1 Rachunek różniczkowy

1.1 Pochodne cząstkowe, różniczkowalność funkcji, przyrosty

Niech $\mathcal{O} \subset \mathbb{R}^N$ będzie zbiorem otwartym, zaś $f : \mathcal{O} \to \mathbb{R}$ – funkcją ciągłą.

Niech $x \in \mathcal{O}$. Wypiszmy jawnie składowe x:

$$x = (x^1, x^2, \dots, x^N), \quad f(x) = f(x^1, x^2, \dots, x^N).$$

Wybierzmy teraz $k \in \{1, 2, \dots, N\}$ i traktujmy zmienne x^l, gdzie $l \neq k$ jako stałe.

Rozważmy granicę:

$$\frac{\partial}{\partial x^k} f(x^1, x^2, \dots, x^N) = \lim_{h \to 0} \frac{f(x^1, \dots, x^{k-1}, x^k + h, x^{k+1}, \dots, x^N) - f(x^1, \dots, x^k, \dots, x^N)}{h} \tag{1}$$

Def. Powyższą granicę nazywamy *pochodną cząstkową* funkcji f po zmiennej x^k (liczoną w punkcie x).

Def. Niech $\mathcal{O} \subset \mathbb{R}^N$ będzie zbiorem otwartym, zaś $f : \mathcal{O} \to \mathbb{R}$ – funkcją ciągłą. (Ten ostatni warunek piszemy też: $f \in C(\mathcal{O})$).

Mówimy, że f jest różniczkowalna r razy w sposób ciągły, jeżeli istnieją wszystkie pochodne cząstkowe aż do rzędu r i są one ciągłe.

Uwaga. (terminologiczna) Ten ostatni warunek zapisujemy krócej jako: $f \in C^r(\mathcal{O})$, gdzie przez $C^r(\mathcal{O})$ oznaczamy zbiór funkcji różniczkowalnych r razy w sposób ciągły. Stosujemy też oznaczenie $C^\infty(\mathcal{O})$ dla funkcji, które posiadają pochodne ciągłe dowolnie wysokiego rzędu. Funkcje takie nazywamy *funkcjami gładkimi* (należą do nich np. wielomiany).

Tw. Niech $\mathcal{O}$ – otwarty podzbiór $\mathbb{R}^N$, oraz $f \in C^1(\mathcal{O})$. Niech $x_0 \in \mathcal{O}$, i niech $h \in \mathbb{R}^N$ – dostatecznie małe, tzn. $||h|| < \epsilon$ dla pewnego ϵ – tak, by $x_0 + h \in \mathcal{O}$ **RYS.**

Zdefiniujmy $r(x_0, h)$ przez:

$$f(x_0 + h) - f(x_0) = \sum_{k=1}^{N} \frac{\partial f}{\partial x^k}(x_0) h^k + r(x_0, h).$$

Wtedy zachodzi:

$$\frac{r(x_0, h)}{||h||} \xrightarrow{h \to 0} 0 \tag{2}$$

Uwaga. Znaczenie tego wzoru: Pozwala on wyznaczać przyrost funkcji: Im mniejsze h, tym lepiej przyrost jest przybliżany przez część liniową.

Dow. Będzie dla (najważniejszego w zastosowaniach) przypadku $N = 3$; dla dowolnego N jest analogiczny.

$$f(x_0 + h) - f(x_0) = f(x_0^1 + h^1, x_0^2 + h^2, x_0^3 + h^3) - f(x_0^1, x_0^2, x_0^3) =$$

$$\begin{aligned} & f(x_0^1 + h^1, x_0^2 + h^2, x_0^3 + h^3) - f(x_0^1, x_0^2 + h^2, x_0^3 + h^3) \qquad I \\ & + f(x_0^1, x_0^2 + h^2, x_0^3 + h^3) - f(x_0^1, x_0^2, x_0^3 + h^3) \qquad II \\ & + f(x_0^1, x_0^2, x_0^3 + h^3) - f(x_0^1, x_0^2, x_0^3) \qquad III \end{aligned}$$

Z twierdzenia Lagrange'a o wartości średniej dla funkcji jednej zmiennej mamy

$$III = \frac{\partial}{\partial x^3} f(x_0^1, x_0^2, y^3) h^3,$$

gdzie y^3 – punkt pomiędzy x_0^3 a $x_0^3 + h^3$;

$$II = \frac{\partial}{\partial x^2} f(x_0^1, y^2, x_0^3 + h^3) h^2, \qquad y^2 \in]x_0^2, x_0^2 + h^2[;$$

$$I = \frac{\partial}{\partial x^1} f(y^1, x_0^2 + h^2, x_0^3 + h^3) h^1, \qquad y^2 \in]x_0^1, x_0^1 + h^1[.$$

Tak więc

$$\begin{aligned} r(x_0, h) = f(x_0 + h) - f(x_0) - \frac{\partial f}{\partial x^1}(x_0) h^1 - \frac{\partial f}{\partial x^2}(x_0) h^2 - \frac{\partial f}{\partial x^3}(x_0) h^3 \\ = \left[\frac{\partial}{\partial x^1} f(y^1, x_0^2 + h^2, x_0^3 + h^3) - \frac{\partial}{\partial x^1} f(x_0^1, x_0^2, x_0^3) \right] h^1 \\ + \left[\frac{\partial}{\partial x^2} f(x_0^1, y^2, x_0^3 + h^3) - \frac{\partial}{\partial x^2} f(x_0^1, x_0^2, x_0^3) \right] h^2 \\ + \left[\frac{\partial}{\partial x^3} f(x_0^1, x_0^2, y^3) - \frac{\partial}{\partial x^3} f(x_0^1, x_0^2, x_0^3) \right] h^3 \end{aligned} \tag{3}$$

Składowe wektora h szacują się przez:

$$\frac{|h^k|}{||h||} \leqslant 1.$$

Ponadto, jeżeli $h \to 0$, to:

$$y^1 \to x^1; \qquad y^2 \to x^2; \qquad y^3 \to x^3.$$

Ponieważ pochodne cząstkowe są ciągłe, to różnice w (3) dążą do zera i mamy

$$\frac{r(x_0, h)}{||h||} \xrightarrow{h \to 0} 0.$$

CBDO

1.2 Pochodna funkcji złożonej

Usystematyzujmy oznaczenia, przydając im, jeśli trzeba, dodatkowe jeszcze wyjaśnienia:

1. $h = \Delta x$ – przyrost zmiennej (-ych);
2. $f(x + h) - f(x) = \Delta f$ – przyrost funkcji;
3. $\sum_{k=1}^{N} \frac{\partial f}{\partial x^k}(x_0) h^k = \mathsf{d}f$ – różniczka funkcji;
4. r – reszta.

Pamiętajmy, że wszystkie powyższe obiekty: $\Delta x, \Delta f, \mathsf{d}f, r$ są funkcjami od x i h.

Mamy też:

$$\Delta f = \mathsf{d}f + r;$$

im mniejsze h, tym mniejsze r i w wielu zastosowaniach fizycznych na ogół przyjmuje się, że dla małych h, r jest zaniedbywalny.

Tw. (Prototyp twierdzenia o pochodnej odwzorowania złożonego) Niech $g : \mathcal{U} \to \mathcal{O}$, gdzie $\mathcal{U} \subset \mathbb{R}$, $\mathcal{O} \subset \mathbb{R}^N$, oraz $f : \mathcal{O} \to \mathbb{R}$. (Pisząc jawnie argumenty, mamy: $f(y^1, y^2, \ldots, y^N)$ oraz $g(x) = g^1(x), g^2(x), \ldots, g^N(x))$. Niech $k = f \circ g$, tzn. $k(x) = f(g(x))$ lub, pisząc bardziej jawnie, ale też bardziej rozwlekle argumenty: $k(x) = f(g^1(x), g^2(x), \ldots, g^N(x))$.

Jeżeli $f \in C^1(\mathcal{O}), g \in C^1(\mathcal{U})$, to $k \in C^1(\mathcal{U})$ oraz

$$\frac{\mathsf{d}}{\mathsf{d}x} k(x) = \sum_{i=1}^{N} \frac{\partial f}{\partial y^i}(g(x)) \frac{\mathsf{d}g^i}{\mathsf{d}x}(x). \tag{4}$$

Dow. Liczymy iloraz różnicowy:

$$\frac{k(x+h) - k(x)}{h} = \frac{f(g(x+h)) - f(g(x))}{h} = \spadesuit$$

(tu $h \in \mathbb{R}$). Oznaczmy: $\Delta y = g(x+h) - g(x)$ (tzn. $\Delta y^i = g^i(x+h) - g^i(x)$).

$$\spadesuit = \frac{f(g(x) + \Delta y)) - f(g(x))}{h} = \frac{\sum_{i=1}^{N} \frac{\partial f}{\partial y^i}(g(x))\Delta y^i + r(g(x), \Delta y)}{h}$$

$$= \frac{\sum_{i=1}^{N} \frac{\partial f}{\partial y^i}(g(x))\Delta y^i}{h} + \frac{r(g(x), \Delta y)}{h} \tag{5}$$

Pierwszy wyraz w powyższym wyrażeniu (5) to jest to co trzeba, ponieważ

$$\frac{\Delta y^i}{h} = \frac{g^i(x+h)) - g^i(x)}{h} \xrightarrow{h \to 0} \frac{\partial g^i}{\partial x}(x)$$

Natomiast drugi wyraz w wyrażeniu (5) okazuje się ,że dąży do 0 gdy $h \to 0$. Bowiem, gdy $\Delta y = 0$, to $\frac{r(g(x), \Delta y)}{h} = 0$. Natomiast gdy $\Delta y \neq 0$, to mamy:

$$\frac{r(g(x), \Delta y)}{h} = \frac{r(g(x), \Delta y)}{||\Delta y||} \cdot \frac{||\Delta y||}{h}$$

W pierwszym czynniku mamy: $\Delta y \xrightarrow{h \to 0}$ i co za tym idzie, cały wyraz też dąży do zera (z własności reszty). Drugi czynnik, tzn. iloraz $\frac{||\Delta y||}{h}$, spełnia:

$$\frac{||\Delta y||}{h} = \left\| \frac{\Delta y}{h} \right\| \xrightarrow{h \to 0} \left\| \frac{\mathsf{d}g}{\mathsf{d}x} \right\|$$

CBDO

Będziemy dalej potrzebować dwu prostych faktów dotyczących normy i odległości.

Stw. Norma jest funkcją ciągłą swoich argumentów (tzn. składowych wektora).

Dow. Przyjrzyjmy się wyrażeniu na normę wektora x:

$$||x|| = \sqrt{(x^1)^2 + (x^2)^2 + \cdots + (x^N)^2}$$

i mamy:

1. podnoszenie do kwadratu jest funkcją ciągłą,

2. sumowanie jest funkcją ciągłą,

3. pierwiastek jest funkcją ciągłą.

Stw. Odległość jest funkcją ciągłą, tzn. jeżeli $x_n \overset{n\to\infty}{\longrightarrow} x$, $y_n \overset{n\to\infty}{\longrightarrow} y$, to

$$d(x_n, y_n) \overset{n\to\infty}{\longrightarrow} d(x, y). \tag{6}$$

Dow. Najsampierw pokażemy następującą *nierówność czworoboku*:

$$d(x, u) \leqslant d(x, y) + d(y, z) + d(z, u);$$

bowiem, pisząc dwukrotnie nierówność trójkąta, mamy:

$$d(x, u) \leqslant d(x, z) + d(z, u) \leqslant d(x, z) + d(z, y) + d(y, u)$$

(szkoda tu pisać **CBDO**) i teraz, korzystając dwukrotnie z nierówności czworoboku:

$$d(x_n, y_n) \leqslant d(x_n, x) + d(x, y) + d(y, y_n) \Longrightarrow d(x_n, y_n) - d(x, y) \leqslant d(x_n, x) + d(y, y_n);$$

$$d(x, y) \leqslant d(x, x_n) + d(x_n, y_n) + d(y_n, y) \Longrightarrow d(x, y) - d(x_n, y_n) \leqslant d(x, x_n) + d(y_n, y);$$

czyli mamy

$$0 \leqslant |d(x, y) - d(x_n, y_n)| \leqslant d(x, x_n) + d(y_n, y);$$

Prawa strona powyższej nierówności z założenia dąży do zera, a z tego wynika (6).

CBDO

1.3 Równość drugich pochodnych mieszanych

Tw. (o równości drugich pochodnych mieszanych). Niech $\mathcal{O} \subset \mathbb{R}^2$ – zbiór otwarty. Niech $f : \mathcal{O} \to \mathbb{R}$, $f \in C^2(\mathcal{O})$. Wtedy:

$$\frac{\partial}{\partial x^1}\frac{\partial f}{\partial x^2} = \frac{\partial}{\partial x^2}\frac{\partial f}{\partial x^1} \tag{7}$$

Dow. Najsampierw oznaczmy x zamiast x^1 oraz y zamiast x^2.

Oznaczmy:

$$\Delta x = h, \qquad \Delta y = k$$

oraz

$$w = f(x + h, y + k) - f(x + h, y) - f(x, y + k) + f(x, y)$$

Ustalmy teraz x i h i zdefiniujmy

$$\phi(y) = f(x + h, y) - f(x, y)$$

Możemy wyrazić w przez ϕ:

$$w = \phi(y + k) - \phi(y) = \phi'(\xi) \cdot k$$

gdzie $\xi \in]y, y + k[$; jest to wniosek z tw. Lagrange'a o wartości średniej. Mamy dalej

$$\phi'(\xi) \cdot k = \left[\frac{\partial f}{\partial y}(x + h, \xi) - \frac{\partial f}{\partial y}(x, \xi)\right] = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(\eta, \xi) \cdot h \cdot k$$

gdzie $\eta \in]x, x+h[$; w ostatnim kroku znów skorzystaliśmy z tw. Lagrange'a o wartości średniej.

Zdefiniujmy teraz

$$\psi(x) = f(x, y+k) - f(x, y)$$

i podobnie jak wyżej, możemy wyrazić w przez ψ. Mamy:

$$w = \psi(x+h) - \psi(x) = \psi'(\tilde{\eta}) \cdot h = \spadesuit$$

(tu $\tilde{\eta} \in]x, x+h[$; znów skorzystaliśmy z tw. Lagrange'a) i dalej

$$\spadesuit = \left[\frac{\partial f}{\partial x}(\tilde{\eta}, y+k) - \frac{\partial f}{\partial x}(\tilde{\eta}, y)\right] = \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(\tilde{\eta}, \tilde{\xi}) \cdot h \cdot k$$

gdzie $\tilde{\xi} \in]y, y+k[$.

W powyższych wyrażeniach liczyliśmy tę samą wielkość w na dwa różne sposoby. Mamy więc:

$$w = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(\eta, \xi) \cdot h \cdot k = \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(\tilde{\eta}, \tilde{\xi}) \cdot h \cdot k$$

Jeśli teraz $h \to 0$, to wtedy $\xi \to y$, $\eta \to x$, oraz $\tilde{\xi} \to y$, $\tilde{\eta} \to x$, zatem otrzymujemy równość pochodnych cząstkowych mieszanych w punkcie (x, y).

CBDO

Przykł.: Nie można opuścić założeń o ciągłości; Nierówność pochodnych mieszanych gdy f nie jest kl. C^2 .

1.4 Wyższe pochodne

Notacja pochodnych cząstkowych. Mówiąc o pochodnej cząstkowej trzeba podać nie tylko jej rząd (ilość różniczkowań), ale też powiedzieć, po jakich zmiennych się różniczkuje. Ogólnie pochodna r-tego rzędu funkcji N zmiennych ma postać:

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y^2},$$

$$\frac{\partial^r f}{(\partial x^1)^{r_1}(\partial x^2)^{r_2}\dots(\partial x^N)^{r_N}}, \quad \text{gdzie} \quad r = r_1 + r_2 + \dots + r_N;$$

i tak wszystkie drugie pochodne funkcji dwóch zmiennych są

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y^2},$$

pochodne trzeciego rzędu:

$$\frac{\partial^3 f}{\partial x^3}, \quad \frac{\partial^3 f}{\partial x^2 \partial y}, \quad \frac{\partial^3 f}{\partial x \partial y^2}, \quad \frac{\partial^3 f}{\partial y^3},$$

itd.

1.5 Różniczkowalność odwzorowań

Powyższe uwagi dotyczyły *funkcji* N zmiennych. Gdy mamy odwzorowania, różniczkowalność tychże definiujemy analogicznie:

Niech T odwzorowuje $\mathcal{O} \subset \mathbb{R}^N$ na $\mathcal{U} \subset \mathbb{R}^M$, $\mathcal{O}$ i $\mathcal{U}$ są otwarte. **RYS.**; niech $x \in \mathcal{O}$, $y \in \mathcal{U}$.

Niech $y = T(x)$. Wypiszmy tę równość jawnie w składowych:

$$y^k = T^k(x^1, x^2, \ldots, x^N), \quad k = 1, 2, \ldots, M.$$

Def. Mówimy, że T jest klasy C^r, jeżeli wszystkie funkcje $T^k \in C^r(\mathcal{O})$.

Można wypisywać wszystkie pochodne cząstkowe rzędu r dla odwzorowania – wzory są podobne jak na pochodną funkcji, tylko nieco bardziej skomplikowane. Będziemy je wypisywać, gdy będzie to potrzebne, a na razie wypiszmy jawnie *pierwszą* pochodną odwzorowania:

$$T'(x) = \begin{bmatrix} (T^1(x))' \\ (T^2(x))' \\ \vdots \\ (T^M(x))' \end{bmatrix} = \begin{bmatrix} \frac{\partial T^1}{\partial x^1} & \frac{\partial T^1}{\partial x^2} & \cdots & \frac{\partial T^1}{\partial x^N} \\ \frac{\partial T^2}{\partial x^1} & \frac{\partial T^2}{\partial x^2} & \cdots & \frac{\partial T^2}{\partial x^N} \\ \vdots & \vdots & \ldots\vdots & \\ \frac{\partial T^M}{\partial x^1} & \frac{\partial T^M}{\partial x^2} & \cdots & \frac{\partial T^M}{\partial x^N} \end{bmatrix} \tag{8}$$

tzn. na skrzyżowaniu i-tego wiersza i j-tej kolumny mamy pochodną $\dfrac{\partial T^i}{\partial x^j}$. Pamiętajmy, że każda z pochodnych $\dfrac{\partial T^i}{\partial x^j}$ jest funkcja N zmiennych $x^1, \ldots, x^N$.

Oznaczenia & terminologia: Pochodną odwzorowania (8) oznacza się też czasem DT. Taka tablica liczb, jak pamiętamy z części algebraicznej wykładu, nazywa się *macierzą*; w tym konkretnym przypadku mamy *macierz pochodnej* odwzorowania albo *macierz Jacobiego.*

Przykł. Macierz Jacobiego zamiany współrzędnych kartezjańskich na biegunowe.

1.6 Pochodna odwzorowania złożonego

Niech: $X \subset \mathbb{R}^N$, $Y \subset \mathbb{R}^M$, $Z \subset \mathbb{R}^k$ oraz $S : X \to Y$, $T : Y \to Z$. Pamiętamy, że *superpozycją* (*złożeniem* odwzorowań S i T nazywamy odwzorowanie $T \circ S : X \to Z$, określone jako: $(T \circ S)(x) = T(S(x))$.

RYS.

Pamiętamy też twierdzenie, że jeśli S i T są odwzorowaniami ciągłymi, to $T \circ S$ też jest odwzorowaniem ciągłym.

Wyprowadziliśmy niedawno wzór (4) na pochodną odwzorowania złożonego w przypadku, gdy X było podzbiorem $\mathbb{R}$, Y podzbiorem $\mathbb{R}^n$, oraz $Z \subset \mathbb{R}$:

RYS.

Zastosujmy teraz ten wzór w przypadku, gdy mamy pochodną $\dfrac{\partial (T \circ S)^k}{\partial x^l}(x)$.

RYS.

Mamy:

$$\frac{\partial (T \circ S)^k}{\partial x^l}(x) = \frac{\partial T^k(S^1(x^1, x^2, \ldots, x^N), S^2(x^1, x^2, \ldots, x^N), \ldots, S^M(x^1, x^2, \ldots, x^N))}{\partial x^l}(x)$$

$$= \frac{\partial T^k}{\partial y^1}(S(x)) \cdot \frac{\partial S^1}{\partial x^l}(x) + \frac{\partial T^k}{\partial y^2}(S(x)) \cdot \frac{\partial S^2}{\partial x^l}(x) + \cdots + \frac{\partial T^k}{\partial y^M}(S(x)) \cdot \frac{\partial S^M}{\partial x^l}(x)$$

$$= \sum_{j=1}^{M} \frac{\partial T^k}{\partial y^j}(S(x)) \cdot \frac{\partial S^j}{\partial x^l}(x). \tag{9}$$

To była konkretna składowa $\frac{\partial (T \circ S)^k}{\partial x^l}(x)$. Dla całej macierzy Jacobiego można wypisać wzór, przypominający jako żywo pochodną funkcji złożonej – ale aby go prawidłowo rozczytać, trzeba pamiętać, co oznaczają poszczególne symbole:

$$(T \circ S)'(x) = T'(S(x)) \cdot S'(x)$$

gdzie kropka oznacza *mnożenie macierzy* pochodnych.

Jeśli zarówno S jak i T są odwzorowaniami różniczkowalnymi klasy C^1, to i ich złożenie też jest odwzorowaniem różniczkowalnym klasy C^1. Wynika to od razu z faktu, że sumy i iloczyny odwzorowań różniczkowalnych typu (4) też są różniczkowalne, a we wzorze (9) są tylko sumy i iloczyny takich wyrażeń.

Analogicznie się argumentuje pokazując, że jeśli S oraz T są odwzorowaniami różniczkowalnymi klasy C^r, to i ich złożenie też jest odwzorowaniem różniczkowalnym klasy C^r.

Powyższe można podsumować w twierdzeniu:

Tw. Jeśli S i T są odwzorowaniami klasy C^r, to również ich złożenie $T \circ S$ jest klasy C^r. Pierwsza pochodna odwzorowania złożonego dana jest wzorem

$$(T \circ S)'(x) = T'(S(x)) \cdot S'(x) \tag{10}$$

or, in more explicit manner,

$$\frac{\partial (T \circ S)^k}{\partial x^l}(x) = \sum_{j=1}^{M} \frac{\partial T^k}{\partial y^j}(S(x)) \cdot \frac{\partial S^j}{\partial x^l}(x). \tag{11}$$

Przykł. $S : \mathbb{R}^2 \ni (r, \phi) \to (x = r\cos\phi, y = r\sin\phi) \in \mathbb{R}^2$, $T : \mathbb{R}^2 \ni (x, y) \to (u = exp(x + y), v = \exp(x - y) \in \mathbb{R}^2$. Policzymy pochodną wprost oraz jako iloczyn macierzy pochodnych S i T. I tu dobrze żeby był jakiś przykład.