Różniczki wyższych rzędów, ich symetria, wzór Taylora

Ostatnio poprawiłem 6 grudnia 2014 r.
Duża część zadań pochodzi od dr Marcina Kuczmy

Niech $f\colon G \longrightarrow \mathbb{R}^l$ oznacza funkcję różniczkowalną we wszystkich punktach zbioru G otwartego w $\mathbb{R}^k$. Jeśli wtedy $Df\colon G \longrightarrow \mathcal{L}(\mathbb{R}^k;\mathbb{R}^l)$ jest funkcją różniczkowalną w punkcie $\mathbf{p} \in G$, to jej różniczkę $D(Df)(\mathbf{p})$ oznaczamy symbolem $D^2f(\mathbf{p})$ i nazywamy drugą różniczką funkcji f w punkcie $\mathbf{p}$. Funkcjami współrzędnymi odwzorowania Df są pochodne cząstkowe funkcji f, czyli funkcje $\frac{\partial f}{\partial x_i}\colon G \longrightarrow \mathbb{R}^l$. Oznaczamy $\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{p}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{p})$, wyrażenie to nazywamy pochodną cząstkową mieszaną drugiego rzędu funkcji f w punkcie $\mathbf{p}$ względem zmiennych x_i i x_j. Jeżeli $i = j$, to piszemy $\frac{\partial^2 f}{\partial x_i^2}(\mathbf{p})$.

Przykład 5.1 $f(\mathbf{x}) = L\mathbf{x}$, tu L oznacza przekształcenie liniowe z $\mathbb{R}^k$ do $\mathbb{R}^l$. Dla każdego punktu $\mathbf{p} \in \mathbb{R}^k$ mamy $Df(\mathbf{p}) = L$, zatem przekształcenie $Df\colon \mathbb{R}^k \longrightarrow \mathcal{L}(\mathbb{R}^k;\mathbb{R}^l)$ jest przekształceniem stałym, więc $D^2f(\mathbf{x}) = \mathbf{0}$ dla każdego $\mathbf{x} \in \mathbb{R}^k$. $\blacksquare$

Przykład 5.2 Niech A będzie macierzą kwadratową wymiaru k i niech $f(\mathbf{x}) = A\mathbf{x}\cdot\mathbf{x}$ dla $\mathbf{x} \in \mathbb{R}^k$. Z twierdzenia o pochodnej przekształcenia dwuliniowego (lub wprost z definicji różniczki) wynika, że $Df(\mathbf{x})\mathbf{h} = A\mathbf{x}\cdot\mathbf{h}+A\mathbf{h}\cdot\mathbf{x} = A\mathbf{x}\cdot\mathbf{h}+\mathbf{h}\cdot A^T\mathbf{x} = (A+A^T)\mathbf{x}\cdot\mathbf{h}$, czyli $Df(\mathbf{x}) = (A + A^T)\mathbf{x}$. Wobec tego $Df\colon \mathbb{R}^k \longrightarrow \mathcal{L}(\mathbb{R}^k;\mathbb{R})$ jest przekształceniem liniowym, zatem $D^2f(\mathbf{x}) = A + A^T$ dla każdego $\mathbf{x} \in \mathbb{R}^k$. Zauważmy jeszcze, że jeśli macierz A jest symetryczna, czyli $A = A^T$, to $D^2f = 2A$, co może przypominać równość $(ax^2)'' = 2a$. $\blacksquare$

Przykład 5.3 Niech $B \in \mathcal{L}(\mathbb{R}^{k_1},\mathbb{R}^{k_2},\ldots,\mathbb{R}^{k_n};\mathbb{R}^l)$ będzie przekształceniem n–liniowym. Bez trudu sprawdzić można, że

$$\big(D^2B(\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n)(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_n)\big)(\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_n) = \\ = \sum_{i\neq j} B(\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_{i-1},\mathbf{h}_i,\mathbf{p}_{i+1},\ldots,\mathbf{p}_{j-1},\mathbf{H}_j,\mathbf{p}_{j+1},\ldots\mathbf{p}_n),$$

oczywiście mamy tu do czynienia z sumowaniem $n(n-1)$ składników, nie wiadomo, czy $i > j$, czy też $j > i$. Można łatwo zauważyć, że ten wzór jest uogólnieniem znanej, np. z pierwszego roku, równości $(x^n)'' = n(n-1)x^{n-2}$. Zestawy $\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_n$ oraz $\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_n$ to argumenty: pierwszy zestaw to argumenty przekształcenia $D^2B(\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n)$, a drugi to argumenty jego wartości, a jest nią przekształcenie $D^2B(\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n)(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_n) \in \mathcal{L}(\mathbb{R}^{k_1+k_2+\cdots+k_n};\mathbb{R}^l)$. $\blacksquare$

Przykład 5.4 Niech $f(x,y) = xy\frac{x^2-y^2}{x^2+y^2}$ dla $(x,y) \neq (0,0)$ i niech $f(0,0) = 0$. Określiliśmy więc na płaszczyźnie funkcję, której wartościami są liczby rzeczywiste. Znajdziemy jej pochodne mieszane w punkcie $(0,0)$. Zauważmy w tym celu, że $\frac{\partial f}{\partial x}(0,y) = -y$ oraz $\frac{\partial f}{\partial y}(x,0) = x$ – wzory tej najłatwiej wywnioskować z definicji pochodnej. Z nich wynika od razu, że $\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial}{\partial x}\big(\frac{\partial f}{\partial y}\big)(0,0) = 1$ oraz $\frac{\partial^2 f}{\partial y \partial x}(0,0) = \frac{\partial}{\partial y}\big(\frac{\partial f}{\partial x}\big)(0,0) = -1$, zatem w tym przypadku pochodne mieszane drugiego rzędu w punkcie $\mathbf{0}$ są różne. $\blacksquare$

Czytelnicy z pewnością zauważyli, że trzeci przykład uogólnia dwa poprzednie. Zajmiemy się teraz twierdzeniem, z którego wynika, że funkcję z przykładu czwartego uznać należy za „wysoce nietypową". Zacznijmy od ważnego stwierdzenia: przekształcenia liniowe z $\mathbb{R}^k$ w $\mathcal{L}(\mathbb{R}^k,\mathbb{R}^l)$ można utożsamiać z przekształceniami dwuliniowymi z $\mathbb{R}^k\times\mathbb{R}^k$ w $\mathbb{R}^l$: $\big(L(\mathbf{u})\big)(\mathbf{v})=\tilde{L}(\mathbf{u},\mathbf{v})$. Przekształceniu liniowemu L przypisujemy przekształcenie dwuliniowe $\tilde{L}$. Czytelnikowi nie sprawi najmniejszych trudności sprawdzenie, że opisane przyporządkowanie jest izomorfizmem przestrzeni liniowej $\mathcal{L}\big(\mathbb{R}^k;\mathcal{L}(\mathbb{R}^k;\mathbb{R}^l)\big)$ i przestrzeni liniowej $\mathcal{L}(\mathbb{R}^k,\mathbb{R}^k;\mathbb{R}^l)$. W dalszym ciągu będziemy utożsamiać drugą różniczkę funkcji w punkcie z przekształceniem dwuliniowym, bo tak jest wygodniej.

Twierdzenie 5.5 Schwarza o symetrii drugiej różniczki[1]

1° Jeśli funkcja $f\colon G\longrightarrow\mathbb{R}^l$ jest dwukrotnie różniczkowalna w punkcie $\mathbf{p}$, to dla dowolnych $\mathbf{u},\mathbf{v}\in\mathbb{R}^k$ zachodzi równość $D^2f(\mathbf{p})(\mathbf{u},\mathbf{v})=D^2f(\mathbf{p})(\mathbf{v},\mathbf{u})$.

2° Jeśli funkcja $f\colon G\longrightarrow\mathbb{R}^l$ ma w zbiorze G pochodne mieszane $\frac{\partial^2 f}{\partial x_i\partial x_j}\colon G\longrightarrow\mathbb{R}^l$ oraz $\frac{\partial^2 f}{\partial x_j\partial x_i}\colon G\longrightarrow\mathbb{R}^l$ i obie te pochodne są ciągłe w punkcie $\mathbf{p}\in G$, to zachodzi równość $\frac{\partial^2 f}{\partial x_i\partial x_j}(\mathbf{p})=\frac{\partial^2 f}{\partial x_j\partial x_i}(\mathbf{p})$.

Przed dowodem zauważmy, że w punktach 1° i 2° założenia są różne: w pierwszym przypadku założona została dwukrotna różniczkowalność funkcji f w jednym punkcie dziedziny funkcji f, a w drugim – założyliśmy istnienie pary pochodnych mieszanych drugiego rzędu w całym otoczeniu ustalonego punktu i ich ciągłość w tym punkcie. Dodatkowa różnica w naszym sformułowaniu polega na tym, że w drugim przypadku mówimy o pochodnych cząstkowych, a w pierwszym o pochodnych kierunkowych, ale to nie jest istotne. Jasne jest, że oba punkty są niezależne, choć dotyczą podobnych sytuacji. Przejdziemy do uzasadnienia twierdzenia.

Dowód. Zaczniemy od punktu 2°. Możemy założyć, że wartościami funkcji f są liczby rzeczywiste, bo wartości interesującej nas funkcji mają l współrzędnych, więc możemy zajmować się kolejno każdą z l funkcji. Niech

$$\varrho(s,t)=f(\mathbf{p}+s\mathbf{e}_i+t\mathbf{e}_j)-f(\mathbf{p}+s\mathbf{e}_i)-f(\mathbf{p}+t\mathbf{e}_j)+f(\mathbf{p}).$$

Funkcję ϱ możemy potraktować jako funkcję zmiennej t. Mamy więc do czynienia z funkcją określoną na pewnym przedziale. Jej wartościami są liczby rzeczywiste. Pozwala to na stosowanie twierdzenia Lagrange'a o wartości średniej. Istnieje więc liczba rzeczywista τ leżąca między liczbami 0 i t taka, że

$$\varrho(s,t)=\varrho(s,t)-\varrho(s,0)=t\Big(\tfrac{\partial f}{\partial x_j}(\mathbf{p}+s\mathbf{e}_i+\tau\mathbf{e}_j)-\tfrac{\partial f}{\partial x_j}(\mathbf{p}+\tau\mathbf{e}_j)\Big).$$

Wyrażenie w nawiasie z prawej strony tej równości potraktujemy teraz jako funkcję zmiennej s. Możemy więc znów zastosować twierdzenie Lagrange'a o wartości średniej. Możemy więc znaleźć między liczbami 0 i s liczbę σ taką, że

$$\varrho(s,t)=ts\tfrac{d}{ds}\Big(\tfrac{\partial f}{\partial x_j}(\mathbf{p}+s\mathbf{e}_i+\tau\mathbf{e}_j)-\tfrac{\partial f}{\partial x_j}(\mathbf{p}+\tau\mathbf{e}_j)\Big)\Big|_{s=\sigma}=ts\tfrac{\partial^2 f}{\partial x_j\partial x_i}(\mathbf{p}+\sigma\mathbf{e}_i+\tau\mathbf{e}_j).$$

Z otrzymanego wzoru wynika od razu, że

[1] Wg. co najmniej niektórych obywateli USA jest to twierdzenie Younga.

$$\lim_{(s,t)\to(0,0)} \frac{\varrho(s,t)}{st} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{p}).$$

Powtarzając to samo rozumowanie z jedną zmianą: najpierw traktujemy wyrażenie $\varrho(s,t)$ jako funkcję zmiennej s o potem jako funkcję zmiennej t otrzymujemy równość

$$\lim_{(s,t)\to(0,0)} \frac{\varrho(s,t)}{st} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{p}).$$

Z dwu ostatnich równości wnioskujemy, że $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{p}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{p})$. Dowód drugiej części twierdzenia został zakończony, więc możemy przejść do dowodu części pierwszej.

1° Tym razem nie będziemy mogli korzystać z twierdzenia o wartości średniej dwukrotnie, bo założenie gwarantuje nam istnienie drugiej różniczki tylko w jednym punkcie, w pozostałych — funkcja może nawet nie mieć pochodnych cząstkowych drugiego rzędu. Skorzystamy więc z twierdzenia o wartości średniej tylko raz, ale dowód przeprowadzimy w taki sposób, by bez żadnych zmian można było go zastosować również do funkcji o wartościach w nieskończenie wymiarowej przestrzeni liniowej.

Niech $\varrho(s,t) = f(\mathbf{p}+s\mathbf{v}+t\mathbf{w}) - f(\mathbf{p}+s\mathbf{v}) - f(\mathbf{p}+t\mathbf{w}) + f(\mathbf{p}) - D^2 f(\mathbf{p})(t\mathbf{w}, s\mathbf{v})$. Funkcja ϱ jest różniczkowalna w pewnym otoczeniu punktu $\mathbf{0} \in \mathbb{R}^2$, bo funkcja f jest różniczkowalna w pewnym otoczeniu punktu $\mathbf{p}$ (wynika to z twierdzenia o różniczce złożenia). Zachodzi wzór

$$\frac{\partial \varrho}{\partial s}(s,t) = Df(\mathbf{p}+s\mathbf{v}+t\mathbf{w})\mathbf{v} - Df(\mathbf{p}+s\mathbf{v})\mathbf{v} - D^2 f(\mathbf{p})(t\mathbf{w}, \mathbf{v})$$

Z tego wzoru i twierdzenia o wartości średniej wynika, że

$$\|\varrho(s,t)\| = \|\varrho(s,t) - \varrho(0,t)\| \leqslant$$
$$\leqslant |s| \sup_\sigma \|Df(\mathbf{p}+\sigma\mathbf{v}+t\mathbf{w})\mathbf{v} - Df(\mathbf{p}+\sigma\mathbf{v})\mathbf{v} - D^2 f(\mathbf{p})(t\mathbf{w}, \mathbf{v})\|,$$

przy czym σ przebiega przedział o końcach 0 i s. Z istnienia drugiej różniczki funkcji f w punkcie $\mathbf{p}$ wynika istnienie takiej funkcji $\mathbf{r}$, że

$$Df(\mathbf{p}+\mathbf{h}) = Df(\mathbf{p}) + D^2 f(\mathbf{p})\mathbf{h} + \|\mathbf{h}\|\mathbf{r}(\mathbf{h}) \text{ i } \lim_{\mathbf{h}\to\mathbf{0}} \mathbf{r}(\mathbf{h}) = \mathbf{0}.$$

Mamy więc

$$\Big\|Df(\mathbf{p}+\sigma\mathbf{v}+t\mathbf{w})\mathbf{v} - Df(\mathbf{p}+\sigma\mathbf{v})\mathbf{v} - D^2 f(\mathbf{p})(t\mathbf{w}, \mathbf{v})\Big\| =$$
$$= \Big\|Df(\mathbf{p}) + \big(D^2 f(\mathbf{p})(\sigma\mathbf{v}+t\mathbf{w})\big)\mathbf{v} + \|\sigma\mathbf{v}+t\mathbf{w}\| \cdot \mathbf{r}(\sigma\mathbf{v}+t\mathbf{w}) -$$
$$- \Big(Df(\mathbf{p}) + \big(D^2 f(\mathbf{p})(\sigma\mathbf{v})\big)(\mathbf{v}) + \|\sigma\mathbf{v}\| \cdot \mathbf{r}(\sigma\mathbf{v})\Big) - D^2 f(\mathbf{p})(t\mathbf{w}, \mathbf{v})\Big\| =$$
$$= \Big\|D^2 f(\mathbf{p})(\sigma\mathbf{v}+t\mathbf{w}, \mathbf{v}) - D^2 f(\mathbf{p})(\sigma\mathbf{v}, \mathbf{v}) - D^2 f(\mathbf{p})(t\mathbf{w}, \mathbf{v}) +$$
$$+ \|\sigma\mathbf{v}+t\mathbf{w}\| \cdot \mathbf{r}(\sigma\mathbf{v}+t\mathbf{w}) - \|\sigma\mathbf{v}\| \cdot \mathbf{r}(\sigma\mathbf{v})\Big\| =$$
$$= \Big\|\|\sigma\mathbf{v}+t\mathbf{w}\| \cdot \mathbf{r}(\sigma\mathbf{v}+t\mathbf{w}) - \|\sigma\mathbf{v}\| \cdot \mathbf{r}(\sigma\mathbf{v})\Big\|.$$

Możemy przyjąć $s = t$. Dla $t > 0$ mamy więc

$$\frac{\|\varrho(t,t)\|}{t^2} \leqslant \frac{1}{t} \cdot \sup_{0<\sigma<t} \Big\|\|\sigma\mathbf{v}+t\mathbf{w}\| \cdot \mathbf{r}(\sigma\mathbf{v}+t\mathbf{w}) - \|\sigma\mathbf{v}\| \cdot \mathbf{r}(\sigma\mathbf{v})\Big\| \xrightarrow[t\to 0]{} 0.$$

Wynika stąd, że $\lim\limits_{t\to 0} \frac{f(\mathbf{p}+t\mathbf{v}+t\mathbf{w}) - f(\mathbf{p}+t\mathbf{v}) - f(\mathbf{p}+t\mathbf{w}) + f(\mathbf{p})}{t^2} = D^2 f(\mathbf{p})(\mathbf{w}, \mathbf{v})$.

Przyjmując, że

$$\tilde{\varrho}(s,t) = f(\mathbf{p}+s\mathbf{v}+t\mathbf{w}) - f(\mathbf{p}+s\mathbf{v}) - f(\mathbf{p}+t\mathbf{w}) + f(\mathbf{p}) - D^2 f(\mathbf{p})(s\mathbf{v}, t\mathbf{w})$$

dowodzimy tak samo [2], że $\lim\limits_{t\to 0} \frac{f(\mathbf{p}+t\mathbf{v}+t\mathbf{w}) - f(\mathbf{p}+t\mathbf{v}) - f(\mathbf{p}+t\mathbf{w}) + f(\mathbf{p})}{t^2} = D^2 f(\mathbf{p})(\mathbf{v}, \mathbf{w})$.

Wynika stąd od razu, że $D^2 f(\mathbf{p})(\mathbf{v}, \mathbf{w}) = D^2 f(\mathbf{p})(\mathbf{w}, \mathbf{v})$, co kończy dowód. ■

[2] jedyna zmiana polega na tym, że teraz zaczynamy od zmiennej t

Uwaga 5.6 Dowód twierdzenie o symetrii drugiej różniczki można przeprowadzić nieco inaczej, co miało miejsce w czasie wykładu 28 listopada 2014 r. Można korzystać z wielowymiarowego twierdzenia o wartości średniej. Pokażemy jak można to zrobić. Niech

$$\varrho(\mathbf{u},\mathbf{v}) = f(\mathbf{p}+\mathbf{u}+\mathbf{v}) - f(\mathbf{p}+\mathbf{u}) - f(\mathbf{p}+\mathbf{v}) + f(\mathbf{p}) - D^2 f(\mathbf{p})(\mathbf{u},\mathbf{v}).$$

Z twierdzenia o wartości średniej wynika, że

$\|f(\mathbf{p}+\mathbf{u}+\mathbf{v}) - f(\mathbf{p}+\mathbf{u}) - f(\mathbf{p}+\mathbf{v}) + f(\mathbf{p}) - D^2 f(\mathbf{p})(\mathbf{u},\mathbf{v})\| \leqslant$

$\leqslant \|\mathbf{v}\| \sup_{0\leqslant s\leqslant 1} \|Df(\mathbf{p}+\mathbf{u}+s\mathbf{v}) - Df(\mathbf{p}+s\mathbf{v}) - D^2 f(\mathbf{u},\bullet)\|$

symbol $\bullet$ oznacza miejsce na argument przekształcenia liniowego, natomiast argument $\mathbf{u}$ jest chwilowo ustalony. Rozpatrujemy liniowo niezależne wektory $\mathbf{u},\mathbf{v}$, bo dla liniowo zależnych nie ma nic do dowodu (wtedy np. $\mathbf{u} = \alpha\mathbf{v}$ dla pewnej liczby $\alpha \in \mathbb{R}$, więc $D^2 f(\mathbf{p})(\mathbf{u},\mathbf{v}) = \alpha D^2 f(\mathbf{p})(\mathbf{v},\mathbf{v}) = D^2 f(\mathbf{p})(\mathbf{v},\alpha\mathbf{v}) = D^2 f(\mathbf{p})(\mathbf{v},\mathbf{u})$). Załóżmy na chwilę, że $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. Niech $c = \inf_{0\leqslant s\leqslant 1} \|\mathbf{u}+s\mathbf{v}\|$, $C = \sup_{0\leqslant s\leqslant 1} \|\mathbf{u}+s\mathbf{v}\|$. Ponieważ funkcja $s \mapsto \|\mathbf{u}+s\mathbf{v}\|$ jest ciągła, więc osiąga swe kresy na zbiorze zwartym $[0,1]$, zatem $0 < c \leqslant C < \infty$. Wynika stąd, że jeśli $\|\mathbf{u}\| = \|\mathbf{v}\|$, to $c\|\mathbf{v}\| \leqslant \|\mathbf{u}+s\mathbf{v}\| \leqslant C\|\mathbf{v}\|$, a stąd z kolei wynika, że $o(\|vg\|) = o(\|\mathbf{u}+s\mathbf{v}\|)$. Z definicji $D^2 f(\mathbf{p})\|$ wynika, że

$Df(\mathbf{p}+\mathbf{u}+s\mathbf{v}) = Df(\mathbf{p}) + D^2 f(\mathbf{p})(\mathbf{u}+s\mathbf{v},\bullet) + o(\|\mathbf{u}+s\mathbf{v})\|) =$

$= Df(\mathbf{p}) + D^2 f(\mathbf{p})(\mathbf{u}+s\mathbf{v},\bullet) + o(\|\mathbf{v})\|)$

oraz

$Df(\mathbf{p}+s\mathbf{v}) = Df(\mathbf{p}) + D^2 f(\mathbf{p})(s\mathbf{v},\bullet) + o(\|s\mathbf{v})\|) = Df(\mathbf{p}) + D^2 f(\mathbf{p})(s\mathbf{v},\bullet) + o(\|\mathbf{v})\|)$.

Wobec tego

$\|Df(\mathbf{p}+\mathbf{u}+s\mathbf{v}) - Df(\mathbf{p}+s\mathbf{v}) - D^2 f(\mathbf{u},\bullet)\| = \|Df(\mathbf{p}) + D^2 f(\mathbf{p})(\mathbf{u}+s\mathbf{v},\bullet) + o(\|\mathbf{v})\|) -$
$-(Df(\mathbf{p}) + D^2 f(\mathbf{p})(s\mathbf{v},\bullet) + o(\|\mathbf{v})\|)) - D^2 f(\mathbf{p})(\mathbf{u},\bullet)\| = \|o(\|\mathbf{v})\|)\|$ — pisząc $o(\|\mathbf{v}\|)$ mamy na myśli: przy $\|\mathbf{u}\| = \|\mathbf{v}\| \longrightarrow 0$. Wobec tego

$$\|f(\mathbf{p}+\mathbf{u}+\mathbf{v}) - f(\mathbf{p}+\mathbf{u}) - f(\mathbf{p}+\mathbf{v}) + f(\mathbf{p}) - D^2 f(\mathbf{p})(\mathbf{u},\mathbf{v})\| = o(\|\mathbf{v}\|^2),$$

więc $\lim_{\|\mathbf{u}\|=\|\mathbf{v}\|\to 0} (f(\mathbf{p}+\mathbf{u}+\mathbf{v}) - f(\mathbf{p}+\mathbf{u}) - f(\mathbf{p}+\mathbf{v}) + f(\mathbf{p}))/\|\mathbf{v}\|^2 = D^2 f(\mathbf{p})(\frac{\mathbf{u}}{\|\mathbf{u}\|},\frac{\mathbf{v}}{\|\mathbf{v}\|})$. Stąd i z symetrii wyrażenia $f(\mathbf{p}+\mathbf{u}+\mathbf{v}) - f(\mathbf{p}+\mathbf{u}) - f(\mathbf{p}+\mathbf{v}) + f(\mathbf{p})$ wynika od razu, że jeśli $\|\mathbf{u}\| = \|\mathbf{v}\|$, to $D^2 f(\mathbf{p})(\mathbf{u},\mathbf{v}) = D^2 f(\mathbf{p})(\mathbf{v},\mathbf{u})$, a z jednorodności przekształcenia $D^2 f(\mathbf{p})$ wynika, że

$D^2 f(\mathbf{p})(\mathbf{u},\mathbf{v}) = \|\mathbf{u}\|\cdot\|\mathbf{v}\| D^2 f(\mathbf{p})(\frac{\mathbf{u}}{\|\mathbf{u}\|},\frac{\mathbf{v}}{\|\mathbf{v}\|}) = \|\mathbf{u}\|\cdot\|\mathbf{v}\| D^2 f(\mathbf{p})(\frac{\mathbf{v}}{\|\mathbf{v}\|},\frac{\mathbf{u}}{\|\mathbf{u}\|}) = D^2 f(\mathbf{p})(\mathbf{v},\mathbf{u})$

dla dowolnych wektorów $\mathbf{u},\mathbf{v}$. ■

Z twierdzenia o symetrii drugiej różniczki wynika, że zajmując się „porządnymi" funkcjami możemy nie przejmować się kolejnością różniczkowań.

Jasne jest, że można indukcyjnie zdefiniować pochodne (różniczki) wyższych rzędów:

$$D^{(n+1)} f(\mathbf{p}) = D(D^{(n)} f)(\mathbf{p}).$$

Różniczkę drugiego rzędu utożsamiamy z przekształceniem dwuliniowym. Podobnie różniczkę n–tego rzędu utożsamiamy z przekształceniem n–liniowym. Z twierdzenia o symetrii drugiej różniczki wynika, że n–ta różniczka również jest symetryczna, natomiast pochodne cząstkowe wyższego rzędu w punktach nieróżniczkowalności nie muszą zachowywać się dobrze chyba, że są ciągłe. Pochodne pozwalają na przybliżanie funkcji jednej zmiennej wielomianami. Okazuje się, że nie tylko w wymiarze 1, w wyższych wymiarach również. Chodzi oczywiście o wzór Taylora. Od tej pory obowiązywać będzie umowa: zamiast $D^n f(\mathbf{p})(\mathbf{h},\mathbf{h},\ldots,\mathbf{h})$ będziemy pisać $D^n f(\mathbf{p})\mathbf{h}^n$ dla wygody i podkreślenia tego, że iloczyn n czynników to szczególny, ale ważny, przypadek przekształcenia n–liniowego.

Twierdzenie 5.7 (wzór Taylora z resztą Peano)
Jeśli funkcja $f\colon G \longrightarrow \mathbb{R}^l$ jest n–krotnie różniczkowalna w punkcie $\mathbf{p}\in G$ i $\mathbf{r}_n$ oznacza n–tą resztę:

$$f(\mathbf{p}{+}\mathbf{h}) = f(\mathbf{p}){+}Df(\mathbf{p})\mathbf{h}{+}\tfrac{1}{2!}D^2f(\mathbf{p})\mathbf{h}^2{+}\cdots{+}\tfrac{1}{n!}D^nf(\mathbf{p})\mathbf{h}^n{+}\mathbf{r}_n(\mathbf{h})\,, \quad \text{to} \lim_{\mathbf{h}\to\mathbf{0}}\frac{\mathbf{r}_n(\mathbf{h})}{\|\mathbf{h}\|^n}=\mathbf{0}.$$

Dowód. Mamy

$$\mathbf{r}_n(\mathbf{h}) = f(\mathbf{p}+\mathbf{h}) - \big(f(\mathbf{p}) + Df(\mathbf{p})\mathbf{h} + \tfrac{1}{2!}D^2f(\mathbf{p})\mathbf{h}^2 + \cdots + \tfrac{1}{n!}D^nf(\mathbf{p})\mathbf{h}^n\big).$$

Stwierdzamy bez trudu, że

$$D\mathbf{r}_n(\mathbf{h}) = Df(\mathbf{p}+\mathbf{h}) - \big(Df(\mathbf{p}) + \tfrac{1}{1!}D^2f(\mathbf{p})\mathbf{h}^1 + \cdots + \tfrac{1}{(n-1)!}D^nf(\mathbf{p})\mathbf{h}^{n-1}\big).$$

Z tego wzoru i z twierdzenia o wartości średniej wynika, że $\|\mathbf{r}_n(\mathbf{h})\| \leqslant$

$$\leqslant \|\mathbf{h}\|\sup_{0<t<1}\big\|Df(\mathbf{p}+t\mathbf{h}) - \big(Df(\mathbf{p}) + \tfrac{1}{1!}D^2f(\mathbf{p})(t\mathbf{h}) + \cdots + \tfrac{1}{(n-1)!}D^nf(\mathbf{p})(t\mathbf{h})^{n-1}\big)\big\|.$$

Zastosowawszy do prawej strony tej równości twierdzenie o wartości średniej i oczywistą nierówność $\|t\mathbf{h}\| \leqslant \|\mathbf{h}\|$ jeszcze $n-2$ razy otrzymujemy

$$\|\mathbf{r}_n(\mathbf{h})\| \leqslant \|\mathbf{h}\|^{n-1}\sup_{0<t<1}\big\|D^{n-1}f(\mathbf{p}+t\mathbf{h}) - D^{(n-1)}f(\mathbf{p}) - D^nf(\mathbf{p})(t\mathbf{h})\big\|.$$

Chciałoby się raz jeszcze skorzystać z twierdzenia o wartości średniej, ale nie można, bo założyliśmy istnienie D^nf tylko w jednym punkcie. Możemy jednak skorzystać wprost z definicji różniczki, by stwierdzić, że

$$\lim_{\mathbf{h}\to\mathbf{0}}\frac{\|D^{n-1}f(\mathbf{p}+\mathbf{h}) - D^{(n-1)}f(\mathbf{p}) - D^nf(\mathbf{p})(\mathbf{h})\|}{\|\mathbf{h}\|} = 0,$$

a stąd i z poprzednio uzyskanych oszacowań teza wynika od razu. ■

W taki sam sposób można wykazać

Twierdzenie 5.8 (Lagrange'a o wzorze Taylora)
Jeśli funkcja $f\colon G \longrightarrow \mathbb{R}^l$ jest $n+1$–krotnie różniczkowalna w otoczeniu punktu $\mathbf{p}\in G$ i $\mathbf{r}_n$ oznacza n–tą resztę, to $\|\mathbf{r}_n(\mathbf{h})\| \leqslant \frac{1}{(n+1)!}\sup\limits_{0<t<1}\big\|D^{(n+1)}f(\mathbf{p}{+}t\mathbf{h})\|\cdot\|\mathbf{h}\|^{n+1}$, przy czym jeśli $l=1$, to istnieje liczba $t\in(0,1)$ taka, że $\mathbf{r}_n(\mathbf{h}) = \frac{1}{(n+1)!}D^{(n+1)}f(\mathbf{p}+t\mathbf{h})\mathbf{h}^{n+1}$.
Dowód. Równość dla funkcji o wartościach rzeczywistych uzyskujemy stosując twierdzenie o reszcie we wzorze Taylora do funkcji przypisującej liczbie t liczbę $f(\mathbf{p}+t\mathbf{h})$. Wartości jej kolejnych pochodnych w punkcie t są równe

$$Df(\mathbf{p}+t\mathbf{h})\mathbf{h},\ D^2f(\mathbf{p}+t\mathbf{h})\mathbf{h}^2,\ \ldots, D^nf(\mathbf{p}+t\mathbf{h})\mathbf{h}^n,\ \ D^{(n+1)}f(\mathbf{p}+t\mathbf{h})\mathbf{h}^{(n+1)}.$$

Stąd teza wynika natychmiast. ■

Definicja 5.9 (wielowskaźnika[3] i operacji z nim związanych)
Wielowskaźnikiem nazywamy dowolny wektor $\mathbf{j}=(j_1,j_2,\ldots,j_k)$, którego współrzędne są liczbami całkowitymi nieujemnymi. Definiujemy

$$|\mathbf{j}| = j_1+j_2+\cdots+j_k,\ \ \mathbf{h}^{\mathbf{j}} = h_1^{j_1}h_2^{j_2}\ldots h_k^{j_k},\ \mathbf{j}! = j_1!j_2!\ldots j_k!,\ \ \frac{\partial^{|\mathbf{j}|}}{\partial\mathbf{x}^{\mathbf{j}}} = \frac{\partial^{j_1+j_2+\cdots+j_k}}{\partial x_1^{j_1}\partial x_2^{j_2}\ldots\partial x_k^{j_k}}.\ \ ■$$

Zadanie 5.1 Wykazać, że $(x_1+x_2+\cdots+x_k)^n = \sum\limits_{|\mathbf{j}|=n}\frac{n!}{\mathbf{j}!}\mathbf{x}^{\mathbf{j}}$.

[3] Niektórzy polubili słowo multiindeks.

Uwaga 5.10 (Inny zapis wielomianu Taylora)

$$f(\mathbf{p}) + Df(\mathbf{p})\mathbf{h} + \tfrac{1}{2!}D^2f(\mathbf{p})\mathbf{h}^2 + \cdots + \tfrac{1}{n!}D^nf(\mathbf{p})\mathbf{h}^n = \sum_{|\mathbf{j}|\leqslant n} \tfrac{1}{\mathbf{j}!}\tfrac{\partial^{|\mathbf{j}|}f}{\partial \mathbf{x}^{\mathbf{j}}}(\mathbf{p})\mathbf{h}^{\mathbf{j}}.$$

Ten wzór wynika z tego, że jeśli $|\mathbf{j}| = n$, to liczba ciągów w których h_s występuje j_s razy równa jest $\frac{n}{\mathbf{j}!}$: j_1 miejsc spośród n miejsc wybieramy na $\binom{n}{j_1}$ sposobów, j_2 miejsca spośród pozostałych $n - j_1$ miejsc wybieramy na $\binom{n-j_1}{j_2}$ sposobów itd. Iloczyn otrzymanych liczb to właśnie $\frac{n!}{\mathbf{j}!}$. ■

Definicja 5.11 Druga pochodna kierunkowa
Jeśli f jest dwukrotnie różniczkowalna w punkcie $\mathbf{p}$ i $g(t) = f(\mathbf{p} + t\mathbf{v})$, to zachodzi równość $g''(0) = D^2f(\mathbf{p})\mathbf{v}^2$.
Dowód. $g'(t) = Df(\mathbf{p} + t\mathbf{v})\mathbf{v}$ – ta równość wynika od razu z twierdzenia o różniczce złożenia. Z tej równości dowodzony wzór wynika od razu, można np. skorzystać z definicji różniczki w punkcie lub z twierdzenia o różniczce złożenia dla trzech przekształceń ($t \to \mathbf{p} + t\mathbf{v}$, $\mathbf{x} \to f(\mathbf{x})$, $\mathbf{y} \to \mathbf{y}\cdot\mathbf{v}$). ■

Sformułujemy teraz twierdzenie o ekstremach lokalnych. Przypomnijmy, że funkcja f, o wartościach rzeczywistych, ma w punkcie $\mathbf{p}$ lokalne minimum wtedy i tylko wtedy, gdy istnieje otoczenie U punktu $\mathbf{p}$ takie, że dla każdego $\mathbf{x} \in U$ zachodzi nierówność $f(\mathbf{x}) \geqslant f(\mathbf{p})$. Jeśli ta nierówność jest ostra dla każdego $\mathbf{x} \in U \setminus \{\mathbf{p}\}$, to mówimy, że minimum jest właściwe. Jeśli funkcja określona jest na jednowymiarowym przedziale, to warunkiem koniecznym na to, by w miała lokalne minimum w punkcie $\mathbf{p}$ jest zerowanie się pierwszej pochodnej, a dostatecznym — by pierwsza pochodna przyjmowała wartość 0 i by druga pochodna była dodatnia. Twierdzenie to ma swój odpowiednik wielowymiarowy.

Twierdzenie 5.12 (o ekstremach lokalnych, warunek dostateczny)
Załóżmy, że funkcja $f\colon G \longrightarrow \mathbb{R}$ jest dwukrotnie różniczkowalna w punkcie $\mathbf{p}$ oraz że $Df(\mathbf{p}) = \mathbf{0}$. W tej sytuacji

a. jeśli $D^2f(\mathbf{p})\mathbf{v}^2 > 0$ dla każdego wektora $\mathbf{v} \neq \mathbf{0}$ ($D^2f(\mathbf{p})$ jest dodatnio określona), to f ma lokalne minimum właściwe w punkcie $\mathbf{p}$;

b. jeśli $D^2f(\mathbf{p})\mathbf{v}^2 < 0$ dla każdego wektora $\mathbf{v} \neq \mathbf{0}$ ($D^2f(\mathbf{p})$ jest ujemnie określona), to f ma lokalne maksimum właściwe w punkcie $\mathbf{p}$;

c. jeśli $D^2f(\mathbf{p})\mathbf{v}^2 > 0 > D^2f(\mathbf{p})\mathbf{w}^2$ dla pewnej pary wektorów $\mathbf{v}, \mathbf{w}$, to f nie ma lokalnego ekstremum w punkcie $\mathbf{p}$.

Dowód. Jeśli funkcja f jest dwukrotnie różniczkowalna w punkcie $\mathbf{p}$ i $Df(\mathbf{p}) = \mathbf{0}$, to zachodzi równość

$$f(\mathbf{p}+\mathbf{h}) = f(\mathbf{p}) + D^2f(\mathbf{p})\mathbf{h}^2 + \|\mathbf{h}\|^2\mathbf{r}_2(\mathbf{h}) = f(\mathbf{p}) + \|\mathbf{h}\|^2\big(D^2f(\mathbf{p})\tfrac{\mathbf{h}^2}{\|\mathbf{h}\|^2} + \mathbf{r}_2(\mathbf{h})\big)$$

przy czym $\mathbf{r}_2(\mathbf{h}) \xrightarrow[\mathbf{h}\to\mathbf{0}]{} 0$. Załóżmy, że różniczka $D^2f(\mathbf{p})$ jest dodatnio określona. Jako przekształcenie dwuliniowe jest ona ciągła. Wobec tego funkcja przypisująca wektorowi $\mathbf{x}$ liczbę $D^2f(\mathbf{p})\mathbf{x}^2$ jest ciągła na sferze jednostkowej i wobec tego przyjmuje na tym zbiorze wartość najmniejszą b. Ta najmniejsza wartość b jest liczbą dodatnią. Stąd wynika, że dla każdego $\mathbf{h} \neq \mathbf{0}$ zachodzi nierówność $D^2f(\mathbf{p})\frac{\mathbf{h}^2}{\|\mathbf{h}\|^2} \geqslant b$ i wobec tego istnieje liczba $\delta > 0$ taka, że jeśli $\mathbf{0} \neq \mathbf{h}$ i $\|\mathbf{h}\| < \delta$, to $D^2 f(\mathbf{p})\frac{\mathbf{h}^2}{\|\mathbf{h}\|^2} + \mathbf{r}_2(\mathbf{h}) > 0$. Część **a** została udowodniona. Dowód części **b** jest taki sam, zresztą jest zbędny, bo można już wykazane twierdzenie zastosować do funkcji $-f$. Pozostała część **c.** W tym przypadku stosujemy twierdzenie jednowymiarowe. Z niego wynika, że funkcja przypisująca liczbie t liczbę $f(\mathbf{p} + t\mathbf{v})$ ma w punkcie 0 lokalne minimum właściwe, zatem funkcja f na pewno nie ma w punkcie $\mathbf{p}$ maksimum, nawet niewłaściwego. Analogicznie funkcja przypisująca liczbie t liczbę $f(\mathbf{p} + t\mathbf{w})$ ma w punkcie 0 lokalne maksimum właściwe i w rezultacie funkcja f nie ma w punkcie $\mathbf{p}$ lokalnego minimum: w dowolnie małym otoczeniu punktu $\mathbf{p}$ funkcja f przyjmuje zarówno wartości większe niż $f(\mathbf{p})$ jak i wartości mniejsze niż $f(\mathbf{p})$. Dowód został zakończony. ■

W przypadku **c** mówimy, że funkcja f ma siodło w punkcie $\mathbf{p}$. Jasne jest, że warunki **a,b,c** nie są konieczne. By się o tym przekonać wystarczy rozważyć funkcje $x_1^4 + x_2^4$, $-x_1^4 - x_2^4$ i $x_1^4 - x_2^4$. We wszystkich trzech przypadkach $Df(\mathbf{0}) = \mathbf{0}$ i $D^2 f(\mathbf{0}) = \mathbf{0}$, pierwsza funkcja ma w punkcie $\mathbf{0}$ lokalne minimum właściwe, druga — lokalne maksimum właściwe, trzecia — siodło. Podobnie jak w przypadku funkcji jednej zmiennej można wykazać twierdzenie opisujące lokalne zachowanie się funkcji w otoczeniu punktu krytycznego w zależności od tego, którego rzędu różniczka przyjmuje wartości niezerowe. Niestety łatwo formułuje się takie twierdzenie, ale jego stosowalność jest ograniczona, bo wcale nie jest łatwo stwierdzić jakiego znaku wartości przyjmuje przekształcenie wieloliniowe. Przypadek przekształceń dwuliniowych symetrycznych był omówiony na GAL-u – twierdzenie Sylvestera. Bardzo elementarny dowód tego twierdzenia podaliśmy w części drugiej.

Definicja 5.13 (funkcji klasy C^r)
Funkcja $f\colon G \longrightarrow \mathbb{R}^l$ jest klasy C^r, $r = 1, 2, \ldots$ wtedy i tylko wtedy, gdy dla każdego $\mathbf{x} \in G$ istnieje różniczka $D^r f(\mathbf{x})$ i odwzorowanie $D^r f\colon G \longrightarrow \mathcal{L}(\mathbb{R}^k, \mathbb{R}^k, \ldots \mathbb{R}^k; \mathbb{R}^l)$ jest ciągłe. ■

Jasne jest, że zamiast mówić o ciągłości odwzorowania $D^r f$ można mówić o ciągłości jego pochodnych r–tego rzędu (z ciągłości pochodnych cząstkowych w pewnym punkcie wynika różniczkowalność odwzorowania w tym punkcie). Dodajmy jeszcze, że z różniczkowalności wynika ciągłość, więc pochodne niższych rzędów też są ciągłe. I wreszcie mówiąc funkcja klasy C^0 mamy na myśli funkcję ciągłą.

Na koniec powiedzmy, że w twierdzeniach o odwracania funkcji, o rózniczce funkcji odwrotnej, o funkcjach uwikłanych, o rzędzie oraz w definicji rozmaitości zanurzonej w przestrzeni kartezjańskiej można zastąpić słowa *klasy* C^1 słowami *klasy* C^r, $r \geqslant 1$. W zasadzie jest to uwaga bardzo prosta, jednak dla przykładu udowodnimy odpowiednią wersję twierdzenia o odwracaniu funkcji. Zaczniemy od sformułowania twierdzenia.

Twierdzenie 5.14 (o odwracaniu funkcji klasy C^r)
Niech $f\colon G \longrightarrow \mathbb{R}^l$ będzie funkcją klasy C^r, $r \geqslant 1$ zbioru otwartego $G \subset \mathbb{R}^k$ i niech dla pewnego punktu $\mathbf{p} \in G$ różniczka $Df(\mathbf{p})$ będzie liniowym izomorfizmem przestrzeni $\mathbb{R}^k$ i $\mathbb{R}^l$. Wtedy istnieje zbiór otwarty $U \subset G$, taki że

(i) $\mathbf{p} \in U$ i funkcja f przekształca zbiór U na otwarty podzbiór przestrzeni $\mathbb{R}^l$,

(ii) funkcja f jest różnowartościowa na zbiorze U,

(iii) funkcja $\left(f_{|U}\right)^{-1}$ odwrotna do ograniczenia $f_{|U}$ funkcji f do zbioru U jest klasy C^r.

Dowód. Ponieważ są spełnione założenia twierdzenia o odwracaniu funkcji klasy C^1, więc pozostaje udowodnić, że funkcja odwrotna jest klasy C^r. Wiemy, że zachodzi wzór $Df^{-1}(\mathbf{y}) = \big(Df(f^{-1}(y))\big)^{-1}$. Operacja $A \mapsto A^{-1}$ jest klasy C^∞, co wynika np. z wzorów Cramera na macierz odwrotną — każdy wyraz macierzy A^{-1} jest ilorazem wielomianów k–zmiennych. Jeśli f jest klasy C^2, to Df jest klasy C^1. Oznacza to, że Df^{-1} jest złożeniem (od wewnątrz): funkcji f^{-1}, która jest klasy C^1 (to już wiemy), funkcji Df (która jest klasy C^1 z założenia) i operacji odwracania macierzy (która jest klasy C^∞). Wobec tego Df^{-1} jest funkcją klasy C^2. Prosta indukcja kończy dowód.

Z twierdzenia o odwracaniu funkcji klasy C^r wynika od razu twierdzenie o funkcjach uwikłanych klasy C^r itd.

Kilka zadań

Zadanie 5.2 Niech

$$f(x,y) = \begin{cases} \frac{\ln[\cos(xy)]}{x\sin^2 y}, & \text{jeśli} \quad 0 < |x|, |y| < 1; \\ -\frac{x}{2}, & \text{jeśli} \quad |x|, |y| < 1 \text{ i } xy = 0. \end{cases}$$

Wyjaśnić, czy funkcja f jest różniczkowalna w punkcie $\mathbf{0} = (0,0)$.
Jeśli jest różniczkowalna w punkcie $\mathbf{0} = (0,0)$, znaleźć $Df(0,0)$.
Jaka jest klasa różniczkowalności funkcji f w zbiorze $(-1,1) \times (-1,1)$.

Zadanie 5.3 Udowodnić, że równanie $x\ln w + w\ln y = 0$ wyznacza w otoczeniu punktu $(x_0,y_0) = (1,1)$ zmienną w jako funkcję pozostałych zmiennych: $w = g(x,y)$ i że ta funkcja jest klasy C^∞.
Napisać drugi wielomian Taylora funkcji g w punkcie $(1,1)$.

Zadanie 5.4 (a) Dowieść, że równanie $xe^{w-1} = 2y + \ln w$ wyznacza zmienną w jako funkcję zmiennych x, y w pewnym otoczeniu punktu $(x_0,y_0,w_0) = (2,1,1)$: $w = g(x,y)$ oraz że g jest funkcją klasy C^∞.
(b) Napisać wielomian Taylora stopnia 2 funkcji g dla otoczenia punktu $(2,1)$.

Zadanie 5.5 Zdefiniujmy funkcję $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ wzorami: $f(x,y) = \frac{x^4+y^4}{x^2+y^2}$, gdy $x^2+y^2 > 0$ oraz $f(0,0) = 0$. Znaleźć największe takie r, że $f \in C^r$.

Zadanie 5.6 Niech $f(x,y) = \frac{1}{3}x^3 - \frac{3}{2}x^2 - xy^2 + \frac{1}{2}y^4$. Znaleźć wszystkie punkty krytyczne funkcji f i wyjaśnić, w których z nich funkcja ma lokalne minima, w których lokalne maksima, a w których lokalnego ekstremum w ogóle nie ma.

Zadanie 5.7 Wyznaczyć wszystkie punkty krytyczne funkcji
$$f(x,y) = 3x^4 - 14x^3 - 27x^2 + 30xy - 15y^2 \,.$$
Dla każdego z nich wyjaśnić, czy funkcja ma w tym punkcie lokalne ekstremum, a jeśli ma, to jakie.

Zadanie 5.8 Niech $f(x,y) = (x^3 - x - y)(2x - y - 2)$.
Wyznaczyć wszystkie punkty krytyczne funkcji f.
Wyjaśnić, w których punktach krytycznych funkcja f ma lokalne maksima, w których ma lokalne minima, a w których lokalnego ekstremu nie ma.

Zadanie 5.9 Dowieść, że obraz K odwzorowania
$$\mathbb{R} \ni t \longmapsto (x,y) = (t^3 - 3t^2 + 3t + 5, t^4 - 3t^3 + 3t^2 - t) \in \mathbb{R}^2$$
jest podrozmaitością w $\mathbb{R}^2$ klasy C^1.
Czy jest też rozmaitością klasy C^2?
Znaleźć równania zbioru $T_{\mathbf{p}}K$ dla tych $\mathbf{p} \in K$, dla których $\overrightarrow{[1241,0]} \in T_{\mathbf{p}}K$.

Zadanie 5.10 Niech $f(x,y) = (2+x)^2(3+y)^3(1-x-y)$. Znaleźć wszystkie te punkty $(x_0,y_0) \in \mathbb{R}^2$, w których funkcja f ma lokalne ekstrema (właściwe lub niewłaściwe).

Zadanie 5.11 Niech $f(x,y) = (x^3 - 3x - y)(y + 2)$.
(a) Wyznaczyć wszystkie punkty krytyczne funkcji f.
(b) W których punktach krytycznych funkcja ma lokalne minima, w których lokalne maksima, a w których lokalnego ekstremum nie ma?.

Zadanie 5.12 Niech $f(x.y) = (\cos x - y)(y - 1)$. Wyznaczyć wszystkie punkty krytyczne funkcji f. Dla każdego z nich wyjaśnić, czy funkcja ma w tym punkcie lokalne ekstremum, jeśli ma , to jakie.

Zadanie 5.13 Udowodnić, że równanie $w^3-2xw+y = 0$ wyznacza w otoczeniu punktu $(x_0,y_0,w_0) = =(1,1,1)$ zmienną w jako funkcję klasy C^∞ zmiennych x i y.
Znaleźć drugi wielomian Taylora funkcji $w = w(x,y)$ w punkcie $(1,1)$.

Zadanie 5.14 Dla jakich r zbiór
$$\{(x,y,z) \in \mathbb{R}^3 : \quad z^3 = x^4 + x^2y^4 + 2y^8\}$$
jest podrozmaitością klasy C^r przestrzeni $\mathbb{R}^3$.

Zadanie 5.15 Niech $f(x,y) = g(x+y)+h(x-y)$, gdzie g i h są funkcjami dwukrotnie różniczkowalnymi. Wykazać, że $\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial y^2} = 0$

Zadanie 5.16 Niech $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ będzie funkcją dwukrotnie różniczkowalną o ciągłych pochodnych cząstkowych drugiego rzędu i niech $\frac{\partial^2 f}{\partial x^2}(x,y) - \frac{\partial^2 f}{\partial y^2}(x,y) = 0$ w każdym punkcie $(x,y) \in \mathbb{R}^2$. Niech $\alpha(u,v) = f(u+v, u-v)$. Wykazać, że $\frac{\partial^2 \alpha}{\partial u \partial v}(u,v) = 0$. Wywnioskować stąd, że istnieją takie funkcje jednej zmiennej g oraz h, że równość $f(x,y) = g(x+y) + h(x-y)$ zachodzi dla każdego $(x,y) \in \mathbb{R}^2$.

Zadanie 5.17 Znaleźć punkty zerowania się gradientu funkcji f i wyjaśnić, w których z nich ma ona lokalne minima, w których – lokalne maksima, a w których nie ma lokalnego ekstremum, jeśli

(a) $f(x,y,z) = x^2+y^2+z^2+2x+4y-6z$;
(b) $f(x,y) = x^3+3xy^2-15x-12y$;
(c) $f(x,y,z) = x^3+y^2+z^2+12xy+2z$;
(d) $f(x,y,z) = x+\frac{4y^2}{x}+\frac{z^2}{y}+\frac{2}{z}$;
(e) $f(x,y,z) = xy^2z^3(6-x-2y-3z)$;
(f) $f(x,y) = 3x^8+3y^8+8x^3y^3$;
(g) $f(x,y) = y^2+y^4+3x^4-4x^3-12x^2$;
(h) $f(x,y) = y^2+3x^2y-x^3y$;
(i) $f(x,y) = x^5y^7(13-x-y)$;
(j) $f(x,y) = -x^4+y^4+4x^2y-2y^2$;
(k) $f(x,y) = x^4-y^4-2x^3-2xy^2+x^2+y^2$.

Zadanie 5.18 Niech $f(x,y,z) = \frac{1}{9}\cdot(3(x+y)^3-18x^2-36xy-54y^2-9z^2+2)$. Znaleźć punkty krytyczne f. Wyjaśnić, w których z tych punktów funkcja f ma lokalne minima, w których lokalne maksima, a w których nie ma lokalnego ekstremum.

Zadanie 5.19 Niech
$$f(x,y) = 6y^5+15y^4-50y^3-90y^2+\tfrac{1}{4}\big(-e^{2x}+(y+1)^2(y-2)^2\big)^2,$$
$$g(x,y) = 6y^5+15y^4-50y^3-90y^2+\tfrac{1}{4}\big(-e^{2x}+y^2(y+3)^2\big)^2,$$
$$h(x,y) = 6y^5+15y^4-50y^3-90y^2+\tfrac{1}{4}\big(-e^{2x}+(y+1)^2(y+3)^2\big)^2.$$
Znaleźć punkty zerowania się gradientu tych funkcji i wyjaśnić, w których punktach funkcje mają lokalne minima, w których lokalne maksima, a w których nie ma lokalnego ekstremum. Wykazać, że funkcje f i g nie są ograniczone ani z góry ani z dołu.

Zadanie 5.20 (a) Wykazać, że funkcja $(1+e^y)\cos x - ye^y$ ma nieskończenie wiele maksimów lokalnych, chociaż nie ma żadnego minimum lokalnego.

(b) Znaleźć punkty, w których zeruje się gradient funkcji
$$2(1-e^{2y}+x^2)^3-3(1-e^{2y}+x^2)^2-24x^2e^{2y}$$
i rozstrzygnąć, w których z nich funkcja f ma lokalne minima, w których lokalne maksima, a w których lokalnych ekstremów nie ma.

Zadanie 5.21 Niech $f\in C^\infty(\mathbb{R}^k;\mathbb{R})$ i niech $f(\mathbf{0})=0$. Wykazać, że istnieją takie funkcje $f_1,f_2,\ldots,f_k\cdots\in C^\infty(\mathbb{R}^k;\mathbb{R})$, że równość $f(\mathbf{x}) = \sum_{j=1}^{k} x_jf_j(\mathbf{x})$ zachodzi dla każdego punktu $\mathbf{x}\in\mathbb{R}^k$.

Zadanie 5.22 Niech $f\in C^\infty(\mathbb{R}^k;\mathbb{R})$ i niech $f(\mathbf{0})=0$, $Df(\mathbf{0})=\mathbf{0}$. Wykazać, że istnieją wtedy takie funkcje $f_{i,j}\in C^\infty(\mathbb{R}^k;\mathbb{R})$, $1\le i,j\le k$, że dla każdego punktu $\mathbf{x}\in\mathbb{R}^k$ zachodzi równość $f(\mathbf{x}) = \sum_{i,j=1}^{k} x_ix_jf_{i,j}(\mathbf{x})$.

Zadanie 5.23 Niech $\partial\colon C^\infty(\mathbb{R}^k;\mathbb{R})\longrightarrow\mathbb{R}$ będzie przekształceniem liniowym takim, że dla dowolnych funkcji $f,g\in C^\infty(\mathbb{R}^k;\mathbb{R})$ zachodzi równość $\partial(fg)=\partial(f)g(0)+f(0)\partial(g)$. Wykazać, że istnieją liczby rzeczywiste $a_1,a_2,\ldots,a_k$ takie, że dla dowolnej funkcji $f\in C^\infty(\mathbb{R}^k;\mathbb{R})$ zachodzi równość $\partial(f) = \sum_{j=1}^{k} a_j\frac{\partial f}{\partial x_i}(0)$.

Przekształcenie ∂ nazywane jest różniczkowaniem.

Zadanie 5.24 Dowieść, że zbiór różniczkowań $\partial\colon C^n(\mathbb{R}^k;\mathbb{R}) \longrightarrow \mathbb{R}$ jest nieskończenie wymiarową przestrzenią liniową dla każdej liczby $n \in \{1,2,3,\dots\}$.

Zadanie 5.25 Niech $\varphi(A) = A^{-1}$ dla każdego izomorfizmu A przestrzeni $\mathbb{R}^k$, każdy taki izomorfizm jest utożsamiany z punktem przestrzeni R^{k^2}. Znaleźć $D\varphi(A)H$ dla każdego izomorfizmu A i każdego $H \in \mathbb{R}^{k^2}$. Znaleźć $D^2\varphi(A)(H_1,H_2)$ dla każdego izomorfizmu A i dowolnych $H_1,H_2 \in \mathbb{R}^{k^2}$.

Zadanie 5.26 Załóżmy, że $f\colon \mathbb{R}^k \longrightarrow \mathbb{R}$ jest funkcja klasy C^2, $Df(\mathbf{0}) = \mathbf{0}$ oraz że różniczka $D^2f(\mathbf{0})$ jest nieosobliwa. Dowieść, że istnieje taka liczba $\delta > 0$ i taki dyfeomorfizm $\varphi\colon B(\mathbf{0},\delta) \longrightarrow f(B(\mathbf{0},\delta)) \subseteq \mathbb{R}^k$, że dla każdego $\mathbf{x} \in B(\mathbf{0},\delta)$ zachodzi równość $f(\varphi(\mathbf{x})) = -x_1^2 - x_2^2 - \ldots - x_i^2 + x_{i+1}^2 + x_{i+2}^2 + \ldots + x_k^2$, być może są same minusy lub same plusy.
Należy przypomnieć sobie jak można sprowadzić trójmian kwadratowy do postaci kanonicznej, ogólniej formę kwadratową, następnie skorzystać z twierdzenia o funkcjach uwikłanych.

Zadanie 5.27 Niech $p(\lambda) = \det(A - \lambda I)$, czyli niech p będzie wielomianem charakterystycznym macierzy $A = (a_{ij})$ wymiaru $k \times k$. Obliczyć $p(0)$, $p'(0)$, $\dots$, $p^{(k)}(0)$. Wynik powinien zostać wyrażony za pomocą minorów odpowiedniego wymiaru (stopnia) macierzy A.

Zadanie 5.28 Niech $f\colon \mathbb{R}^k \setminus \{\mathbf{0}\} \longrightarrow \mathbb{R}$ będzie dana wzorem $f(\mathbf{x}) = \frac{1}{\|x\|_2} = \frac{1}{\|x\|}$. Znaleźć $Df(\mathbf{x})\mathbf{h}$ i $D^2f(\mathbf{x})(\mathbf{h},\mathbf{H})$ dla dowolnego $\mathbf{x} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$ i dowolnych $\mathbf{h},\mathbf{H} \in \mathbb{R}^k$. Wynik wyrazić bez użycia współrzędnych $\mathbf{x}$, $\mathbf{h}$, $\mathbf{H}$.

Zadanie 5.29 Załóżmy, że funkcje f,g są dwukrotnie różniczkowalne w punktach $g(\mathbf{p})$ i $\mathbf{p}$ oraz że określone jest złożenie $f \circ g$. Wykazać, że funkcja $f \circ g$ jest różniczkowalna w punkcie $\mathbf{p}$. Znaleźć $D^2(f \circ g)(\mathbf{p})$, czyli dla dowolnych wektorów $\mathbf{h}$ i $\mathbf{H}$ podać wartość $D^2(f \circ g)(\mathbf{p})(\mathbf{h},\mathbf{H})$. Znaleźć $\frac{\partial^2(f\circ g)}{\partial x_i \partial x_j}(\mathbf{p})$ dla dowolnych $i,j \in \{1,\dots,k\}$.

Zadanie 5.30 Udowodnić, że dla każdej funkcji ciągłej $\alpha\colon \mathbb{R}^k \longrightarrow (0,\infty)$ istnieje taka funkcja $\eta\colon \mathbb{R}^k \longrightarrow (0,\infty)$, klasy C^∞, że dla każdego $\mathbf{x} \in \mathbb{R}^k$ spełniona jest nierówność $0 < \eta(\mathbf{x}) < \alpha(\mathbf{x})$.
Można korzystać z twierdzenia o najpaskudniejszej poziomicy w wersji C^∞.

Zadanie 5.31 Dowieść, że dla każdej pary takich funkcji ciągłych $\alpha,\beta\colon \mathbb{R}^k \longrightarrow \mathbb{R}$, że $\beta(\mathbf{x}) < \alpha(\mathbf{x})$ dla wszystkich $\mathbf{x} \in \mathbb{R}^k$, istnieje taka funkcja $\eta\colon \mathbb{R}^k \longrightarrow (0,\infty)$, klasy C^∞, że dla każdego $\mathbf{x} \in \mathbb{R}^k$ spełniona jest nierówność $\beta(\mathbf{x}) < \eta(\mathbf{x}) < \alpha(\mathbf{x})$.

Zadanie 5.32 Dowieść, że dla każdej pary takich funkcji ciągłych $\alpha,\beta\colon \mathbb{R}^k \longrightarrow \mathbb{R}$, że $\beta(\mathbf{x}) < \alpha(\mathbf{x})$ dla wszystkich $\mathbf{x} \in \mathbb{R}^k$, istnieje dyfeomorfizm klasy C^∞
$$F\colon \{(\mathbf{x},y)\colon \quad \beta(\mathbf{x}) < y < \alpha(\mathbf{x})\} \xrightarrow{na} \mathbb{R}^{k+1}.$$
Jedno z możliwych rozwiązań polega na skorzystaniu z własności rozwiązań równania różniczkowego $y'(t) = \eta(x,y(t))$, gdzie η jest funkcją z poprzedniego zadania.

Zadanie 5.33 Czy istnieje funkcja różniczkowalna $F\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dwukrotnie różniczkowalna wyłącznie w punkcie $(0,0)$?