Interpretacja geometryczna gradientu funkcji różniczkowalnej
Niech $D \subset \mathbb{R}^2$, $(x^0, y^0) \in \operatorname{Int} D$, $f : D \to \mathbb{R}$. Jeżeli f jest różniczkowalna w punkcie (x^0, y^0), to równanie płaszczyzny stycznej do wykresu funkcji f w punkcie $(x^0, y^0, f(x^0, y^0))$ wyraża się wzorem

$$z - f(x^0, y^0) = f'_x(x^0, y^0)(x - x^0) + f'_y(x^0, y^0)(y - y^0).$$

Ogólnie dla $k \in \mathbb{N}$, $k \geq 2$, $D \subset \mathbb{R}^k$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}$. Jeżeli f jest różniczkowalna w punkcie x^0, to równanie hiperpłaszczyzny stycznej do wykresu funkcji f w punkcie $(x^0, f(x^0))$ wyraża się wzorem

$$y - f(x^0) = \nabla f(x^0)(x - x^0).$$

Uwaga 1. i) Gradient funkcji różniczkowalnej f w punkcie $x^0 \in \operatorname{Int} D$ jest kierunkiem najszybszego wzrostu funkcji f w tym punkcie.
ii) Gradient funkcji różniczkowalnej f w punkcie $x^0 \in \operatorname{Int} D$ jest wektorem prostopadłym do poziomicy funkcji f zawierającej punkt x^0 tzn. $\nabla f(x^0) \perp \operatorname{lev}_{=f(x^0)} f$.

Zastosowania różniczkowalności do obliczeń przybliżonych

Stwierdzenie 1. *Niech $f : K((x^0, y^0), r) \to \mathbb{R}$ dla pewnego $r > 0$. Jeżeli f jest różniczkowalna w punkcie (x^0, y^0), to dla dowolnego $(h_1, h_2) \in K((0, 0), r)$ zachodzi*

$$f(x^0 + h_1, y^0 + h_2) - f(x^0, y^0) = f'_x(x^0, y^0)h_1 + f'_y(x^0, y^0)h_2 + \varepsilon_1(h_1, h_2)h_1 + \varepsilon_2(h_1, h_2)h_2$$

gdzie $\varepsilon_1, \varepsilon_2 : K((0, 0), r) \to \mathbb{R}$ są pewnymi funkcjami takimi, że $\lim\limits_{(h_1,h_2)\to(0,0)} \varepsilon_1(h_1, h_2) = 0$ oraz $\lim\limits_{(h_1,h_2)\to(0,0)} \varepsilon_1(h_1, h_2) = 0$.

Wniosek 1. *Niech $f : K((x^0, y^0), r) \to \mathbb{R}$ dla pewnego $r > 0$. Jeżeli f jest różniczkowalna w punkcie (x^0, y^0), to dla (h_1, h_2) „bliskich” $(0, 0)$ mamy, że*

$$f(x^0 + h_1, y^0 + h_2) \approx f(x^0, y^0) + f'_x(x^0, y^0)h_1 + f'_y(x^0, y^0)h_2.$$

Analogicznie rezultaty do powyższego stwierdzenia i wniosku obowiązują dla funkcji k zmiennych, które zapiszemy w trochę innej postaci.

Stwierdzenie 2. *Niech $f : K_{\mathbb{R}^k}(x^0, r) \to \mathbb{R}$ dla pewnego $r > 0$. Jeżeli f jest różniczkowalna w punkcie x^0, to dla „dostatecznie małego r”*

$$\forall\, x \in K(x^0, r) \quad f(x) - f(x^0) \approx \nabla f(x^0)(x - x^0) =: df(x^0)(x - x^0),$$

przy czym błąd przybliżenia $\delta(x - x^0) := f(x) - f(x^0) - df(x^0)(x - x^0)$ zbiega do 0 szybciej niż $\|x - x^0\|$ tzn. $\lim\limits_{x\to x^0} \frac{\delta(x-x^0)}{\|x-x^0\|} = 0$. Wyrażenie $df(x^0)$ nazywamy ***różniczką****.*

Reguły różniczkowania

Twierdzenie 1. *Niech $D \subset \mathbb{R}^k$, $x^0 \in \operatorname{Int} D$, $f, g : D \to \mathbb{R}$, $c \in \mathbb{R}$. Wówczas*

(1) Jeżeli funkcje f, g są różniczkowalne w x^0, to funkcje $f+g$, $c \cdot f$, $f \cdot g$ są różniczkowalne w x^0.

(2) Jeżeli funkcje f, g są różniczkowalne w x^0 oraz $g(x) \neq 0$, $x \in D$, to funkcja $\frac{f}{g}$ jest różniczkowalna w x^0.

Twierdzenie 2. ***(warunek dostateczny różniczkowalności (mocnej, w sensie Frécheta))*** *Niech $D \subset \mathbb{R}^k$, $f : D \to \mathbb{R}$, $x^0 \in \operatorname{Int} D$. Jeżeli f'_{x_i}, $i = 1, \ldots, k$ istnieją w pewnym otoczeniu punktu x^0 oraz są ciągłe w punkcie x^0, to funkcja f jest różniczkowalna (w sposób mocny, w sensie Freécheta) w punkcie x^0.*

Dowód. Dowód przeprowadzony na wykładzie. □

Rozważmy przykład

Przykład 1. Rozważmy funkcję $f : \mathbb{R}^2 \to \mathbb{R}$ określoną wzorem

$$f(x,y) = \begin{cases} xy \sin(\frac{1}{x^2+y^2}) & \text{dla } (x,y) \neq (0,0) \\ 0 & \text{dla } (x,y) = (0,0). \end{cases}$$

Wówczas łatwo sprawdzić, że

$$f'_x(x,y) = \begin{cases} y \sin(\frac{1}{x^2+y^2}) - \frac{2x^2y}{(x^2+y^2)^2} \cos(\frac{1}{x^2+y^2}) & \text{dla } (x,y) \neq (0,0) \\ 0 & \text{dla } (x,y) = (0,0) \end{cases}$$

oraz

$$f'_y(x,y) = \begin{cases} x \sin(\frac{1}{x^2+y^2}) - \frac{2xy^2}{(x^2+y^2)^2} \cos(\frac{1}{x^2+y^2}) & \text{dla } (x,y) \neq (0,0) \\ 0 & \text{dla } (x,y) = (0,0). \end{cases}$$

Rozważmy $(x_n, y_n)_{n \in \mathbb{N}}$, gdzie $(x_n, y_n) = (\frac{1}{\sqrt{4n\pi}}, \frac{1}{\sqrt{4n\pi}}) \in D_{f'_x}$, $n \in \mathbb{N}$. Wówczas

$$\lim_{n \to +\infty} (x_n, y_n) = (0,0),$$

$$\lim_{n \to +\infty} f'_x(x_n, y_n) = \lim_{n \to +\infty} \frac{1}{\sqrt{4n\pi}} \sin(2n\pi) - \frac{\frac{1}{n\sqrt{n}\sqrt{\pi^3}}}{\frac{1}{n^2\pi^2}} \cos(2n\pi) = -\infty$$

Zatem f'_x nie jest ciągła w $(0,0)$. Analogicznie pokazuje się, że f'_y nie jest ciągła w $(0,0)$. Z drugiej strony

$$\forall (h_1,h_2) \neq (0,0) \quad 0 \leq |\frac{f(h_1,h_2) - f(0,0) - f'_x(0,0)h_1 - f'_y(0,0)h_2}{\|(h_1,h_2)\|}| = |\frac{h_1 h_2 \sin(\frac{1}{h_1^2+h_2^2})}{\sqrt{h_1^2+h_2^2}}| \leq \sqrt{h_1^2+h_2^2},$$

zatem f jest różniczkowalna w $(0,0)$. Ten przykład pokazuje, że warunek dostateczny różniczkowalności funkcji w $\mathbb{R}^k$, nie jest warunkiem koniecznym.

Różniczkowalność funkcji wektorowych

Uwaga 2. Z algebry wiadomo, że odwzorowanie liniowe $A : \mathbb{R}^k \to \mathbb{R}^m$ jest jednoznacznie wyznaczone przez macierz $[a_{ij}]_{i\le m, j\le k}$ wymiaru (kształtu) $m \times k$, przy czym dla $x = [x_1, \ldots, x_k]^T \in \mathbb{R}^k$ mamy, że

$$A(x) = [a_{ij}] \bullet [x_1, \ldots, x_k]^T = \left[\sum_{j=1}^{k} a_{1j}x_j, \ldots, \sum_{j=1}^{k} a_{mj}x_j \right] \in \mathbb{R}^m,$$

gdzie $\bullet$ oznacza iloczyn macierzy. Łatwo wykazać, że A jest odzworowaniem liniowym i ciągłym. Jeżeli rozważamy zbiór wszystkich odwzorowań liniowych i ciągłych przekształcających $\mathbb{R}^k$ w $\mathbb{R}^m$, to można wprowadzić w tym zbiorze strukturę przestrzeni liniowej i oznaczamy ją $\mathcal{L}(\mathbb{R}^k, \mathbb{R}^m)$.

Definicja 1. *Niech $k, m \in \mathbb{N}$, $m \ge 2$. Niech $D \subset \mathbb{R}^k$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}^m$.* ***Pochodną (mocną, w sensie Frécheta) funkcji*** f ***w punkcie*** x^0 *nazywamy odzworowanie $A \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^m)$ spełniające warunek*

$$\lim_{\|h\|_{\mathbb{R}^k} \to 0} \tfrac{1}{\|h\|_{\mathbb{R}^k}} (f(x^0 + h) - f(x^0) - A(h)) = 0_{\mathbb{R}^m}$$

Wówczas oznaczamy $A = f'(x^0)$ i mówimy, że funkcja f jest ***różniczkowalna (w sposób mocny, w sensie Frécheta) w punkcie*** x^0.

Następujące twierdzenie jest analogiem twierdzeń dla funkcji wielu zmiennych, w związku z tym dowód pomijamy.

Twierdzenie 3. *Niech $k, m \in \mathbb{N}$, $m \ge 2$. Niech $D \subset \mathbb{R}^k$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}^m$. Wtedy*

(1) Odwzorowanie $A \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^m)$ jest pochodną mocną funkcji f w punkcie x^0 wtedy i tylko wtedy, gdy dla każdego takiego $h \in \mathbb{R}^k$, że $x^0 + h \in D$ mamy

$$f(x^0 + h) - f(x^0) = A(h) + r(h),$$

gdzie $r : \{h \in \mathbb{R}^k : x^0 + h \in D\} \to \mathbb{R}^m$ jest pewną funkcją spełniającą warunek $\lim\limits_{h \to 0_{\mathbb{R}^k}} \frac{r(h)}{\|h\|} = 0_{\mathbb{R}^m}$

(2) Jeżeli f jest różniczkowalna w punkcie x^0, to f jest ciągła w punkcie x^0.

Uwaga 3. Niech $D \subset \mathbb{R}^k$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}^m$. Załóżmy, że f jest różniczkowalna w punkcie x^0, co wprost z definicji oznacza, że $f'(x^0) \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^m)$. Z uwagi 2 oznacza to, że istnieje macierz kształu $m \times k$- $[a_{ij}]$, którą można utożsamiać z $f'(x^0)$. Tę macierz dla $f : \mathbb{R}^k \to \mathbb{R}^m$ wyznaczamy przy pomocy pochodnych cząstkowych współrzędnych wektora $f = (f_1, \ldots, f_m)$.

Definicja 2. *Niech $D \subset \mathbb{R}^k$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$. Załóżmy, że f jest różniczkowalna w punkcie x^0. Wówczas macierz*

$$[(f_i)'_{x_j}(x^0)]_{i \le m, j \le k}$$

nazywamy ***macierzą Jacobiego funkcji*** f ***w punkcie*** x^0 *i oznaczamy $f'(x^0)$. Jeżeli dodatkowo $k = m$, to wyznaczynik macierzy Jacobiego funkcji f w punkcie x^0 nazywamy* ***jakobianem funkcji*** f ***w punkcie*** x^0 *i oznaczamy $Jf(x^0)$.*

Definicja 3. *Niech $k, m \in \mathbb{N}$. Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $f : U \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$. Mówimy, że* ***funkcja f jest klasy C^1 na*** *U, co zapisujemy $f \in C^1(U)$, gdy funkcja f jest różniczkowalna na U oraz $(f_i)'_{x_j}$ są ciągłe na U dla $i = 1, \ldots, m$, $j = 1, \ldots, k$.*

Twierdzenie 4. ***(pochodna superpozycji)*** *Niech $k, m, l \in \mathbb{N}$. Niech $D \subset \mathbb{R}^k$, $E \subset \mathbb{R}^m$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}^m$, $f[D] \subset E$, $f(x^0) \in \operatorname{Int} E$, $g : E \to \mathbb{R}^l$. Jeżeli f jest różniczkowalna w punkcie x^0, a g różniczkowalna w punkcie $f(x^0)$, to funkcja założona $g \circ f : D \to \mathbb{R}^l$ jest różniczkowalna w x^0 oraz zachodzi wzór*

$$(g \circ f)'(x^0) = g'(f(x^0)) \circ f'(x^0). \tag{1}$$

Uwaga 4. Symbol $\circ$ oznacz superpozycję (złożenie odwzorowań). W szczególności $g'(f(x^0)) \circ f'(x^0)$ we wzorze (1) oznacza złożenie odwzorowania liniowego i ciągłego z przestrzeni $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^l)$ z odwzorowaniem z $\mathcal{L}(\mathbb{R}^k, \mathbb{R}^m)$.

Dowód. **Dowód nie obowiązuje na egzaminie.**
Z istnienia $f'(x^0)$ oraz twierdzenia 3 wynika, że istnieje takie $t > 0$, że dla $h \in \mathbb{R}^k$ takich, że $x^0 + h \in K(x^0, t) \subset D$ zachodzi

$$f(x^0 + h) - f(x^0) = f'(x^0)(h) + r_1(h), \quad \lim_{h \to 0_{\mathbb{R}^k}} \tfrac{r_1(h)}{\|h\|_{\mathbb{R}^k}} = 0_{\mathbb{R}^m}, \tag{2}$$

gdzie r_1 jest pewną funkcją określoną na $K(0_{\mathbb{R}^k}, t)$ o wartościach w $\mathbb{R}^m$. Z istnienia $g'(f(x^0))$ oraz twierdzenia 3 wynika, że istnieje takie $s > 0$, że dla $d \in \mathbb{R}^m$ takich, że $f(x^0) + d \in K(f(x^0), s) \subset E$ zachodzi

$$g(f(x^0) + d) - g(f(x^0)) = g'(f(x^0))(d) + r_2(d), \quad \lim_{d \to 0_{\mathbb{R}^m}} \tfrac{r_2(d)}{\|d\|_{\mathbb{R}^m}} = 0_{\mathbb{R}^l}, \tag{3}$$

gdzie r_2 jest pewną funkcją określoną na $K(0_{\mathbb{R}^m}, s)$ o wartościach w $\mathbb{R}^l$.
Połóżmy

$$d(h) = f'(x^0)(h) + r_1(h), \text{ dla } h \in K(0_{\mathbb{R}^k}, t). \tag{4}$$

Zauważmy, że

$$\lim_{\|h\|_{\mathbb{R}^k} \to 0} \big(f'(x^0)(h) + r_1(h)\big) = \lim_{\|h\|_{\mathbb{R}^k} \to 0} \big(f'(x^0)(h) + \tfrac{r_1(h)}{\|h\|_{\mathbb{R}^k}} \|h\|_{\mathbb{R}^k}\big) = 0_{\mathbb{R}^m}.$$

Wobec faktu, że $d(0_{\mathbb{R}^k}) = 0_{\mathbb{R}^m}$ oraz powyższego dostajemy, że funkcja d jest ciągła w $0_{\mathbb{R}^k}$. Zatem istnieje taka $0 < \delta < t$, że $f(x^0) + d(h), f(x^0 + h) \in K(f(x^0), s)$ dla wszystkich takich $h \in \mathbb{R}^k$, że $\|h\|_{\mathbb{R}^k} < \delta$. Zatem dla $h \in \mathbb{R}^k$ takich, że $\|h\|_{\mathbb{R}^k} < \delta$ ze wzorów (2), (3) mamy, że

$$\begin{aligned}
&(g \circ f)(x^0 + h) - (g \circ f)(x^0) = g(f(x^0 + h)) - g(f(x^0)) = \\
&g(f(x^0) + f'(x^0)(h) + r_1(h)) - g(f(x^0)) = \\
&g'(f(x^0))(f'(x^0)(h) + r_1(h)) + r_2(f'(x^0)(h) + r_1(h)) = \\
&g'(f(x^0))(f'(x^0)(h)) + g'(f(x^0))(r_1(h)) + r_2(d(h)) = \\
&g'(f(x^0)) \circ f'(x^0)(h) + r_3(h),
\end{aligned}$$

gdzie

$$r_3(h) = g'(f(x^0))(r_1(h)) + r_2(d(h)), \text{ dla } h \in K(0_{\mathbb{R}^k}, \delta).$$

W myśl twierdzenia 3 aby otrzymać tezę wystarczy pokazać, że

$$\lim_{h \to 0_{\mathbb{R}^k}} \frac{r_3(h)}{\|h\|_{\mathbb{R}^k}} = 0_{\mathbb{R}^l}. \tag{5}$$

Istotnie.

$$\forall\, 0 < \|h\|_{\mathbb{R}^k} < \delta \quad 0 \leq \frac{\|r_3(h)\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} \leq \frac{\|g'(f(x^0))(r_1(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} + \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}}. \tag{6}$$

Z faktu, że odwzorowanie liniowe i ciągłe $T : \mathbb{R}^m \to \mathbb{R}^l$ jest ograniczone tzn., że

$$\exists\, c > 0\ \forall\, y \in \mathbb{R}^m \quad \|T(y)\|_{\mathbb{R}^l} \leq c\|y\|_{\mathbb{R}^m}$$

mamy, że

$$\exists\, c > 0\ \forall\, \|h\|_{\mathbb{R}^k} < \delta \quad \|g'(f(x^0))(r_1(h))\|_{\mathbb{R}^l} \leq c\|r_1(h)\|_{\mathbb{R}^m}. \tag{7}$$

Zatem wobec powyższego oraz (2) dostajemy, że

$$\lim_{h \to 0_{\mathbb{R}^k}} \frac{\|g'(f(x^0))(r_1(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} = 0. \tag{8}$$

Ponadto

$$\forall\, (0 < \|h\|_{\mathbb{R}^k} < \delta \wedge\ d(h) \neq 0_{\mathbb{R}^m}) \quad \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} = \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|d(h)\|_{\mathbb{R}^m}} \cdot \frac{\|d(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^k}}.$$

Korzystając z (3) mamy, że dla $0 < \|h\|_{\mathbb{R}^k} < \delta \wedge\ d(h) \neq 0_{\mathbb{R}^m}$

$$\frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} = \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|d(h)\|_{\mathbb{R}^m}} \cdot \frac{\|f'(x^0)(h)+r_1(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^k}} \leq \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|d(h)\|_{\mathbb{R}^m}} \cdot \frac{\|f'(x^0)(h)\|_{\mathbb{R}^m}+\|r_1(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^k}}. \tag{9}$$

Ponieważ $f'(x^0)$ jest odzworowaniem liniowym i ciągłym działającym z $\mathbb{R}^k$ do $\mathbb{R}^m$, więc

$$\exists\, c_1 > 0\ \forall\, h \in \mathbb{R}^k \quad \|f'(x^0)(h)\|_{\mathbb{R}^m} \leq c_1\|h\|_{\mathbb{R}^k}. \tag{10}$$

Zatem z (9), (10) dostajemy

$$\forall\, (0 < \|h\|_{\mathbb{R}^k} < \delta \wedge\ d(h) \neq 0_{\mathbb{R}^m}) \quad \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} \leq \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|d(h)\|_{\mathbb{R}^m}} \cdot \left(c_1 + \frac{\|r_1(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^k}}\right). \tag{11}$$

Pamiętając, że $d(0_{\mathbb{R}^k}) = 0_{\mathbb{R}^m}$ oraz $r_2(0_{\mathbb{R}^m}) = 0_{\mathbb{R}^l}$ oraz wobec ciągłości funkcji d w $0_{\mathbb{R}^m}$ oraz (2) i (3) mamy

$$\lim_{h \to 0_{\mathbb{R}^k}} \frac{\|r_2(d(h))\|_{\mathbb{R}^l}}{\|h\|_{\mathbb{R}^k}} = 0. \tag{12}$$

Wobec (8) otrzymujemy tezę, czyli (5). $\square$

Uwaga 5. Niech $f : \mathbb{R}^k \to \mathbb{R}^m$, $f = (f_1, \dots, f_m)$, $g : \mathbb{R}^m \to \mathbb{R}^l$, $g = (g_1, \dots, g_l)$ są różniczkowalne odpowiednio w punkcie $x^0 \in \mathbb{R}^k$ i $f(x^0) \in \mathbb{R}^m$ tzn.

$$f'(x^0) = [(f_i)'_{x_j}(x^0)]_{i \le m, j \le k}, \quad g'(f(x^0)) = [(g_r)'_{y_i}(f(x^0))]_{r \le l, i \le m}. \tag{13}$$

Wówczas z powyższego twierdzenia mamy, że $\varphi = g \circ f : \mathbb{R}^k \to \mathbb{R}^l$, $\varphi = (\varphi_1, \dots, \varphi_l)$ jest różniczkowalna w x^0 oraz $\varphi'(x^0) = g'(f(x^0)) \bullet f'(x^0)$ tzn., że

$$\forall\, r = 1, \dots, l \;\forall\, j = 1, \dots, k \quad (\varphi_r)'_{x_j}(x^0) = \sum_{i=1}^{m} (g_r)'_{y_i}(f(x^0)) \cdot (f_i)'_{x_j}(x^0)$$

albo równoważnie

$$\forall\, r = 1, \dots, l \;\forall\, j = 1, \dots, k \quad \tfrac{\partial \varphi_r}{\partial x_j}(x^0) = \sum_{i=1}^{m} \tfrac{\partial g_r}{\partial y_i}(f(x^0)) \cdot \tfrac{\partial f_i}{\partial x_j}(x^0).$$

Pochodne wyższych rzędów funkcji wielu zmiennych

Definicja 4. *Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $f : U \to \mathbb{R}$, $x^0 \in U$. Załóżmy, że pochodna w sensie Frécheta funkcji f istnieje w każdym punkcie zbioru U. Jeżeli odzworowanie*

$$U \ni x \to f'(x) \in \mathcal{L}(\mathbb{R}^k, \mathbb{R})$$

jest różniczkowalne w sensie Frécheta w punkcie x^0, to mówimy, że ***funkcja** f **jest dwukrotnie różniczkowalna (w sensie Frécheta, w sposób mocny) w punkcie** x^0. Druga pochodna funkcji f w punkcie x^0 jest odwzorowaniem liniowym i ciągłym $A : \mathbb{R}^k \to \mathcal{L}(\mathbb{R}^k, \mathbb{R})$ i oznaczamy ją $A = f''(x^0)$.*

Uwaga 6. Zauważmy, że $f''(x^0) \in \mathcal{L}(\mathbb{R}^k, \mathcal{L}(\mathbb{R}^k, \mathbb{R}))$. Do tej pory na tym wykładzie nie definiowaliśmy pochodnej odzworowań działających z $\mathbb{R}^k$ w $\mathcal{L}(\mathbb{R}^k, \mathbb{R})$. W praktyce sprowadza się to liczenia pochodnych cząstkowych odpowiednich funkcji wektorowych będących pochodnymi cząstkowymi wyjściowej funkcji, a więc do liczenia pochodnych cząstkowych rzędu drugiego.

Definicja 5. *Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $f : U \to \mathbb{R}$, $x^0 \in U$, $h^{(1)}$, $h^{(2)} \in \mathbb{R}^k$. Pochodną* ***kierunkową rzędu drugiego funkcji** f **w punkcie** x^0 **w kierunku wektorów** $h^{(1)}$, $h^{(2)}$ nazywamy pochodną kierunkową funkcji*

$$U \ni x \to f'_{h^{(1)}}(x)$$

w punkcie x^0 w kierunku wektora $h^{(2)}$ i oznaczamy $f''_{h^{(1)}h^{(2)}}(x^0)$. W szczególności $f''_{e_i e_j}(x^0)$ nazywamy ***pochodną cząstkową rzędu drugiego funkcji** f **w punkcie** x^0 **względem zmiennych** x_i, x_j, gdzie $i, j \in \{1, \dots, k\}$ i oznaczamy $f''_{x_i x_j}(x^0)$ lub $\frac{\partial f}{\partial x_i \partial x_j}(x^0)$.*

Uwaga 7. Z powyższej definicji wynika, że

$$f''_{h^{(1)}h^{(2)}}(x^0) = \left(f'_{h^{(1)}}\right)'_{h^{(2)}}(x^0).$$

W analogiczny (indukcyjny) sposób definiujemy pochodne kierunkowe i cząstkowe rzędu co najmniej trzeciego.

Uwaga 8. Można również stosować notację

$$f''_{x_i x_j}(x^0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0), \quad \frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial f}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

Twierdzenie 5. *(o pochodnej rzędu drugiego) Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $f : U \to \mathbb{R}$, $x^0 \in U$. Wtedy*

(1) Jeśli istnieje pochodna mocna $f''(x^0) \in \mathcal{L}(\mathbb{R}^k, \mathcal{L}(\mathbb{R}^k, \mathbb{R}))$, to dla dowolnych wektorów $h^{(1)}$, $h^{(2)} \in \mathbb{R}^k$ istnieje pochodna kierunkowa $f''_{h^{(1)}h^{(2)}}(x^0)$ przy czym

$$f''_{h^{(1)}h^{(2)}}(x^0) = \sum_{i=1}^{k}\sum_{j=1}^{k} f''_{x_i x_j}(x^0) h_i^{(1)} h_j^{(2)}, \quad \text{gdzie } h^{(l)} = [h_1^{(l)}, \dots, h_k^{(l)}]^T,\ l = 1, 2.$$

(w szczególności istnieją pochodne cząstkowe rzędu drugiego funkcji f w punkcie x^0).

(2) Jezeli w otoczeniu punktu x^0 istnieją wszystkie pochodne cząstkowe rzędu drugiego funkcji f i są ciągłe w w punkcie x^0, to funkcja f jest dwukrotnie różniczkowlana w punkcie x^0.

Uwaga 9. Można pokazać, że $\mathcal{L}(\mathbb{R}^k, \mathcal{L}(\mathbb{R}^k, \mathbb{R}))$ jest izmorficzna z $\mathcal{L}_2(\mathbb{R}^k, \mathbb{R})$, czyli zbiorem form dwuliniowych i ciągłych $A : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$. Z algebry wiadomo, że forma dwuliniowa $A : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ ma postać

$$A(h^{(1)}, h^2) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij} h_i^{(1)} h_j^{(2)},$$

gdzie $h^{(l)} = [h_1^{(l)}, \dots, h_k^{(l)}]^T$, $l = 1, 2$ oraz $[a_{ij}]_{i,j \le k}$ jest taką macierzą, że $a_{ij} = A(e_i, e_j)$. Zatem jeżeli f jest dwukrotnie różniczkowalna w punkcie x^0, to $[a_{ij}]_{i,j\le k} = [f''_{x_i x_j}(x^0)]_{i,j\le k}$.

Twierdzenie 6. ***(Schwarza)*** *Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $x^0 \in U$, $f : U \to \mathbb{R}$. Jeżeli $f''_{x_i x_j}$ istnieją w U i są ciągłe w punkcie x^0 dla $i, j = 1, \dots, k$, $i \neq j$ to*

$$\forall (i, j \in \{1, \dots, k\},\ i \neq j) \quad f''_{x_i x_j}(x^0) = f''_{x_j x_i}(x^0).$$

Dowód. Dowód przeprowadzony na wykładzie. □

Rozważmy przykład funkcji, dla której pochodne mieszane rzędu drugiego nie są równe.

Przykład 2. Rozważmy funkcję $f : \mathbb{R}^2 \to \mathbb{R}$ określoną wzorem

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

Wówczas łatwo sprawdzić, że

$$f'_x(0, y) = -y,\ y \in \mathbb{R}; \quad f''_{xy}(0, 0) = -1$$

oraz

$$f'_y(x,0) = x,\ x \in \mathbb{R};\quad f''_{yx}(0,0) = 1$$

Zatem

$$f''_{xy}(0,0) \neq f''_{yx}(0,0).$$

Definicja 6. *Niech $\underline{U \text{ będzie otwartym podzbiorem } \mathbb{R}^k}$, $f : U \to \mathbb{R}$, $\underline{x^0 \in U}$. Jeżeli f jest dwukrotnie różniczkowalna w punkcie x^0, to macierz drugich pochodnych cząstkowych funkcji f w punkcie x^0 $[f''_{x_ix_j}(x^0)]_{i,j\leq k}$ nazywamy **macierzą Hessego funkcji** f **w punkcie** x^0 i oznaczamy ją $Hf(x^0)$. Wyznaczynik macierzy Hessego funkcji f w punkcie x^0 nazywamy **hesjanem funkcji** f **w punkcie** x^0.*

Uwaga 10. Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $f : U \to \mathbb{R}$, $x^0 \in U$. Jeżeli $f''_{x_ix_j}$ istnieją w pewnym otoczeniu punktu x^0 i są ciągłe w x^0, dla $i, j = 1, \ldots, k$, $i \neq j$, to z twierdzenia 6 Schwarza mamy, że $Hf(x^0)$ jest macierzą symetryczną.

Analogicznie jak w przypadku funkcji $f : \mathbb{R} \to \mathbb{R}$ pochodną Frécheta $n + 1$ rzędu funkcji f w punkcie x^0 określamy jako pochodną odzworowania będącego pochodną n-tego rzędu.

Definicja 7. *Niech $\underline{U \text{ będzie otwartym podzbiorem } \mathbb{R}^k}$, $f : U \to \mathbb{R}$, $\underline{x^0 \in U}$, $n \in \mathbb{N}$, $n > 2$. Jeżeli f jest n-krotnie różniczkowalna U oraz odwzorowanie $U \ni x \to f^{(n)}(x)$ jest różniczkowalne w punkcie x^0, to mówimy, że **funkcja** f **jest** $(n+1)$**-krotnie różniczkowalna (w sensie Frécheta, w sposób mocny) w punkcie** x^0 i jest ona odwzorowaniem $(n+1)$-liniowym tzn. $f^{(n+1)}(x^0) = (f^n)'(x^0) \in \mathcal{L}(\mathbb{R}^k, \mathcal{L}_n(\mathbb{R}^k, \mathbb{R}))$.*

Uwaga 11. Jeżeli funkcja f jest n-krotnie różniczkowalna w punkcie x^0, to dla dowolnych $h^{(l)} = \left[h_1^{(l)}, \ldots, h_k^{(l)}\right]^T \in \mathbb{R}^k$, $l = 1, \ldots, n$ mamy, że

$$f^{(n)}(x^0)(h^{(1)}, \ldots, h^{(n)}) = f^{(n)}_{h^{(1)},\ldots,h^{(n)}}(x^0) = \sum_{i_1=1}^{k} \cdots \sum_{i_n=1}^{k} f^{(n)}_{x_{i_1},\ldots,x_{i_n}}(x^0) \cdot h^{(1)}_{i_1} \cdot \ldots \cdot h^{(n)}_{i_n}.$$

Wzór Taylora dla funkcji wielu zmiennych i jego zastosowania

Definicja 8. *Niech $\underline{U \text{ będzie otwartym podzbiorem } \mathbb{R}^k}$, $f : U \to \mathbb{R}$, $n \in \mathbb{N}$. Mówimy, że **funkcja** f **jest klasy** $C^{(n)}$ **na** U, gdy w zbiorze U istnieją wszystkie pochodne cząstkowe rzędu n oraz są one funkcjami ciągłymi w zbiorze U.*

Twierdzenie 7. ***(wzór Taylora dla funkcji** k **zmiennych)** Niech U będzie otwartym podzbiorem $\mathbb{R}^k$, $x^0 \in U$, $n \in \mathbb{N}$. Jeżeli $I(x^0, x^0 + h)$-odcinek o końcach x^0, $x^0 + h$, gdzie $h \in \mathbb{R}^k$ zawiera się w U, to dla dowolnej funkcji $f : U \to \mathbb{R}$ klasy $C^{(n)}$ na U istnieje taka liczba $\theta \in (0, 1)$, że*

$$f(x^0 + h) = f(x^0) + \tfrac{1}{1!}f'(x^0)(h) + \tfrac{1}{2!}f''(x^0)(h, h) + \ldots, + \tfrac{1}{(n-1)!}f^{(n-1)}(x^0)\underbrace{(h, \ldots, h)}_{(n-1)-razy} +$$

$$+ \tfrac{1}{n!}f^{(n)}(x^0 + \theta h)\underbrace{(h, \ldots, h)}_{n-razy}.$$

Dowód. Dowód przeprowadzony na wykładzie. $\square$

Uwaga 12. Wzór w twierdzeniu 7 nazywamy wzorem Taylora rzędu n. W szczególności wzór Taylora rzędu 2 ma postać

$$f(x^0+h) = f(x^0) + \nabla f(x^0)h + \tfrac{1}{2}h^T Hf(x^0+\theta h)h,$$

gdzie $h = [h_1, \dots, h_k]^T \in \mathbb{R}^k$.

Definicja 9. *Niech A będzie symetryczną macierzą kwadratową wymiaru (kształtu) $k \times k$. Funkcję $Q_A : \mathbb{R}^k \to \mathbb{R}$ określoną następująco*

$$Q_A(x) := \big\langle x, Ax \big\rangle, \quad x \in \mathbb{R}^k$$

*nazywamy **formą kwadratową wyznaczoną przez macierz** A.*

Uwaga 13. Zauważmy, że dla A macierzy symetrycznej kszałtu $k \times k$

$$\big\langle x, Ax \big\rangle = x^T Ax, \quad x \in \mathbb{R}^k.$$

Definicja 10. *Niech A będzie symetryczną macierzą kwadratową kszałtu $k \times k$,a Q_A formą kwadratową wyznaczoną przez macierz A. Mówimy, że macierz A (forma kawadratowa Q_A) jest*
*(i) **dodatnio określona**, gdy $Q_A(x) > 0$, $\forall\ x \in \mathbb{R}^k \setminus \{0_{\mathbb{R}^k}\}$;*
*(ii) **ujemnie określona**, gdy $Q_A(x) < 0$, $\forall\ x \in \mathbb{R}^k \setminus \{0_{\mathbb{R}^k}\}$;*
*(iii) **nieokreślona**, gdy $\exists\ x, y \in \mathbb{R}^k$, $Q_A(x) < 0$, $Q_A(y) > 0$.*

Wprowadźmy oznaczenie: Dla symetrycznej macierzy kwadratowej A wymiaru $k \times k$ symbolem Δ_i, $i = 1, \dots, k$ oznaczać będziemy minory główne macierzy A.

Twierdzenie 8 (**Sylvestera**)**.** *Niech A będzie symetryczną macierzą kwadratową kszałtu $k \times k$. Wówczas*
(i) $\Delta_i > 0$, $i = 1 \dots, k$, wtedy i tylko wtedy, gdy macierz A oraz Q_A- forma kwadratowa wyznaczona przez macierz A jest dodatnio określona.
(ii) $(-1)^i \Delta_i > 0$, $i = 1 \dots, k$, wtedy i tylko wtedy, gdy macierz A oraz Q_A-forma kwadratowa wyznaczona przez macierz A jest ujemnie określona.

Konsekwencją wzoru Taylora rzędu dwa jest następujące twierdzenie:

Twierdzenie 9. ***(warunek dostateczny istnienia ekstremum lokalnego)*** *Niech $D \subset \mathbb{R}^k$, $\operatorname{Int} D \neq \emptyset$, $x^0 \in \operatorname{Int} D$, $f : D \to \mathbb{R}$ klasy $C^{(2)}$ na pewnym otoczeniu punktu x^0. Załóżmy, że x^0 jest punktem stacjonarnym funkcji f. Wówczas:*

(1) funkcja f osiąga w punkcie x^0 ścisłe minimum lokalne, gdy $Hf(x^0)$ jest dodatnio określona.

(2) funkcja f osiąga w punkcie x^0 ścisłe maksimum lokalne, gdy $Hf(x^0)$ jest ujemnie określona.

(3) funkcja f nie osiąga w punkcie x^0 ekstremum lokalnego, gdy $Hf(x^0)$ jest nieokreślona.

Dowód. Dowód przeprowadzony na wykładzie. $\square$

Przykład 3. Niech $D \subset \mathbb{R}^2$, $\operatorname{Int} D \neq \emptyset$, $(x_0, y_0) \in \operatorname{Int} D$, $f : D \to \mathbb{R}$ klasy $C^{(2)}$ na $\operatorname{Int} D$ oraz $\nabla f(x_0, y_0) = [0, 0]$. Dla punktu stacjonarnego (x_0, y_0) mamy:
i) jeżeli $\Delta_1(x_0, y_0) > 0 \ \wedge \ \Delta_2(x_0, y_0) > 0$, to funkcja f osiąga ścisłe mimium lokalne w punkcie (x_0, y_0).
ii) jeżeli $\Delta_1(x_0, y_0) < 0 \ \wedge \ \Delta_2(x_0, y_0) > 0$, to funkcja f osiąga ścisłe maksimum lokalne w punkcie (x_0, y_0).
iii) jeżeli $\Delta_2(x_0, y_0) < 0$, to funkcja f nie osiąga ekstremum lokalnego w (x_0, y_0).
iv) jeżeli $\Delta_2(x_0, y_0) = 0$, to NIC nie można powiedzieć o zachowaniu funkcji f w punkcie x_0. (Przykłady: $f_1(x, y) = x^3 + y^3$ brak ekstremum lokalnego w punkcie $(0, 0)$, $f_2(x, y) = x^2 + y^4$ - f_2 osiąga ścisłe minimum globalne w punkcie $(0, 0)$, a tym samym ścisłe minimum lokalne. Zauważmy, że obu przypadkach hesjan każdej z funkcji w punkcie $(0, 0)$ jest równy 0, czyli $\Delta_2(0, 0) = 0$).