Pochodne wyższego rzędu

Załóżmy, że funkcja $f: \mathcal{U} \to Y$, gdzie $\mathcal{U} \subset X$, jest różniczkowalna na $\mathcal{U}$. Wówczas pochodna Df przypisuje każdemu punktowi $x \in \mathcal{U}$ przekształcenie liniowe $Df(x) \in B(X,Y)$, więc

$$Df : \mathcal{U} \longrightarrow B(X,Y).$$

Przestrzeń $B(X,Y)$ przekształceń liniowych ciągłych jest przestrzenią Banacha, możemy więc znów pytać o to, czy Df jest funkcją ciągłą na $\mathcal{U}$ (mówimy wtedy, że $f \in C^1(\mathcal{U}, Y)$), czy jest na $\mathcal{U}$ różniczkowalna?

Z definicji Df jest różniczkowalna w $x \in \mathcal{U}$, jeżeli istnieje $T \in B(X, B(X,Y))$ takie, że

$$\forall_{y \in \mathcal{U}} \quad Df(y) = Df(x) + T(y-x) + R(y-x),$$

gdzie $\dfrac{R(y-x)}{\|y-x\|_X} \xrightarrow[y \to x]{} 0.$

Czym w ogóle jest T? Przekształcenie liniowe (i ciągłe) o wartościach w przestrzeni przekształceń liniowych ciągłych...

Każdemu $v \in X$ przypisujemy przekształcenie liniowe ciągłe $T(v)(\cdot) : X \to Y$ — i przypisanie to jest liniowe i ciągłe. Mamy więc

$T : X \times X \to Y$, które jest liniowe i ciągłe względem każdej ze zmiennych.

Druga pochodna w ustalonym punkcie $x \in U$ jest zatem przekształceniem dwuliniowym, ciągłym względem obu argumentów.

Co oznacza jego ciągłość? Dla przekształcenia liniowego ciągłość = ograniczoność:

$T(\cdot,\cdot)$ jest ciągłe względem pierwszej zmiennej v, $T(\cdot,\cdot) : X \to B(X,Y)$, gdy istnieje $C>0$ t.ż.:

$$\forall_{v \in X} \quad \underbrace{|||T(v,\cdot)|||}_{\text{norma w } B(X,Y)} \le C\|v\|_X$$

Dalej, dla każdego $v \in X$ $T(v,\cdot)$ ma być liniowe i ciągłe, zatem $\forall_{w \in X} \ \|T(v,w)\|_Y \le |||T(v,\cdot)||| \, \|w\|_X$

Zbierając jedną i drugą nierówność, dostajemy, że T jest dwuliniowe ciągłe, gdy istnieje $C>0$ taka, że $\forall_{v,w \in X}$ $\|T(v,w)\|_Y \le C\|v\|_X \|w\|_X$.

Analogicznie, trzecia pochodna f w punkcie $x \in \mathcal{U}$, o ile istnieje, jest przekształceniem liniowym ciągłym

$$D^3 f(x) : X \to B(X, B(X, Y))$$

i łatwo można sprawdzić, że możemy $D^3 f(x)$ utożsamić z przekształceniem trzy-liniowym i $\forall_{u,v,w \in X} \; \|D^3 f(x)(u,v,w)\|_Y \leq C \|u\|_X \|v\|_X \|w\|_X$

i tak dalej, k-ta pochodna f w $x \in \mathcal{U}$, zdefiniowana jako pochodna $D^{k-1} f : \mathcal{U} \to \left\{\begin{array}{l}\text{przekszt.}\\ \text{(k-1)-liniowe}\\ \text{ciągłe}\\ \text{z } X \text{ w } Y\end{array}\right.$ jest przekształceniem k-liniowym ciągłym $D^k f(x) : \underbrace{X \times X \times X \times \ldots \times X}_{k \text{ razy}} \longrightarrow Y$

i istnieje $C > 0$ t.ż

$$\forall_{v_1, \ldots, v_k \in X} \quad \|D^k f(x)(v_1, \ldots, v_k)\|_Y \leq C \|v_1\|_X \cdot \ldots \cdot \|v_k\|_X$$

Tu trzeba jeszcze odpowiedzieć, że przestrzeń przekształceń k-liniowych ciągłych z X w Y jest przestrzenią Banacha z normą

$$|||T||| = \sup_{\substack{\|v_1\| \leq 1 \\ \vdots \\ \|v_k\| \leq 1}} \|T(v_1, \ldots, v_k)\|_Y$$

Dlaczego nie musimy już tego dowodzić?

<u>Uwaga:</u> Pochodna $D^k f$ jest funkcją $(k+1)$ zmiennych:

$x \in U \subset X$ oraz $v_1, \ldots, v_k \in X$

punkt   wektory

Żądamy, by $D^k f$ była liniowa i ciągła względem zmiennych $v_1, \ldots, v_k$. Względem x nie musi być ciągła (ani tym bardziej liniowa) – chyba, że chcemy liczyć kolejne pochodne ...

Jak to wygląda w skończenie wielu wymiarach? $f: U \to \mathbb{R}^k$, $U \subset \mathbb{R}^m$ otwarty, $x \in U$.

$D^2 f(x): \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^k$ jest dwuliniowe, więc istnieje jego macierz (a_{ij}) w bazie standardowej: $\mathbb{R}^k \ni a_{ij} = D^2 f(x)(e_i, e_j)$

Wiemy też, że $D^2 f(x)(e_i, \cdot) = D_{e_i} Df(x)(\cdot) =$

$$= \frac{\partial}{\partial x_i} Df(x) = \quad \text{, gdzie}$$

$$Df: U \to B(\mathbb{R}^m, \mathbb{R}^k) \cong M_{m \times k}$$

więc Df możemy utożsamić z jej macierzą (czyli macierzą Jacobiego f).

$$= \frac{\partial}{\partial x_i} \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial f_k}{\partial x_m} \end{pmatrix}.$$

Wiemy już, że w $\mathbb{R}^l$ zbieżność (a więc wszystkie przejścia graniczne, w tym różniczkowanie) jest po współrzędnych, zatem

$D^2 f(x)(e_i, \cdot)$, jako przekształcenie liniowe,

ma macierz $\begin{pmatrix} \frac{\partial}{\partial x_i}\frac{\partial f_1}{\partial x_1} & \frac{\partial}{\partial x_i}\frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial}{\partial x_i}\frac{\partial f_1}{\partial x_m} \\ \vdots & & \ddots & \vdots \\ \frac{\partial}{\partial x_i}\frac{\partial f_k}{\partial x_1} & \frac{\partial}{\partial x_i}\frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial}{\partial x_i}\frac{\partial f_k}{\partial x_m} \end{pmatrix}$

zatem $D^2 f(x)(e_i, e_j) = \left(\frac{\partial}{\partial x_i}\frac{\partial f_1}{\partial x_j}, \frac{\partial}{\partial x_i}\frac{\partial f_2}{\partial x_j}, \ldots, \frac{\partial}{\partial x_i}\frac{\partial f_k}{\partial x_j}\right)$.

Będziemy, dla uproszczenia, pisać

$\frac{\partial^2 f_1}{\partial x_i \partial x_j}$ zamiast $\frac{\partial}{\partial x_i}\frac{\partial f_1}{\partial x_j}$ itd.

Gdy $f : U \to \mathbb{R}$ (czyli $k=1$), jest troszkę prościej: a_{ij} to wtedy liczby, nie wektory i $D^2 f(x)(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}$

Macierz drugiej pochodnej, $D^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1}^{m}$ nazywamy <u>macierzą Hessa</u>, a jej wyznacznik – hessjanem funkcji f w punkcie x; pochodne $\frac{\partial^2 f}{\partial x_i \partial x_j}$ (i ogólniej $\frac{\partial^l f_i}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_l}}$) nazywamy drugimi (odp. l-tymi) pochodnymi cząstkowymi funkcji f.

Jak widać, już dla $f: \mathbb{R}^m \to \mathbb{R}^k$ w punkcie $x \in \mathbb{R}^m$ mamy m^2 różnych drugich pochodnych cząstkowych, z których każda ma k współrzędnych.

To zoo pomaga nam nieco opanować twierdzenie Schwarza:

<u>Twierdzenie</u> (Hermanna Schwarza o równości pochodnych mieszanych), dwa warianty

Niech $U \subset \mathbb{R}^n$ będzie otwarty, $f \in C^1(U, \mathbb{R}^m)$.

Załóżmy dodatkowo, że

1) dla pewnych $i, j \in \{1, \dots, n\}$ we wszystkich $x \in U$ istnieje (druga) pochodna cząstkowa $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ i jest ona ciągłą funkcją $x \in U$.

lub

2) f jest dwukrotnie różniczkowalna w $a \in U$.

Wtedy

1) $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ istnieje we wszystkich punktach $x \in U$ i zachodzi równość

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

lub

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

Nim przejdziemy do dowodu, przyjrzyjmy się klasycznemu przykładowi, pochodzącemu od Giuseppe Peano, który dowodzi konieczności założeń w tw. Schwarza:

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

Oczywiście poza $(0,0)$ f jest ciągła i różniczkowalna. A w $(0,0)$? Wykażemy, że owszem, jest różniczkowalna, a jej pochodną jest przekształcenie zerowe. Mamy bowiem

przekształcenie zerowe z $\mathbb{R}^2$ w $\mathbb{R}$

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

$$\leq \frac{\frac{1}{2}(x^2+y^2)(x^2+y^2)}{(x^2+y^2)^{3/2}} = \frac{1}{2}\sqrt{x^2+y^2} \xrightarrow[(x,y)\to(0,0)]{} 0.$$

Wiemy w szczególności stąd, że $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$.

Aby zbadać istnienie i wyznaczyć wartości $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ i $\frac{\partial^2 f}{\partial y \partial x}(0,0)$, musimy znać wartości $\frac{\partial f}{\partial x}(0,y)$ i $\frac{\partial f}{\partial y}(x,0)$ dla (x,y) bliskich punktu $(0,0)$.

Mamy dla $x \neq 0$ $\frac{\partial f}{\partial y}(x,0) = \frac{\partial}{\partial y}\left(xy\frac{x^2-y^2}{x^2+y^2}\right)\Big|_{y=0} =$

$$= \underbrace{\frac{\partial}{\partial y}(xy)\Big|_{y=0}}_{= x} \cdot \underbrace{\frac{x^2-y^2}{x^2+y^2}\Big|_{y=0}}_{= 1} + \underbrace{xy\Big|_{y=0}}_{= 0} \cdot \underbrace{\frac{\partial}{\partial y}\frac{x^2-y^2}{x^2+y^2}\Big|_{y=0}}_{\text{coś}}$$

$= x$

i analogicznie $\frac{\partial f}{\partial x}(0,y) = -y$

(funkcja $f(x,y)$ przy zamianie $x \leftrightarrow y$ zmienia znak)

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

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

pochodne mieszane $\frac{\partial^2 f}{\partial x \partial y}$ i $\frac{\partial^2 f}{\partial y \partial x}$ nie są więc w $(0,0)$ równe.

<u>Zadanie:</u> proszę sprawdzić bezpośrednio, że funkcja f nie spełnia założeń 1) ani 2) twierdzenia Schwarza.

Dowód tw. Schwarza

Krok 1 Zauważmy, że gdy $m>1$, to $f=(f_1,\dots,f_m)$, gdzie $f_i : U \to \mathbb{R}$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \left(\frac{\partial^2 f_1}{\partial x_i \partial x_j}, \dots, \frac{\partial^2 f_m}{\partial x_i \partial x_j}\right),$$

więc wystarczy udowodnić tw. Schwarza dla funkcji $f: U \to \mathbb{R}$, a potem zastosować je oddzielnie do każdej ze zmiennych.

Krok 2 Tw. Schwarza dla $n=2$, przypadek 1)

Niech $U \subset \mathbb{R}^2$ otwarty, $Q=[a,b]\times[c,d] \subset U$.

Załóżmy, że $f \in C^1(U,\mathbb{R})$ ma pochodną cząstkową $\frac{\partial^2 f}{\partial x \partial y}$, która jest ciągła w Q. Wówczas w Q istnieje $\frac{\partial^2 f}{\partial y \partial x}$ i dla $(x,y)\in Q$ $\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$.

Dowód: Ustalmy $(x,y) \in Q$.

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

Różniczkujemy obie strony względem x:

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

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

Zadanko: proszę sprawdzić, że spełnione są założenia twierdzenia o różniczkowaniu pod znakiem całki.

Stąd

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

$$= \frac{\partial^2 f}{\partial x \partial y}(x,\xi) \quad \text{dla pewnego } \xi \in [y,t].$$

zerowe tw. o wart. średniej, $\frac{\partial^2 f}{\partial x \partial y}$ jest ciągła na Q.

Biorąc $t \to y$ dostajemy $\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y)$. □.

Krok 3: Twierdzenie Schwarza dla $n > 2$, przypadek 1)

Zauważamy, że wartości $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ i $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ zależą tylko od wartości f na płaszczyźnie afinicznej $x + \mathrm{span}(e_i, e_j)$; przyjmujemy

$\tilde{f}(t,s) = f(x + te_i + se_j)$ i stosujemy krok 2, dowodząc, że

$$\begin{array}{ccc} \frac{\partial^2 \tilde{f}}{\partial t \partial s}(0,0) & = & \frac{\partial^2 \tilde{f}}{\partial s \partial t}(0,0) \\ \| & & \| \\ \frac{\partial^2 f}{\partial x_i \partial x_j}(x) & & \frac{\partial^2 f}{\partial x_j \partial x_i}(0,0). \end{array}$$

Krok 1.
Dowód części 2)

Ustalmy $v, w \in \mathbb{R}^n$ i niech $\varphi(s,t) = f(a+tw+sv) -$
$- f(a+tw) - f(a+sv) + f(a) - ts\, D^2 f(a)(w,v)$

Naszym celem jest wykazanie, że $\frac{\varphi(t,t)}{t^2} \to 0$.
Dlaczego to wystarcza?

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

$\downarrow$
0

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

Prawa strona jest symetryczna ze względu na zamianę v i w, więc i lewa też musi być symetryczna.
To znaczy, że macierz $D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1}^n$ jest macierzą symetryczną - a to jest teza punktu 2).

No to oszacujmy najpierw $\|\varphi(s,t)\|$. Wiemy, że $\varphi(0,0) = 0$, więc $\|\varphi(s,t)\| = \|\varphi(s,t) - \varphi(0,0)\| \leq$

tw. o wart. średniej $\longrightarrow$ $\leq |s| \cdot \sup_{\sigma \in [0,s]} \left\| \frac{\partial \varphi}{\partial s}(\sigma,t) \right\|$ (☆)

No to obliczamy $\frac{\partial \varphi}{\partial s}(\sigma,t)$:

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

$$= Df(a)v + D^2 f(a)(\sigma v + tw, v) + R(\sigma v + tw) \cdot v$$
$$- Df(a)v - D^2 f(a)(\sigma v, v) - R(\sigma v) \cdot v - t\, D^2 f(a)(w,v) =$$

$$= t D^2 f(a)(w,v) + R(\sigma v + tw)\cdot v - R(\sigma v)\cdot v - t D^2 f(a)(w,v) =$$

$$= R(\sigma v + tw)\cdot v - R(\sigma v)\cdot v$$

Stąd $$\frac{\|\varphi(t,t)\|}{t^2} \underset{\text{z } (*),\ \text{przy } s=t}{\le} \frac{\sup\limits_{\sigma\in[0,t]} \left\| \frac{\partial \varphi}{\partial s}(\sigma,t)\right\|}{t} \le$$

$$\le \sup_{\sigma\in[0,t]} \left(\underset{\substack{\downarrow\, t\to 0 \\ 0}}{\frac{|||R(\sigma v + tw)|||}{\|\sigma v + tw\|}} \cdot \underbrace{\frac{\|\sigma v + tw\|}{t}}_{\le \|v\| + \|w\|} + \underset{\substack{\downarrow\, t\to 0 \\ 0}}{\frac{|||R(\sigma v)|||}{\|\sigma v\|}} \cdot \underbrace{\frac{\|\sigma v\|}{t}}_{\le \|v\|} \right)$$

$\xrightarrow[t\to 0]{} 0$, co dowodzi, że $\dfrac{\|\varphi(t,t)\|}{t^2} \xrightarrow[t\to 0]{} 0$. $\square$.

<u>Wniosek</u>: Jeżeli $f \in C^1(\mathcal{U}, \mathbb{R}^m)$ ma w $\mathcal{U}$ wszystkie drugie pochodne cząstkowe i są one ciągłe w $\mathcal{U}$, to f jest dwukrotnie różniczkowalna w $\mathcal{U}$ i $D^2 f(x)$ jest symetryczna dla wszystkich $x \in \mathcal{U}$ (czyli pochodne mieszane są równe).

<u>Dowód</u>: Df ma w $\mathcal{U}$ ciągłe pochodne cząstkowe

$\Rightarrow Df$ jest różniczkowalna w $x \in \mathcal{U}$ $\xRightarrow[\text{2), tw. Schwarza}]{}$ f ma w $x \in \mathcal{U}$ pochodne mieszane równe.

$\square$.

Twierdzenie (wzór Taylora z resztą w postaci Peano)

Niech X, Y będą Banacha, $\mathcal{U} \subset X$ otwarty, $f : \mathcal{U} \to Y$. Załóżmy dalej, że f jest $(k-1)$ razy różniczkowalna w $\mathcal{U}$ i dla pewnego $x \in \mathcal{U}$ istnieje $D^k f(x)$.
Wówczas dla każdego $y \in \mathcal{U}$

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

gdzie $\underbrace{R_k(y-x) / \|y-x\|^k \xrightarrow[y \to x]{} 0}_{\text{a to jest teza twierdzenia.}}$.

ta część to definicja R_k

Nim udowodnimy to ważne twierdzenie, przyjrzyjmy się jego nieco słabszym wersjom

- niech $Y = \mathbb{R}$, ustalmy $v \in X$, $v \neq 0$.

Rozważmy $g(t) = f(x + tv)$. Funkcja g, określona na $(-\varepsilon, \varepsilon)$ i o wartościach w $\mathbb{R}$, jest k-krotnie różniczkowalna (jako złożenie funkcji k-krotnie różniczkowalnych) w $t = 0$. Możemy zatem wypisać

dla ε t że: $B(x, \varepsilon) \subset \mathcal{U}$

dla g wzór Taylora z resztą Peano:

$$g(t) = g(0) + g'(0)t + \frac{1}{2!}g''(0)t^2 + \ldots + \frac{1}{k!}g^{(k)}(0)t^k + r_k(t)$$

Obliczmy kolejne pochodne g:

$$g'(t) = Df(x+tv)v \quad \Rightarrow \quad g'(0) = Df(x)v$$

$$g''(t) = D^2f(x+tv)(v,v) \quad \Rightarrow \quad g''(0) = D^2f(x)(v,v)$$

$$\vdots$$

$$g^{(k-1)}(t) = D^{k-1}f(x+tv)(\underbrace{v,\dots,v}_{k-1 \text{ razy}}) \Rightarrow g^{(k-1)}(0) = D^{k-1}f(x)(v,\dots,v)$$

i $g^{(k)}(0) = D^kf(x)(\underbrace{v,\dots,v}_{k \text{ razy}})$.

skąd

$$f(x+tv) = g(t) = g(0) + g'(0)t + g''(0)\cdot\frac{t^2}{2!} + \dots + g^{(k)}(0)\cdot\frac{t^k}{k!} + r_k(t)$$

$(\star)$

$$= f(x) + tDf(x)v + \frac{t^2}{2!}D^2f(x)(v,v) + \dots + \frac{t^k}{k!}D^kf(x)(v,\dots,v) + r_k(t),$$

gdzie $\frac{r_k(t)}{t^k} \xrightarrow[t\to 0]{} 0$. (oczywiście r zależy nie tylko od t, ale również od v, a także od x, będę więc dalej pisać $r_k(v,t)$.

Porównując $(\star)$ z wzorem Taylora widzimy, że $r_k(v,t) = R_k(v\cdot t)$, więc to, że $\frac{r_k(t)}{t^k} \to 0$ oznacza, że $\frac{R_k(v\cdot t)}{\|vt\|^k} \to 0$ (o ile $v \neq 0$),

czyli $\frac{R_k(z)}{\|z\|^k} \xrightarrow{z\to 0} 0$ o ile z zbiega do zera po promieniach! Wiemy już, że to mniej, niż $\lim_{z\to 0} \frac{R_k(z)}{\|z\|^k} = 0$, więc tą drogą nie udowodnimy wzoru Taylora.

Załóżmy jednak nieco więcej, że $f \in C^2(\mathcal{U}, \mathbb{R})$, i wypiszmy dla g wzór Taylora rzędu 1 z resztą w postaci Lagrange'a (możemy, bo $g \in C^2$):

$$g(t) = g(0) + g'(0)t + \frac{t^2}{2} g''(\tau) \quad \text{dla pewnego } \tau \in (0,t)$$

czyli

$$f(x+tv) = f(x) + t\,Df(x)v + \frac{t^2}{2} D^2 f(x+\tau v)(v,v).$$

Oznaczając $h = tv$, $\vartheta = \frac{\tau}{t}$, otrzymujemy

<u>Wniosek</u>: Jeżeli $f \in C^2(\mathcal{U}, \mathbb{R})$, $x \in \mathcal{U}$ i $h \in X$ jest takie, że $[x, x+h] \subset \mathcal{U}$, to istnieje $\vartheta \in [0,1]$ taka, że

$$f(x+h) = f(x) + Df(x)h + \frac{1}{2} D^2 f(x+\vartheta h)(h,h).$$

<u>Dowód wzoru Taylora</u>

<u>Lemat</u>: Jeżeli $T: \underbrace{Y \times Y \times \ldots \times Y}_{k \text{ razy}} \to Z$ jest przekształceniem k-liniowym ciągłym, $f_i : \mathcal{U} \to Y$ dla $i = 1, 2, \ldots, k$ są różniczkowalne w $x \in \mathcal{U}$, to $T(f_1, f_2, \ldots, f_k) : \mathcal{U} \to Z$ jest różniczkowalne w x i zachodzi

$$D(T(f_1, \ldots, f_k))(x)v = T(Df_1(x)v, f_2(x), \ldots, f_k(x)) + T(f_1(x), Df_2(x)v, \ldots, f_k(x)) + \ldots + T(f_1(x), \ldots, f_{k-1}(x), Df_k(x)v).$$

Dowód:

- albo powtarzamy dowód dla przekształceń dwuliniowych ciągłych,
- albo zauważamy, że np. przekszt. trzyliniowe ciągłe $T: Y\times Y\times Y \to Z$ możemy utożsamić z przekszt. dwuliniowym ciągłym o wartościach w przekształceniach liniowych: $T: Y\times Y \to B(Y,Z)$, a k-liniowe – przekszt. $(k-1)$-liniowym ciągłym o wartościach w $B(Y,Z)$ – i użyć indukcji.

No to niech $R(h) = f(x+h) - \big(f(x) + Df(x)h + \frac{1}{2!}D^2f(x)(h,h) +$

$$+ \ldots + \frac{1}{k!}D^kf(x)(h,\ldots,h)\big).$$

R jest określone dla wszystkich $h\in X$ takich, że $y = x+h \in \mathcal{U}$, w szczególności istnieje $\varepsilon > 0$ tż. R jest dobrze określona w $B(0,\varepsilon)\subset X$. Obliczmy pochodne R względem h:

dla ustalonego $v\in X$

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

Wiemy już, że druga pochodna przekształcenia różniczkowalnego f jest przekształceniem 2-liniowym symetrycznym. Również wyższe pochodne są symetryczne, bo są drugimi pochodnymi (pewnych pochodnych f).
Stąd np. $D^2f(x)(v,h) = D^2f(x)(h,v)$ i mamy

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

w szczególności $DR(0) = 0$

W ten sam sposób, różniczkując dalej, dostajemy

$$D^2R(h)(v_1,v_2) = D^2f(x+h)(v_1,v_2) - D^2f(x)(v_1,v_2) - \frac{1}{1!}D^3f(x)(h,v_1,v_2) + \frac{1}{2!}D^4f(x)(h,h,v_1,v_2) + \dots$$
$$\dots + \frac{1}{(k-2)!}D^kf(x)(h,\dots,h,v_1,v_2),$$ więc $D^2R(0)=0$

i tak dalej, aż do

$$D^{k-1}R(h)(v_1,\dots,v_{k-1}) = D^{k-1}f(x+h)(v_1,\dots,v_{k-1}) - D^{k-1}f(x)(v_1,\dots,v_{k-1}) - D^kf(x)(h,v_1,\dots,v_{k-1}) \Rightarrow D^{k-1}R(0)=0$$

Dalej różniczkować po h nie możemy, bo nie mamy gwarancji, czy $D^kf(x+h)$ istnieje. Oszacujmy zatem $R(h)$:

$$\|R(h)\|_Y = \|R(h) - R(0)\|_Y \underset{\substack{\uparrow \\ \text{o ile } \|h\|<\varepsilon, \\ \text{czyli } [x, x+h] \subset \mathcal{U}}}{\leq} \|h\|_X \cdot \sup_{t\in[0,1]} |||DR(th)||| =$$

Funkcja $t \longmapsto |||DR(th)|||$ jest ciągła na $[0,1]$, więc osiąga swoje supremum — w jakimś punkcie $\theta_1 \in [0,1]$

$$= \|h\|_X \cdot |||DR(\theta_1 h)||| = \|h\|_X \, |||DR(\theta_1 h) - DR(0)||| \leq$$

$$\leq \|h\|_X^2 \cdot \sup_{t\in[0,\theta_1]} |||D^2R(th)||| \leq \|h\|_X^2 \sup_{t\in[0,1]} |||D^2R(th)||| =$$

$$= \|h\|_X^2 \, |||D^2R(\theta_2 h)||| = \ldots \quad \text{i tak dalej, aż do}$$

dla pewnego $\theta_2 \in [0,t]$

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

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

$$= r(th), \text{ gdzie } \frac{r(z)}{\|z\|} \underset{z\to 0}{\to} 0,$$

(to są równości w przestrzeni Banacha przekształceń $(k-1)$-liniowych z X w Y) bo $D^{k-1}f$ jest różniczkowalna w x.

$$\leq \|h\|_X^k \cdot \sup_{t\in[0,1]} \frac{|||r(th)|||}{\|th\|}$$

więc ostatecznie

$$\frac{\|R(h)\|_Y}{\|h\|_X^k} \leq \sup_{t\in[0,1]} \frac{|||r(th)|||}{\|th\|_X} \xrightarrow[h\to 0]{} 0.$$

□.

Przypomnienie z GALu:

Jeżeli $T: X \times X \to \mathbb{R}$ jest formą symetryczną dwuliniową,

to mówimy, że T jest

oznaczenie

- dodatnio półokreślona, gdy $\forall_{v \in X}\ T(v,v) \geq 0$ | $T \geq 0$
- dodatnio określona, gdy $\forall_{v \in X, v \neq 0}\ T(v,v) > 0$ | $T > 0$
- ujemnie półokreślona, gdy $\forall_{v \in X}\ T(v,v) \leq 0$ | $T \leq 0$
- ujemnie określona, gdy $\forall_{v \in X, v \neq 0}\ T(v,v) < 0$. | $T < 0$

Gdy X jest skończenie wymiarowa, T ma macierz, która, jako symetryczna, jest diagonalizowalna;

T jest

- dodatnio półokreślona, gdy wartości własne tej macierzy są nieujemne
- dodatnio określona, gdy są dodatnie
- ujemnie półokreślona, gdy są niedodatnie

i - ujemnie określona, gdy są ujemne.

Określoność macierzy możemy badać np. przy pomocy kryterium Sylvestera.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

Niech $d_1 = a_{11}$

$$d_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$d_3 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

itd, $d_n = \det A$.

Jeżeli dla $k = 1, 2, \dots, n$ $\quad d_k > 0$, to $A > 0$

$\exists_k \; d_k < 0$, to $A \not> 0$

$(-1)^k d_k > 0$, to $A < 0$

$\exists_k \; (-1)^k d_k < 0$, to $A \not< 0$.

Wnioski z wzoru Taylora

Warunki konieczne i dostateczne istnienia ekstremum lokalnego:

Twierdzenie (warunek konieczny). Niech $U \subset \mathbb{R}^n$.
Załóżmy, że $f: U \to \mathbb{R}$ ma w $a \in U$
a) minimum lokalne b) maksimum lokalne
oraz że f jest dwukrotnie różniczkowalna w a. (wystarczy, że istnieje w a druga poch. Gateaux)
Wówczas a) $D^2 f(a) \geqslant 0$ b) $D^2 f(a) \leqslant 0$

Dowód: a)
Z tw. Fermata $Df(a) = 0$. Postępując jak w pierwszym podejściu do dowodu wzoru Taylora, ustalamy $v \in \mathbb{R}^n \setminus \{0\}$ i kładziemy $g(t) = f(a+tv)$.
Wtedy

$$f(a+tv) = g(t) = g(0) + g'(0)t + g''(0)\frac{t^2}{2} + r(t)$$
$$= f(a) + t\underbrace{Df(a)v}_{=0} + \frac{t^2}{2} D^2 f(a)(v,v) + r(t)$$

skąd $$D^2 f(a)(v,v) + 2\frac{r(t)}{t^2} = 2\,\frac{f(a+tv) - f(a)}{t^2},$$
gdzie o $r(t)$ wiemy, że $r(0) = 0 = \lim_{t\to 0} \frac{r(t)}{t^2}$.

Gdyby $D^2 f(a)(v,v) < 0$, to istniałoby $\delta > 0$ takie, że $\forall_{t \in (-\delta,\delta)}$ $\left|\frac{r(t)}{t^2}\right| < -\frac{1}{4} D^2 f(a)(v,v)$

dowód (oczywiście wystarczy 1) i 2)).

1). Załóżmy, że $D^2f(a) > 0$, tzn. $\forall_{v \in \mathbb{R}^n \setminus \{0\}} \; D^2f(a)(v,v) > 0$.

Funkcja
$$\begin{array}{ccc} S^{n-1} & \longrightarrow & \mathbb{R} \\ \Psi & & \Psi \\ v & \longmapsto & D^2f(a)(v,v) \end{array}$$
jest
- ciągła (jako wielomian kwadratowy od współrzędnych v)
- dodatnia.

Skoro jest ciągła, to przyjmuje na (zwartej) S^{n-1} swoje minimum: istnieje $v_0 \in \mathbb{R}^n$, $\|v_0\| = 1$ (tzn. $v_0 \in S^{n-1}$) takie, że $\forall_{v \in S^{n-1}} \; D^2f(a)(v,v) \geq \underbrace{D^2f(a)(v_0,v_0)}_{\text{oznaczmy to przez } \alpha \,(>0)} > 0$

Niech $w \in \mathbb{R}^n$, $w \neq 0$. Wtedy
$$D^2f(a)(w,w) = D^2f(a)\Big(\underbrace{\tfrac{w}{\|w\|}}_{\in S^{n-1}}, \tfrac{w}{\|w\|}\Big)\|w\|^2 \geq \alpha\|w\|^2$$
(ten wzór zachodzi również dla $w = 0$).

Jeżeli $a+h \in \mathcal{U}$, to z wzoru Taylora (i tw. Fermata)
$$f(a+h) - f(a) = \frac{1}{2}D^2f(a)(h,h) + R(h), \text{ gdzie } \frac{R(h)}{\|h\|^2} \xrightarrow[h\to 0]{} 0$$

stąd istnieje $\delta > 0$ takie, że jeżeli $\|h\| < \delta$, to
$$\left|\frac{R(h)}{\|h\|^2}\right| < \frac{\alpha}{4} \quad \left(\text{czyli } |R(h)| < \frac{\alpha}{4}\|h\|^2\right)$$

Wtedy, dla $x \in B(a,\delta)$, biorąc $h = x - a$,
$$f(x) - f(a) = \frac{1}{2}D^2f(a)(h,h) + R(h) \geq \frac{1}{2}\alpha\|h\|^2 - \frac{\alpha}{4}\|h\|^2 = \frac{1}{4}\alpha\|h\|^2$$