0,0)=0$.

Chcemy teraz wyznaczyć $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ i $\frac{\partial^2 f}{\partial y \partial x}(0,0)$

Potrzebujemy do tego wartości $\frac{\partial f}{\partial y}(x,0)$ i $\frac{\partial f}{\partial x}(0,y)$ dla x,y bliskich zeru $(i \neq 0)$.

$$\frac{\partial f}{\partial y}(x,0)=\frac{\partial}{\partial y}\left(xy\frac{x^2-y^2}{x^2+y^2}\right)\Bigg|_{\cancel{(x,0)}\ y=0}=\overbrace{\frac{\partial}{\partial y}(xy)\Big|_{y=0}}^{=x}\cdot\underbrace{\frac{x^2-y^2}{x^2+y^2}\Big|_{y=0}}_{=1}+$$

$$+\underbrace{xy\Big|_{y=0}}_{=0}\cdot\frac{\partial}{\partial y}\left(\frac{x^2-y^2}{x^2+y^2}\right)\Bigg|_{y=0}=x$$

$$\frac{\partial f}{\partial x}(0,y)=\frac{\partial}{\partial x}\left(xy\frac{x^2-y^2}{x^2+y^2}\right)\Bigg|_{x=0}=\overbrace{\frac{\partial}{\partial x}(xy)\Big|_{x=0}}^{=y}\cdot\overbrace{\frac{x^2-y^2}{x^2+y^2}\Big|_{x=0}}^{=-1}$$

$$+\underbrace{xy\Big|_{x=0}}_{=0}\frac{\partial}{\partial x}\left(\frac{x^2-y^2}{x^2+y^2}\right)\Bigg|_{x=0}=-y.$$

Stąd

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \cancel{\lim} \frac{\partial}{\partial x}\Big|_{x=0} \left(\frac{\partial f}{\partial y}(x,0)\right) = \frac{\partial}{\partial x}\Big|_{x=0} x = 1$$

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \frac{\partial}{\partial y}\Big|_{y=0} \left(\frac{\partial f}{\partial x}(0,y)\right) = \frac{\partial}{\partial y}\Big|_{y=0} (-y) = -1.$$

Zatem $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

Proszę sprawdzić, że $\frac{\partial^2 f}{\partial x \partial y}$ nie jest ciągła w $(0,0)$.

(Wskazówka: wyznaczyć $\frac{\partial^2 f}{\partial x \partial y}(x,0)$ i $\frac{\partial^2 f}{\partial x \partial y}(0,y)$).

Dowód tw. Schwarza.

1) Dowód za skryptem p. Strzeleckiego.
Potrzebujemy dwóch lematów, z których pierwszy na pewno przyda nam się jeszcze w przyszłości.

<u>Lemat 1</u> (o różniczkowaniu pod znakiem całki)

Niech $Q = [a,b] \times [c,d] \subset \mathbb{R}$ i niech
$f : Q \to \mathbb{R}^m$ będzie funkcją ciągłą.

Załóżmy też, że f ma ciągłą pochodną cząstkową $\frac{\partial f}{\partial y} : Q \to \mathbb{R}^m$.

Wówczas funkcja $\Phi : [c,d] \to \mathbb{R}^m$ dana wzorem $\Phi(y) = \int_a^b f(x,y)\,dx$

jest różniczkowalna i $\Phi'(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx$.

Innymi słowy, zachodzi wówczas równość

$$\frac{d}{dy} \int_a^b f(x,y)\,dx = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx.$$

<u>Uwaga</u>: zakładamy, że $i \neq j$, bo dla $i = j$ teza twierdzenia jest oczywista.

109/

Dowód

Oszacujemy różnicę $\left\| \frac{\Phi(y+h)-\Phi(y)}{h} - \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx \right\| = \star$

i wykażemy, że $\forall_{\varepsilon>0} \exists_{\delta>0}$ taka, że jeżeli $|h|<\delta$, to $\star < \varepsilon$. To dowiedzie, że

$$\Phi'(y) = \lim_{h\to 0} \frac{\Phi(y+h)-\Phi(y)}{h} = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx.$$

Na początek zauważmy, że

$$\frac{\Phi(y+h)-\Phi(y)}{h} = \cancel{\int_a^b f(x,y)\,dx} \; \frac{\int_a^b f(x,y+h)\,dx - \int_a^b f(x,y)\,dx}{h} =$$

$$= \int_a^b \frac{1}{h}\left(f(x,y+h) - f(x)\right)dx.$$

Różnicę $f(x,y+h) - f(x)$ możemy natomiast przedstawić jako całkę z odpowiedniej pochodnej

$$f(x,y+h) - f(x,y) = \int_0^1 \frac{d}{ds}\left(f(x,y+sh)\right)ds =$$

$$= \int_0^1 \frac{\partial f}{\partial y}(x,y+sh)\cdot h\,ds = h\int_0^1 \frac{\partial f}{\partial y}(x,y+sh)\,ds$$

Łącząc te wyniki możemy zapisać

$$\star = \left\| \int_a^b \int_0^1 \frac{\partial f}{\partial y}(x,y+sh)\,ds\,dx - \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx \right\| =$$

$$= \left\| \int_a^b \left(\int_0^1 \frac{\partial f}{\partial y}(x,y+sh)\,ds - \frac{\partial f}{\partial y}(x,y)\right)dx \right\| =$$

110)

$$= \left\| \int_a^b \int_0^1 \left(\frac{\partial f}{\partial y}(x,y+sh) - \frac{\partial f}{\partial y}(x,y) \right) ds\, dx \right\|$$

$$\leq \int_a^b \int_0^1 \left\| \frac{\partial f}{\partial y}(x,y+sh) - \frac{\partial f}{\partial y}(x,y) \right\| = ☺$$

Funkcja $\frac{\partial f}{\partial y}$ jest na $[a,b]\times[c,d]$ ciągła, a ów prostokąt jest zwarty, jest więc na nim jednostajnie ciągła. ~~Mam~~ Dla każdego więc $\varepsilon>0$ umiemy dobrać $\delta>0$ tak, by jeżeli tylko $\underset{u,v\in Q}{\|u-v\|}<\delta$, to $\left\|\frac{\partial f}{\partial y}(u)-\frac{\partial f}{\partial y}(v)\right\| < \frac{\varepsilon}{b-a}$

W ten sposób, jeżeli $|h|<\delta$, to $\|(x,y+sh)-(x,y)\| \underset{\text{bo } s\in[0,1]}{\leq} |h| < \delta$, zatem

$$\left\| \frac{\partial f}{\partial y}(x,y+sh) - \frac{\partial f}{\partial y}(x,y) \right\| < \frac{\varepsilon}{b-a}, \text{ więc}$$

$$☺ < \int_a^b \int_0^1 \frac{\varepsilon}{b-a}\, ds\, dx = \int_a^b \frac{\varepsilon}{b-a}\, dx = \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

$\square$.

Lemat 2 (tw. Schwarza w dwóch wymiarach)

Niech $\mathcal{U} \subset \mathbb{R}^2$ będzie zbiorem otwartym i niech $Q = [a,b] \times [c,d] \subset \mathcal{U}$.

Załóżmy, że $f \in C^1(\mathcal{U}, \mathbb{R}^m)$ ma pochodną cząstkową $\frac{\partial^2 f}{\partial x \partial y}$ w Q i że jest ona na Q ciągła.

Wówczas w Q istnieje $\frac{\partial^2 f}{\partial y \partial x}$ i dla $(x,y) \in Q$ zachodzi równość

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y)$$

Dowód

Przedstawimy dowód dla $m=1$; dla $m>1$ wystarczy powtórzyć go oddzielnie dla każdej współrzędnej f_i funkcji $f = (f_1, \dots, f_m)$.

Niech $(x,y) \in Q$.

Mamy $f(x,t) = f(x,y) + \int_y^t \frac{\partial f}{\partial y}(x,z)\,dz$ dla t takich, że $(x,t) \in Q$.

Stąd $\cancel{\left(\frac{\partial f}{\partial x}(x,t) - \frac{\partial f}{\partial x}(x,y)\right) / (t-y)}$

Zróżniczkujmy obie strony równości ~~p~~ względem x. Na mocy pierwszego lematu mamy

$$\frac{\partial f}{\partial x}(x,t) = \frac{\partial f}{\partial x}(x,y) + \int_y^t \frac{\partial^2 f}{\partial x \partial y}(x,z)\,dz$$

112)

Stąd

$$\frac{\frac{\partial f}{\partial x}(x,t) - \frac{\partial f}{\partial x}(x,y)}{t-y} = \frac{1}{t-y}\int_y^t \frac{\partial^2 f}{\partial x \partial y}(x,z)\,dz$$

z tw. o wartości średniej (zerowego, całkowego) prawa strona jest równa $\frac{\partial^2 f}{\partial x\partial y}(x,\xi)$

dla pewnego $\xi\in[y,t]$

(tu korzystamy z ciągłości $\frac{\partial^2 f}{\partial x\partial y}$ na Q).

Jeżeli teraz $t\to y$, to

$$\begin{array}{ccc} \dfrac{\frac{\partial f}{\partial x}(x,t) - \frac{\partial f}{\partial x}(x,y)}{t-y} & = & \dfrac{\partial^2 f}{\partial x\partial y}(x,\xi) \\ \big\downarrow\, t\to y & & \big\downarrow\, t\to y \ \text{ więc i } \xi\to y \text{ bo } \xi\in[y,t] \\ \dfrac{\partial^2 f}{\partial y\partial x}(x,y) & & \dfrac{\partial^2 f}{\partial x\partial y}(x,y) \end{array}$$

to dowodzi tezy lematu. $\square$.

113)

Teraz dowód 1) z tw. Schwarza wynika już z prostej obserwacji, że wartość $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ i $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ zależy tylko od wartości f na płaszczyźnie afinicznej $x + \mathrm{span}(e_i, e_j)$.

Rozważając $\tilde{f}(t,s) = f(x + te_i + se_j)$ widzimy, że $\tilde{f}$ spełnia założenia Lematu 2, mamy

więc
$$\frac{\partial^2 \tilde{f}}{\partial t \partial s}(0,0) = \frac{\partial^2 \tilde{f}}{\partial s \partial t}(0,0)$$

ale
$$\begin{array}{ccc} \| & & \| \\ \frac{\partial^2 f}{\partial x_i \partial x_j}(x) & & \frac{\partial^2 f}{\partial x_j \partial x_i}(x). \end{array}$$

$\square$.

Dowód 2) z tw. Schwarza

Ustalmy $v, w \in \mathbb{R}^n$ i rozważmy pomocniczą funkcję

$$\varphi(s,t) = f(a + tw + sv) - f(a+tw) - f(a+sv) + f(a) - $$
$$- ts\, D^2 f(a)(w,v).$$

Oznaczmy $M = \max(\|v\|, \|w\|)$.

114)

Z definicji φ widzimy, że $\varphi(0,t)=0$, więc na mocy twierdzenia o wartości średniej

$$\|\varphi(s,t)\| = \|\varphi(s,t)-\varphi(0,t)\| \le |s| \sup_{\sigma\in[0,s]} \left\|\frac{\partial\varphi}{\partial s}(\sigma,t)\right\|$$

Obliczmy zatem $\frac{\partial\varphi}{\partial s}(\sigma,t)$:

$$\frac{\partial\varphi}{\partial s}(\sigma,t) = Df(a+tw+\sigma v)\cdot v - Df(a+\sigma v)v - tD^2f(a)(w,v) \underset{\uparrow}{=}$$

tu korzystamy z istnienia $D^2f(a)$

$$= Df(a)v + D^2f(a)(tw+\sigma v, v) + \|tw+\sigma v\| r(tw+\sigma v)v - Df(a)v - D^2f(a)(\sigma v, v) - \|\sigma v\| r(\sigma v)v - tD^2f(a)(w,v) = \|tw+\sigma v\| r(tw+\sigma v)v - \|\sigma v\| r(\sigma v) v$$

więc $(tw+\sigma v,\ \sigma v \in B(0, \max(|t|,|s|)M))$

$$\sup_{\sigma\in[0,s]} \left\|\frac{\partial\varphi}{\partial s}(\sigma,t)\right\| \le \max(|t|,|s|)M^2 \sup\{|||r(u)||| : u\in B(0,\max(|t|,|s|)M$$

Biorąc $s=t$ dostajemy

$$\|\varphi(t,t)\| \le |t|\cdot|t|\cdot M^2 \sup\{|||r(u)||| : u\in B(0,|t|M)\} = t^2M^2 \sup\{|||r(u)||| : u\in B(0,|t|M)\}$$

wiemy, że $r(u) \xrightarrow[u\to 0]{} 0$, więc $\sup\{|||r(u)||| : u\in B(0,|t|M)\} \xrightarrow[t\to 0]{} 0$.

Stąd $\frac{\|\varphi(t,t)\|}{t^2} \xrightarrow{t\to 0} 0$.

To oznacza, że

$$\frac{\varphi(t,t)}{t^2} = \frac{f(a+tw+sv)-f(a+tw)-f(a+sv)+f(a)}{t^2} - D^2f(a)(w,v)$$

$$\xrightarrow[t\to 0]{} 0, \text{ zatem}$$

$$D^2f(a)(w,v) = \lim_{t\to 0} \frac{f(a+tw+sv)-f(a+tw)-f(a+sv)+f(a)}{t^2}$$

Prawa strona tego wzoru jest symetryczna ze względu na rolę wektorów w i v – nie zmienia się, gdy je zamienimy. Dlatego i lewa musi być symetryczna: $D^2f(a)(v,w) = D^2f(a)(w,v)$.

Macierz przekształcenia $D^2f(a)$ jest macierzą Jacobiego macierzy $Df(a)$, skąd łatwo widać, że

$$\underset{\displaystyle \frac{\partial^2 f}{\partial x_i \partial x_j}(a)}{\underset{\shortparallel}{D^2f(a)(e_i,e_j)}} = \underset{\displaystyle \frac{\partial^2 f}{\partial x_j \partial x_i}(a)}{\underset{\shortparallel}{D^2f(a)(e_j,e_i)}}.$$

Wniosek z 1)

Załóżmy, że $f \in C^1(\mathcal{U}, \mathbb{R}^m)$ ma w $\mathcal{U}$ wszystkie drugie pochodne cząstkowe i że są one ciągłe na $\mathcal{U}$. Wówczas f jest dwukrotnie różniczkowalna na $\mathcal{U}$ i $\forall_{x \in \mathcal{U}}$ $D^2 f(x)$ ma macierz symetryczną ($\Leftrightarrow$ jest przekształceniem dwuliniowym symetrycznym)

Dowód: Dwukrotna różniczkowalność (tj. różniczkowalność Df) wynika z ciągłości $\frac{\partial}{\partial x_i} Df(x)$ (w bazach standardowych współrzędne $\frac{\partial}{\partial x_i} Df(x)$ to $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$) i z twierdzenia o tym, że ciągłość pochodnych cząstkowych w otoczeniu pewnego punktu pociąga za sobą różniczkowalność w tym punkcie.
Symetryczność macierzy $D^2 f(x)$ wynika wtedy z (któregokolwiek podpunktu) twierdzenia Schwarza.

Wzór Taylora

Wzór Taylora z resztą w postaci Peano

Twierdzenie: Niech $\mathcal{U}\subset\mathbb{R}^n$ będzie otwarty, $f:\mathcal{U}\to\mathbb{R}^m$. Załóżmy, że f jest $(k-1)$-krotnie różniczkowalna w $\mathcal{U}$, a dla pewnego $x\in\mathcal{U}$ istnieje $D^k f(x)$. Niech $r>0$ będzie takie, że $B(x,r)\subset\mathcal{U}$. Wówczas dla każdego $h \in \mathbb{R}^n$ tż. $\|h\|<r$

$$f(x+h) = f(x) + Df(x)h + \frac{1}{2!}D^2 f(x)(h,h) + \frac{1}{3!}D^3 f(x)(h,h,h) + \ldots + \frac{1}{k!}D^k f(x)(h,\ldots,h) + R(h)$$

to jest definicja $R(h)$

gdzie $$\frac{R(h)}{\|h\|^k} \longrightarrow 0 \quad \text{gdy } h\to 0$$

a to teza twierdzenia

Pierwsze podejście do dowodu. Dla uproszczenia weźmy $m=1$ i ustalmy $v \in \mathbb{R}^n \setminus \{0\}$

Rozważmy $g:(-\varepsilon,\varepsilon) \to \mathbb{R}$ (ε dostatecznie małe, by $[x-\varepsilon v, x+\varepsilon v] \subset \mathcal{U}$)
$$g(t) = f(x+tv).$$

Funkcja g jest w $t=0$ k-krotnie różniczkowalna, jako złożenie funkcji k-krotnie różniczkowalnych.

Mamy

$$g'(t) = Df(x+tv)v \Rightarrow g'(0) = Df(x)v$$
$$g''(t) = D^2f(x+tv)(v,v) \Rightarrow g''(0) = D^2f(x)(v,v)$$
$$\vdots$$
$$g^{(k-1)}(t) = D^{k-1}f(x+tv)(v,\dots,v) \Rightarrow g^{(k-1)}(0) = D^{k-1}f(x)(v,\dots,v)$$
$$g^{(k)}(0) = D^k f(x)(v,\dots,v).$$

Stąd
$$g(t) = g(0) + g'(0)t + \frac{1}{2!}g''(0)t^2 + \dots + \frac{1}{k!}g^{(k)}(0)t^k + r(t),$$
gdzie $r(t)/t^k \xrightarrow[t\to 0]{} 0$.

czyli
$$f(x+tv) = f(x) + t\cdot Df(x)v + \frac{t^2}{2!}D^2f(x)(v,v) + \dots + \frac{t^k}{k!}D^kf(x)(v,\dots,v) + r(t)$$

Funkcja g, więc również i r, zależą od v, więc zamiast $r(t)$ będę pisał $r(v,t)$.

Oznaczając $h=tv$ mamy
$$f(x+h) = f(x) + Df(x)h + \frac{1}{2!}D^2f(x)(h,h) + \dots + \frac{1}{k!}D^kf(x)(h,\dots,h) + r(v,t).$$

119)

Jeżeli teraz $\overset{\neq}{v}$ jest ustalone i $h = tv \to 0$, to równoważnie $t \to 0$

Stąd $$\lim_{\substack{h\to 0 \\ h = t\cdot v}} \frac{R(h)}{\|h\|^k} = \lim_{t\to 0} \frac{r(v,t)}{t^k} \cdot \frac{1}{\|v\|^k} = 0.$$

Widzimy więc, że gdy z h dążymy do 0 po „promieniach", tj po półprostej wyznaczonej przez (dowolny, ale ustalony) wektor $v \neq 0$, to $\frac{R(h)}{\|h\|^k} \to 0$. To jednak, jak już wiemy, nie gwarantuje tego, że $\lim_{h\to 0} \frac{R(h)}{\|h\|^k} = 0$.

Z tego podejścia możemy jednak, przy odrobinie silniejszych założeniach (pozwalających na użycie reszty w post. Lagrange'a) uzyskać pożyteczny

Wniosek: Jeżeli $f \in C^2(\mathcal{U}, \mathbb{R})$, $a \in \mathcal{U}$ i h jest takie, że $[a, a+h] \subset \mathcal{U}$, ~~Wówczas~~ to istnieje $\theta \in [0,1]$ takie, że

$$f(a+h) = f(a) + Df(a)h + \frac{1}{2} D^2 f(a+\theta h)(h,h)$$

Dowód: Wypisać wzór Taylora (dla g) z $k=1$ i resztą w postaci Lagrange'a.

120)

Porządny dowód

Na początek przypomnienie: jeżeli $T: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^\ell$ jest dwuliniowe, $\mathcal{U} \subset \mathbb{R}^n$ otwarty, $f, g : \mathcal{U} \to \mathbb{R}^m$ różniczkowalne w $x \in \mathcal{U}$, to $y \mapsto T(f(y), g(y))$ jest różniczkowalna w $y = x$ i $\forall_{v \in \mathbb{R}^n}$

$$D(T(f,g))(x)v = T(Df(x)v, g(x)) + T(f(x), Dg(x)v).$$

Dokładnie tak samo dowodzi się analogiczne twierdzenie dla przekształceń k-liniowych: jeżeli $T: \underbrace{\mathbb{R}^m \times \mathbb{R}^m \times \ldots \times \mathbb{R}^m}_{k\text{ - krotnie}} \to \mathbb{R}^\ell$ jest k-liniowe, $\mathcal{U} \subset \mathbb{R}^n$ otwarty, $f_1, \ldots, f_k : \mathcal{U} \to \mathbb{R}^m$ różniczkowalne w $x \in \mathcal{U}$, to

$$D(T(f_1, f_2, \ldots, f_k))(x)v =$$

$$T(Df_1(x)v, f_2(x), \ldots, f_k(x)) + T(f_1(x), Df_2(x)v, f_3(x), \ldots, f_k(x)) + \\ + \ldots + T(f_1(x), f_2(x), \ldots, Df_k(x)v).$$

Niech teraz, zgodnie z treścią twierdzenia,

$$R(h) = f(x+h) - \left(f(x) + Df(x)h + \frac{1}{2!}D^2f(x)(h,h) + \ldots + \frac{1}{k!}D^kf(x)(h,\ldots,h)\right).$$

Oczywiście $R(0)=0$.

Pamiętając, że x jest ustalonym punktem $\mathcal{U}$ obliczmy kolejne pochodne $R(h)$ względem h

Dla $v \in \mathbb{R}^n$

$$DR(h)v = Df(x+h)\cdot id \cdot v - \Big(Df(x)\cdot v + \frac{1}{2!}\big(D^2f(x)(id\cdot v, h) + D^2f(x)(h, id\cdot v)\big) + \frac{1}{3!}\big(D^3f(x)(id\cdot v, h, h) + D^3f(x)(h, id\cdot v, h) + D^3f(x)(h,h,id\cdot v)\big) + \ldots + \frac{1}{k!}\big(D^kf(x)(id\cdot v, h, \ldots, h) + D^kf(x)(h, id\cdot v, h,\ldots,h) + \ldots + D^kf(x)(h,h,\ldots,id\cdot v)\big)\Big) = \star.$$

Wiemy już, że $D^2f(x)$ jest przekształceniem dwuliniowym symetrycznym, a więc $D^2f(x)(v,h) = D^2f(x)(h,v)$.

Również wyższe pochodne są przekształceniami symetrycznymi (zagadka: dlaczego?!), więc $D^jf(x)(v,h,\ldots,h) = D^jf(x)(h,v,h,\ldots,h) = \ldots = D^jf(x)(h,h,\ldots,h,v)$.

Stąd $\star$ możemy przekształcić do

$$DR(h)v = \star = Df(x+h)v - Df(x)v - D^2f(x)(h,v) - \frac{1}{2!}D^3f(x)(h,h,v) - \ldots - \frac{1}{(k-1)!}D^kf(x)(h,h,\ldots,h,v).$$

W szczególności

$DR(0) = 0$.

12-2)

Różniczkując w ten sam sposób dalej dostajemy $\forall v_1, v_2 \in \mathbb{R}^n$

$$D^2R(h)(v_1,v_2) = D^2f(x+h)(v_1,v_2) - D^2f(x)(v_1,v_2) -$$

$$- \frac{1}{1!}D^3f(x)(h,v_1,v_2) - \frac{1}{2!}D^4f(x)(h,h,v_1,v_2) - \ldots -$$

$$- \frac{1}{(k-2)!}D^kf(x)(h,\ldots,h,v_1,v_2)$$

więc $D^2R(0)=0$

i dalej, krok po kroku dowodzimy, że $D^jR(0)=0$ dla $j=1,2,\ldots,k-1$:

$$D^{k-1}R(h)(v_1,v_2,\ldots,v_{k-1}) = D^{k-1}f(x+h)(v_1,\ldots,v_{k-1}) -$$

$$- D^{k-1}f(x)(v_1,\ldots,v_{k-1}) - D^kf(x)(h,v_1,\ldots,v_{k-1})$$

więc $D^{k-1}R(0)=0$.

Dalej nie możemy już różniczkować, bo nie wiemy, czy $D^kR(x+h)$ istnieje dla $h\neq 0$. Przejdźmy więc do szacowania $\|R(h)\|$:

$$\|R(h)\| = \|R(h)-R(0)\| \underset{\substack{\uparrow \\ \text{tw. o wartości średniej}}}{\leq} \|h\| \sup_{t\in[0,1]} \|\!|\!| DR(th) |\!|\!\|$$

$\|\!|\!| DR(th) |\!|\!\|$ jest ciągłą funkcją zmiennej t, więc $\exists_{\theta\in[0,1]}$ tż.

$$\sup_{t\in[0,1]} \|\!|\!| DR(th)|\!|\!\| = \|\!|\!| DR(\theta h)|\!|\!\| = \|\!|\!| DR(\theta h) - DR(0)|\!|\!\| \underset{\substack{\uparrow \\ \text{znów tw. o wart. średniej}}}{\leq}$$

$$\leq \|\theta h\| \cdot \sup_{t\in[0,\theta]} \|\!|\!| D^2R(th)|\!|\!\|$$

$$\leq \|h\| \sup_{t\in[0,1]} \|\!|\!| D^2R(th)|\!|\!\|, \text{ więc } \|R(h)\| \leq \|h\|^2 \sup_{t\in[0,1]} \|\!|\!| D^2R(th)|\!|\!\|$$

i dalej, tak samo,

$$\|R(h)\| \leq \|h\|^{k-1} \sup_{t\in[0,1]} |||D^{k-1}R(th)|||$$

Wiemy, że

$$D^{k-1}R(th) = D^{k-1}f(x+th) - D^{k-1}f(x) - D^k f(x)th,$$

$\{$ to jest równość przekształceń $(k-1)$-liniowych,
$D^k f(x)th(v_1,\dots,v_{k-1}) := D^k f(x)(th, v_1, \dots, v_{k-1})$ $\}$

a funkcja $D^{k-1}f$ jest różniczkowalna w x, zatem

$$|||D^{k-1}R(th)||| = |||D^{k-1}f(x+th) - D^{k-1}f(x) - D^k f(x)\cdot th||| \leq$$

$= \|th\|\, |||r(th)|||$, gdzie $r(th) = D^{k-1}f(x+th) - D^{k-1}f(x) - D^k f(x)th$ ma tę własność, że

$$r(v) \xrightarrow{v\to 0} 0$$

stąd $\sup_{t\in[0,1]} |||D^{k-1}R(th)||| \leq \|h\| \cdot \sup_{v\in B(0,\|h\|)} |||r(v)|||$

ostatecznie

$$\|R(h)\| \leq \|h\|^k \sup_{v\in B(0,\|h\|)} |||r(v)|||$$

i $$\frac{\|R(h)\|}{\|h\|^k} = \sup_{v\in B(0,\|h\|)} |||r(v)||| \xrightarrow[h\to 0]{} 0.$$

$\square$