WYKŁAD 8

Pochodna $f:\mathbb{R}^k \supset \mathcal{O} \to \mathbb{R}^l$ rzędu n, $f^{(n)}(a) \in L(\underbrace{\mathbb{R}^k, \dots, \mathbb{R}^k}_{n\text{-razy}}; \mathbb{R}^l)$ - odwzorowanie n-liniowe na $\mathbb{R}^k \times \dots \times \mathbb{R}^k$ o wartościach w $\mathbb{R}^l$. Przykładowo dla $k=2$, $l=1$, $n=2$ $f^{(2)}(a)$ jest zadana macierzą 2-gich pochodnych cząstkowych

$$f^{(2)}(a)(h_1, h_2) = h_1^T \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a), & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a), & \frac{\partial^2 f}{\partial x_2^2}(a) \end{bmatrix} h_2$$

gdzie $h_1, h_2 \in \mathbb{R}^2$ $h_1 = \begin{bmatrix} h_{1,1} \\ h_{1,2} \end{bmatrix}$ $h_2 = \begin{bmatrix} h_{2,1} \\ h_{2,2} \end{bmatrix}$. Zatem $f^{(2)}(a)(h_1, h_2) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_{1,i} h_{2,j}$

Zauważmy, że $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ implikuje

$$f^{(2)}(a)(h_1, h_2) = f^{(2)}(a)(h_2, h_1) \quad (*)$$

Ogólniej $f^{(n)}(a)(h_1, \dots, h_n) = \underbrace{\sum_{i_1, \dots, i_n}}_{\text{suma po wszystkich wskaźnikach } i_1, \dots, i_n} \frac{\partial^n f(a)}{\partial x_{i_1} \dots \partial x_{i_n}} h_{1,i_1} \dots h_{n,i_n}$

Tutaj analogiem $(*)$ jest symetria $f^{(n)}(a)$ ze względu na przestawianie $h_1, \dots, h_n$ - permutacje $1, \dots, n$. Niech $\sigma \in S(n)$ - grupa permutacji zbioru indeksów $1, \dots, n$. $\sigma = \begin{pmatrix} 1, & \dots, & n \\ \sigma(1), & \dots, & \sigma(n) \end{pmatrix}$ Wówczas $f^{(n)}(a)(h_1, \dots, h_n) = f^{(n)}(a)(h_{\sigma(1)}, \dots, h_{\sigma(n)})$.

<u>Wielowskaźnik</u> $i = (i_1, \dots, i_k)$ to ciąg o współrzędnych całkowitych nieujemnych. Długością wielowskaźnika i nazywamy wielkość $|i| = i_1 + \dots + i_k$. Symbol $\frac{\partial^{|i|} f(a)}{\partial x^i}$ oznaczać będzie pochodną cząstkową $\frac{\partial^{i_1}}{\partial x_1^{i_1}} \dots \frac{\partial^{i_k}}{\partial x_k^{i_k}} f(a)$. Ponadto dla wektora $h = \begin{bmatrix} h_1 \\ \vdots \\ h_k \end{bmatrix} \in \mathbb{R}^k$ piszemy $h^i = h_1^{i_1} \dots h_k^{i_k}$

Przykład $f(x_1, x_2, x_3)$ - funkcja 3 zmiennych

$i = (1,2,3)$: $\frac{\partial^{|i|}}{\partial x^i} f(a) = \frac{\partial^6 f(a)}{\partial x_1 \partial x_2^2 \partial x_3^3}$

$j = (0,0,1)$ $\frac{\partial^{|j|}}{\partial x^j} f(a) = \frac{\partial f(a)}{\partial x_3}$

Uwaga $f^{(n)}(a)(\underbrace{h, \dots, h}_{n\text{-razy}}) = \sum_{\substack{i\text{ - wielowskaźnik} \\ \text{długości } n \text{ czyli} \\ i=(i_1,\dots,i_k) \,\&\, i_1+\dots+i_k=n}} \frac{\partial^{|i|} f(a)}{\partial x^i} h^i$

Stwierdzenie 1

Przypuśćmy, że $f: \mathbb{R}^k \supset O \to \mathbb{R}$ ma pochodne cząstkowe na O rzędu 2 które są ciągłe (mówimy wtedy, że f jest klasy C^2 na O i piszemy $f \in C^2(O, \mathbb{R})$) Jeśli $a, h \in O$ są takie, że odcinek $[a,h] = \{a + t(h-a) \quad t \in [0,1]\}$ jest zawarty w O to istnieje $\theta \in (0,1)$ (zależny od a i h) że $f(a+h) = f(a) + f'(a)h + \frac{1}{2} f''(a + \theta h)(h,h)$.

Uwaga 1 w dalszej części wykładu piszemy $f''(b)(h, \dots, h) = f''(b) h^2$ Ogólniej $f^{(n)}(b)(h, \dots, h) \overset{\text{ozn}}{=} f^{(n)}(b) h^n$.

Uwaga 2 Ogólniejsza wersja stw 1 wygląda następująco: jeśli $f \in C^n(O, \mathbb{R})$ to

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2!} f''(a) h^2 + \dots + \frac{1}{n!} f^{(n)}(a + \theta h) h^n.$$

Ma to więc postać wzoru Taylora z resztą Lagrange'a dla funkcji jednej zmiennej.

Dowód (Stwierdzenia 1 (v2 - dowód analogiczny)) (3)

Rozważmy funkcję $g:[0,1] \to \mathbb{R}$ daną wzorem $g(t) = f(a+th) = f(a_1+th_1, \dots, a_k+th_k)$

Obliczmy $g'(t)$. Z reguły łańcucha (pochodna funkcji złożonej) dostajemy $g'(t) = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(a+th) \cdot h_i = f'(a+th) \circ h$

Podobnie $g''(t) = \sum_{i_1,i_2=1}^{k} \frac{\partial^2 f(a+th)}{\partial x_{i_1} \partial x_{i_2}} h_{i_1} h_{i_2} = f''(a)(h,h)$

Zatem ze wzoru Taylora z resztą w postaci Lagrange'a wynika, że istnieje $\theta \in (0,1)$ t. że $g(1) = g(0) + g'(0) + \frac{1}{2} g''(\theta)$

co daje $f(a+h) = f(a) + f'(a)h + \frac{1}{2} f''(a+\theta h) h^2$.

Uwaga: Powyższy dowód nie przechodzi dla $f(x) = \begin{bmatrix} f_1(x_1,\dots,x_k) \\ \vdots \\ f_\ell(x_1,\dots,x_k) \end{bmatrix}$ gdyż dla każdej f_i mielibyśmy inne $\theta_i \in (0,1)$.

Prawdziwe jest następujące twierdzenie

Twierdzenie (Wzór Taylora z resztą Peano)

Załóżmy, że $f: \mathbb{R}^k \supset \theta \to \mathbb{R}^\ell$ jest $n-1$-krotnie różniczkowalna na θ, $K(a,r) \subset \theta$ oraz $f^{(n)}(a)$ istnieje.

Wówczas dla $\|h\| < r$ mamy

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2!} f''(a) h^2 + \dots + \frac{1}{n!} f^{(n)}(a) h^n + R(a,h)$$

gdzie $R(a,h) \in \mathbb{R}^\ell$ spełnia $\lim_{\|h\| \to 0} \frac{\|R(a,h)\|}{\|h\|^k} = 0$

Dowód Skorzystamy z wykazanej wcześniej (4)
nierówności $\|f(x+h)-f(x)\| \overset{(**)}{\leq} \|h\| \cdot \sup_{\theta\in[0,1]} \|f'(x+\theta h)\|$
w kontekście odwzorowania (za f wstawio-my R)
$K(a,r)\ni h \longmapsto R(a,h)\in\mathbb{R}^l$. W dalszej części piszemy

$$R(a,h) \overset{\text{ozn}}{=} R(h) = f(a+h) - f(a) - f'(a)h - \dots - \frac{1}{n!} f^{(n)}(a) h^n$$

gdzie $f^{(j)}(a)\, h^j = f^{(j)}(a)\underbrace{(h,\dots,h)}_{j\text{- razy}}$ $h\in\mathbb{R}^k$

$f^{(j)}(a) \in L(\underbrace{\mathbb{R}^k,\dots,\mathbb{R}^k}_{j\text{- razy}}, \mathbb{R}^l)$.

Obliczając pochodne $R(h)$ względem h

Skorzystamy z następujących wzorów:
rozważmy np. odwzorowanie $g: h \longmapsto f''(a)(h,h)$
Jak obliczyć pochodną? Musimy znaleźć część liniową przyrostu:

$$f''(a)(h+\tilde{h}, h+\tilde{h}) = f''(a)(h,h) + f''(a)(h,\tilde{h}) + f''(a)(\tilde{h},h) + f''(a)(\tilde{h},\tilde{h}) \underset{\substack{\text{symetria}\\ \text{2- poch}}}{=\!=} f''(a)(h,h) + \underbrace{2f''(a)(h,\tilde{h})}_{\text{liniowe w } \tilde{h}} + \underbrace{f''(a)(\tilde{h},\tilde{h})}_{\text{kwadratowe}}$$

Wniosek $g'(h)\tilde{h} = 2f''(a)(h,\tilde{h})$ co symbolicznie zapiszemy $g'(h) = 2f''(a)h$.

Ogólniej biorąc $\tilde{g}(h) = f^{(j)}(a)\underbrace{(h,\dots,h)}_{j\text{- razy}}$ dostaniemy

$$\tilde{g}'(h)(\tilde{h}) = j\, f^{(j)}(a)(\overbrace{h,\dots,h}^{j-1 \text{ razy}}, \tilde{h})$$ i symbolicznie

$$\tilde{g}'(h) = j \cdot f^{(j)}(a)\, h^{j-1}$$

Stosując te reguły / tę notację dostajemy

$$R'(h) = f'(a+h) - f'(a) - f''(a)h - \dots - \frac{1}{(n-1)!} f^{(n)}(a) h^{n-1}$$

$\vdots$

$$R^{n-1}(h) = f^{(n-1)}(a+h) - f^{(n-1)}(a) - f^{(n)}(a) h$$

Mamy więc $R(0) = R'(0) = \dots = R^{n-1}(0) = 0$

Korzystając z (**) (p. 4 wiersz 2) z funkcją R mamy

$$\|R(h)\| = \|R(h) - R(0)\| \leq \|h\| \sup_{\theta\in(0,1)} \|R'(\theta h)\| \leq$$

$$\leq \|h\|^2 \sup_{\theta\in(0,1)} \|R''(\theta h)\| \leq \dots \leq \|h\|^{n-1} \sup_{\theta\in(0,1)} \|R^{(n-1)}(\theta h)\|$$

$$= \|h\|^{n-1} \sup_{\theta\in(0,1)} \|f^{(n-1)}(a+\theta h) - f^{(n-1)}(a) - f^{(n)}(a)\theta h\| =$$

$$= \|h\|^{n-1} \sup_{\theta\in(0,1)} \frac{\|f^{(n-1)}(a+\theta h) - f^{(n-1)}(a) - f^{(n)}(a)\theta h\|}{|\theta|\,\|h\|} \cdot \theta\|h\| \leq$$

$$\|h\|^n \sup_{\theta\in(0,1)} \frac{\|f^{(n-1)}(a+\theta h) - f^{(n-1)}(a) - f^{(n)}(a)\theta h\|}{\theta\|h\|} \quad (***)$$

Skoro istnieje n-ta pochodna $f^{(n)}(a)$ to

$\forall \varepsilon > 0 \ \exists \delta > 0 \ \ \|\tilde{h}\| < \delta$ to $\dfrac{\|f^{(n-1)}(a+\tilde{h}) - f^{(n-1)}(a) - f^{(n)}(a)\tilde{h}\|}{\|\tilde{h}\|} < \varepsilon$

W szczególności dla $\|h\| \leq \delta$ wyrażenie (***) $< \varepsilon$.

Wniosek $\dfrac{\|R(h)\|}{\|h\|^n} \xrightarrow[\|h\|\to 0]{} 0$ $\square$

Ekstrema lokalne funkcji $f: \mathbb{R}^k \supset O \to \mathbb{R}$ ⑥
wielu zmiennych $f(x_1, \dots, x_k)$.

Definicja: Mówimy, że $a \in O$ jest punktem krytycznym funkcji jeśli $f'(a) = 0$ (w szczególności będziemy zakładać że $f'(a)$ istnieje oraz $f'(a) = \left[\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_k}(a)\right] = 0$)

Mówimy, że w $a \in O$ jest minimum (maksimum) lokalne funkcji f jeśli istnieje r takie, że dla $h \in K(0,r)$ zachodzi $f(a+h) \geq f(a)$ $(f(a+h) \leq f(a))$.

Lemat: Przypuśćmy, że $f: O \to \mathbb{R}$ jest różniczk w $a \in O$. Jeśli w a jest lokalne minimum to a jest punktem krytycznym funkcji f.

Dowód Rozważmy funkcję $g(t) = f(a + th)$ określoną dla $t \in \mathbb{R}$ z pewnego otoczenia zera. Z reguły Łańcucha wynika że g jest różniczkowalna w $t = 0$. Ponadto g ma w $t = 0$ minimum. Z lematu Fermata wnioskujemy, że $g'(0) = 0$. Z reguły Łańcucha $g'(0) = f'(a)h = 0$ dla wszystkich h.

Zatem $f'(a) = \underbrace{[0, \dots, 0]}_{k\text{-razy}}$, czyli a jest pktem krytycznym f.

Warunek dostateczny ekstremów lokalnych.

Jeśli f jest funkcją jednej zmiennej, $a \in \mathbb{R}$ jest punktem krytycznym f i $f''(a) > 0$ to f ma w p-cie a minimum lokalne.

Ogólniej 2-ga pochodna $f''(a)$ jest formą 2-liniową symetryczną na $\mathbb{R}^k \times \mathbb{R}^k$ o wartościach w $\mathbb{R}$. Przypomnijmy wzór Taylora

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

$$\left\{ \begin{array}{l} a - \text{pt. kryt.} \\ f'(a) = 0 \end{array} \right\} \overset{(****)}{=} f(a) + \frac{1}{2} f''(a+\theta h)(h,h).$$

Jeśli $f''(x)(h,h) \geq 0$ dla x z małego otoczenia a to dla dostatecznie małych h $f(a+h) \geq f(a)$ i w $a \in O$ f przyjmuje minimum lokalne

Mamy więc pierwszy warunek

Twierdzenie

Jeśli $f: \mathbb{R}^k \supset O \to \mathbb{R}$, $a \in O$ - pkt krytyczny, $f \in C^2(O, \mathbb{R})$ oraz dla x z pewnego otoczenia a zachodzi $f''(x)(h,h) \geq 0$ to w f przyjmuje w a minimum lokalne.

Dowód: $(****)$ - spełniony jest przy zał. tw. stąd $f(a+h) - f(a) \geq 0$ dla dostatecznie małego h.

Twierdzenie powyższe posiada też wersję w (8)
które zakłada się, że f ma pierwszą pochodną na O, a jest p-tem krytycznym oraz $f''(a)(h,h) > 0$ dla wszystkich $h \in \mathbb{R}^k \setminus \{0\}$
Nie zakłada się istnienia drugiej pochodnej wokół a.

Definicja: Mówimy, że druga pochodna $f''(a)$ jest ściśle dodatnio określona jeśli dla $h \in \mathbb{R}^k \setminus \{0\}$ zachodzi $f''(a)(h,h) > 0$. (Piszemy $f''(a) > 0$)

Twierdzenie
Niech a będzie punktem krytycznym $f: O \to \mathbb{R}$ i przypuśćmy, że $f''(a)$ istnieje oraz $f''(a) > 0$.
Wówczas f przyjmuje w a minimum lokalne.

Dowód: Oznaczmy macierz drugiej poch. symbolem A.

$$f''(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]_{i,j=1,\dots,k}$$

$$f''(a)(h,h) = [h_1, \dots, h_k]\, A \begin{bmatrix} h_1 \\ \vdots \\ h_k \end{bmatrix} \quad \text{gdzie } h = \begin{bmatrix} h_1 \\ \vdots \\ h_k \end{bmatrix}$$

Zauważmy, że $\inf\limits_{\|h\|=1} h^T A h > 0$. Rzeczywiście funkcja $h \mapsto h^T A h$ przyjmuje kres na sferze $S(0,1) = \{h : \|h\| = 1\}$. Czyli $\inf h^T A h = v^T A v \overset{\text{ozn}}{=} c > 0$ dla pewnego v t. że $\|v\| = 1$.

Wzór Taylora: $$0 = \lim_{\|h\| \to 0} \frac{\| f(a+h) - f(a) - \frac{1}{2} h^T A h \|}{\|h\|^2} = 0 = \lim_{h \to 0} \frac{\|R(h)\|}{\|h\|^2}$$

Istnieje więc $\delta > 0$ takie, że dla $\|h\| < \delta$

mamy $\frac{|R(h)|}{\|h\|^2} < \frac{c}{4}$ gdzie $c = \inf_{\|\xi\|=1} \xi^T A \xi$

Podsumowując

$$f(a+h) = f(a) + \frac{1}{2} h^T A h + R(h).$$

$$= \frac{1}{2} \underbrace{\frac{h^T A h}{\|h\|^2}}_{\geq c} \|h\|^2 + R(h)$$

$$\geq \frac{c}{2}\|h\|^2 - \frac{c}{4}\|h\|^2 \geq \frac{c}{4}\|h\|^2 > 0$$

dla $\|h\| < \delta$

Wniosek w a jest minimum lokalne

Uwaga
Podobnie pokazujemy, że jeśli $f''(a)(h,h) < 0$ dla $h \in \mathbb{R}^n \setminus \{0\}$ to punkt krytyczny $a \in O$ jest lokalnym maksimum.

Jak sprawdzać, czy $f''(a)(h,h) > 0$?

Patrz wykład z algebry