Oznaczenia

Umawiamy się, że

$$(x_1, \ldots, x_N) := [x_1 \ldots x_N]^\mathsf{T} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

Przez e_k oznaczamy k-ty wektor bazowy ze standardowej bazy przestrzeni $\mathbb{R}^N$, a więc

$$e_k := (0, 0, \ldots, 0, \overbrace{1}^{k\text{-ta pozycja}}, 0, \ldots, 0, 0) = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Jeżeli mamy dany zbiór G i punkt $x \in G$, to $G - x = \{y - x \colon y \in G\}$. Wówczas

$$y \in G - x \iff y + x \in G.$$

Pochodne kierunkowe i cząstkowe

Definicja 1. *Pochodną cząstkową* funkcji $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ po zmiennej x_k definiujemy jako granicę (jeśli istnieje)

$$\lim_{h\to 0} \frac{f(x_1, x_2, \ldots, x_k + h, \ldots, x_n) - f(x_1, x_2, \ldots, x_k, \ldots, x_n)}{h}$$

i oznaczamy jako

$$\frac{\partial f}{\partial x_k}(x_1, \ldots, x_n) \text{ lub } \frac{\partial f(x_1, \ldots, x_n)}{\partial x_k}$$

Definicja 2 (Pochodna Kierunkowa)**.** Niech dana będzie funkcja $f\colon U \to \mathbb{R}$, gdzie $U \subseteq \mathbb{R}^n$ jest zbiorem niepustym i otwartym. *Pochodną kierunkową* funkcji f w punkcie a w kierunku niezerowego wektora $u \in U$ nazywamy (jeżeli istnieje) granicę

$$\lim_{t\to 0} \frac{f(a + tu) - f(a)}{t}.$$

Oznaczamy powyższą granicę symbolem $\nabla_u f(a)$ lub $\partial_u f(a)$.

Stosowane są także oznaczenia $\dfrac{\partial f}{\partial u}(a)$ i $f'_u(a)$.
Oczywisty jest związek

$$\frac{\partial f}{\partial x_k}(a) = \nabla_{e_k} f(a).$$

Definicja 3. Definiujemy *gradient* funkcji f w punkcie a jako

$$\operatorname{grad} f(a) = \nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a) \right).$$

Symbol ∇ (tzw. nabla) oznacza „operator różniczkowy”, możemy rozumieć, że danej funkcji przyporządkowuje on funkcję wektorową postaci:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n} \right).$$

Gradient funkcji w punkcie jest wektorem wskazującym kierunek najszybszego wzrostu funkcji w danym punkcie. Zatem pochodna kierunkowa liczona w kierunku gradientu będzie mieć największą wartość.

Różniczka i pochodna Frécheta

Rozważamy przestrzenie unormowane $(X, \|\cdot\|_X)$ i $(Y, \|\cdot\|_Y)$.

Niech $f\colon G \to Y$ odwzorowaniem, gdzie $G \subseteq X$ i niech $a \in \operatorname{int} G$. Mówimy, że funkcja f jest różniczkowalna w sensie Frécheta w punkcie f, gdy istnieje takie odwzorowanie liniowe i **ciągłe** $A\colon X \to Y$, że

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - Ah\|_Y}{\|h\|_X} = 0. \tag{F}$$

Odwzorowanie A nazywamy *różniczką* odwzorowania f w punkcie a i oznaczamy $A = \mathrm{d}f\,(a)$. Stosowane są różne inne oznaczenia: $f'(a)$, $\mathrm{d}_a f$, etc.

Twierdzenie 1. *Nastepujące warunki są równoważne:*

(i) funkcja $f\colon G \to Y$ jest różniczkowalna w sensie Frécheta w punkcie a;

(ii) istnieje taka funkcja $r\colon X \to Y$, że

$$f(a+h) - f(a) = Ah + r(h), \quad \text{dla } h \in G - a \tag{1}$$

przy czym

$$\lim_{h\to 0} \frac{r(h)}{\|h\|_X} = 0; \tag{1.1}$$

(iii) istnieje taka funkcja $r\colon X \to Y$, ciągła w zerze, że $r(0) = 0$ oraz dla każdego $h \in G - a$ zachodzi

$$f(a+h) - f(a) = Ah + \|h\|_X\, r(h). \tag{2}$$

Dowód Twierdzenia 1. $(i) \Rightarrow (ii)$. Wystarczy położyć

$$r(h) = \begin{cases} \dfrac{\|f(a+h) - f(a) - Ah\|}{\|h\|}, & \text{dla } h \in G - a \setminus \{0\}; \\ 0, & \text{w pozostałych przypadkach.} \end{cases}$$

Wówczas,

$$\lim_{h\to 0} r(h) = \lim_{h\to 0} \frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} \overset{\text{(F)}}{=} 0 = r(0).$$

$(ii) \Rightarrow (iii)$. Zakładamy, że istnieje odwzorowanie liniowe A oraz funkcja r_1 spełniająca warunki (1) i (1.1), tj. ciągła w zerze, $r_1(0) = 0$ oraz $f(a+h) - f(a) = Ah + \|h\|_X\, r_1(h)$ dla $h \in G - a$. Wówczas, warunek (iii)

jest spełniony z tym samym odwzorowaniem A i z funkcją r zdefiniowaną wzorem $r(h) = \|h\| \, r_1(h)$, $h \in X$. Warunek (2) jest oczywiście spełniony oraz mamy

$$\lim_{h\to 0} \frac{r(h)}{\|h\|} = \lim_{h\to 0} r_1(h) = r_1(0) = 0,$$

co było do wykazania.

$(iii) \Rightarrow (i)$. Warunek (2) zapisujemy w postaci

$$r(h) = f(a+h) - f(a) - Ah.$$

Stąd,

$$\frac{f(a+h) - f(a) - Ah}{\|h\|} = \frac{r(h)}{\|h\|},$$

a ponieważ $\frac{r(h)}{\|h\|} \xrightarrow{h\to 0} 0$, to również $\left\| \frac{r(h)}{\|h\|} \right\| \xrightarrow{h\to 0} 0$ i mamy

$$\frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} = \frac{\|r(h)\|}{\|h\|} = \left\| \frac{r(h)}{\|h\|} \right\| \xrightarrow{h\to 0} 0. \qquad \square$$

Twierdzenie 2. *Jeżeli $f\colon U \to \mathbb{R}$, gdzie $U \subseteq \mathbb{R}^N$ jest niepustym zbiorem otwartym, jest funkcją różniczkowalną w sensie Frécheta w punkcie $a \in U$, to wówczas*

(i) w punkcie a istnieje pochodna kierunkowa $\nabla_u f(a)$ funkcji f w kierunku dowolnego wektora $u \in \mathbb{R}^N$, a ponadto $\nabla_u f(a) = \mathrm{d}f\,(a)\,u$;

(ii) istnieją wszystkie pochodne cząstkowe $\frac{\partial f}{\partial x_i}(a), i \in \{1, \ldots, N\}$ i zachodzi równość

$$\mathrm{d}f\,(a)\,h = \sum_{k=1}^{N} \frac{\partial f}{\partial x_k}(a) h_k,$$

dla dowolnego $h = (h_1, \ldots, h_N) \in \mathbb{R}^N$.

Dowód. Z założenia, istnieje $\varepsilon > 0$, takie że kula $B\,(a, \varepsilon)$ zawiera się w U. Dla $|t| < \varepsilon \, \|u\|^{-1}$ punkt $a + tu$ leży w kuli $B\,(a, \varepsilon)$ i wobec założenia o różniczkowalności f, mamy

$$f(a+tu) - f(a) = \mathrm{d}f\,(a)\,tu + r(tu),$$

gdzie r jest taką funkcją, że $\dfrac{r(tu)}{\|tu\|} \to 0$ przy $\|tu\| \to 0$. Mamy, dla odpowiednio małych t,

$$\begin{aligned}
\nabla_u f(a) &= \lim_{t\to 0} \frac{f(a+tu)-f(a)}{t} \\
&= \lim_{t\to 0} \frac{\mathrm{d}f\,(a)\,(tu)-r(tu)}{t} \\
&= \lim_{t\to 0} \left(\frac{t\,\mathrm{d}f\,(a)\,u}{t} - \frac{r(tu)}{t}\right) \\
&= \lim_{t\to 0} \left(\mathrm{d}f\,(a)\,u - \frac{r(tu)}{t}\cdot\frac{\|u\|}{\|u\|}\right) \\
&= \mathrm{d}f\,(a)\,u - \|u\| \lim_{t\to 0} \frac{r(tu)}{t\,\|u\|},
\end{aligned}$$

a ponieważ $\lim\limits_{t\to 0} \dfrac{r(tu)}{t\,\|u\|} = \lim\limits_{t\to 0} \operatorname{sgn}(t)\dfrac{r(tu)}{\|tu\|} = \pm \lim\limits_{h\to 0} \dfrac{r(h)}{\|h\|} = 0$, to mamy porządaną równość.

Aby wykazać (ii), zauważmy, że dowolny wektor $h = (h_1, \ldots, h_N) \in \mathbb{R}^N$ możemy zapisać w postaci

$$h = \sum_{k=1}^{N} h_k e_k,$$

gdzie e_k jest k-tym wektorem bazy standardowej $\mathbb{R}^N$. Mamy

$$\begin{aligned}
\mathrm{d}f\,(a)\,h = \mathrm{d}f\,(a) \sum_{k=1}^{N} h_k e_k &= \sum_{k=1}^{N} h_k\, \mathrm{d}f\,(a)\, e_k \\
&= \sum_{k=1}^{N} h_k \nabla_{e_k} f(a) = \sum_{k=1}^{N} \frac{\partial f}{\partial x_k}(a) h_k. \qquad \square
\end{aligned}$$

Twierdzenie 3. *Niech $f\colon G \to \mathbb{R}$, gdzie $G \subseteq \mathbb{R}^N$ jest zbiorem otwartym. Wówczas, jeżeli funkcja f ma wszystkie pochodne cząstkowe pierwszego rzędu i są one ciągłe w punkcie $a \in G$, to funkcja f jest różniczkowalna w sensie Frécheta.*