16 Różniczki i ich własności

Twierdzenie. *Jeśli pochodne cząstkowe $\frac{\partial f_i}{\partial x_j}$ istnieją $(\forall_{i\leqslant k}, \forall_{j\leqslant d})$ w pewnym otoczeniu punktu i są ciągłe w tym punkcie, to odwzorowanie $F = (f_1, \ldots, f_k)$ jest w tym punkcie różniczkowalne.*

Dowód był naszkicowany dla 2 zmiennych, k=1. Spróbujmy to zrobić w sytuacji ogólnej.

Dowód. Jak już wiemy, założenie $k = 1$ nie zmniejsza ogólności -to rezultat analogicznego faktu dla granic. Punkt z wnętrza dziedziny, w którym badamy różniczkowalność oznaczmy przez P, zaś wektor przyrostu zmiennej niezależnej przez h, czyli w bazie kanonicznej wektorów 0-1 -kowych $e_1, \ldots, e_d$ przestrzeni $\mathbb{R}^d$ mamy $h = h_1e_1 + \cdots + h_de_d$ i rozważamy przyrost funkcji f (bo $k = 1$) -czyli różnicę: $\Delta f = f(P + h) - f(P)$, którą rozkładamy na sumę przyrostów w kierunkach poszczególnych osi (czyli wektorów e_k): niech $P_0 = P$, zaś

$$P_1 = P+h_1e_1, P_2 = P_1+h_2e_2 = P+(h_1, h_2, 0, \ldots, 0), ..., P_d = P_{d-1}+h_de_d = P+h.$$

Teraz dla $\Delta_k f := f(P_k) - f(P_{k-1})$ mamy

$$\Delta f = \Delta_1 f + \Delta_2 f + \cdots + \Delta_d f.$$

Dla każdego z "częściowych" przyrostów $\Delta_k f$ stosujemy twierdzenie Lagrange'a znajdując punkt pośredni pomiędzy P_k oraz P_{k-1} - oznaczmy go $Q_k = P_{k-1} + \theta_k h_k e_k$ dla pewnej liczby $\theta_k \in (0, 1)$ tak, by

$$\Delta_k f = h_k \frac{\partial f}{\partial x_k}(Q_k).$$

Faktycznie, jeśli na odcinku $[0, 1]$ określimy funkcję $\phi_k(t) := f(P_{k-1}+th_ke_k)$, to jej przyrost pomiędzy punktami 0 oraz 1 wynosi $\phi_k(1) - \phi_k(0) = (1-0)\phi_k'(\theta_k)$ dla pewnej $\theta_k \in (0, 1)$. Ten przyrost, to dokładnie $\Delta_k f$. Natomiast pochodna $\phi_k'(t) = h_k \frac{\partial f}{\partial x_k}(P_{k-1} + th_ke_k)$- bo z dokładnością do podstawienia $t \mapsto th_k$ (pochodną tego podstawienia jest stała h_k), to jest funkcja, której pochodną jest właśnie pochodna cząstkowa z f.

Licznikiem w ułamku z definicji różniczkowalności będzie

$$\Delta f - \sum_{k=1}^{d} \frac{\partial f}{\partial x_k}(P) \cdot h_k = \sum_{k=1}^{d} h_k \big\{ \frac{\partial f}{\partial x_k}(Q_k) - \frac{\partial f}{\partial x_k}(P) \big\}.$$

Gdy podzielimy wszystko przez $\|h\|$, to przed nawiasami { } pojawią się czynniki $\frac{h_k}{\|h\|}$, których moduł nie przekracza 1. Każdy z nawiasów, to różnica zmierzająca do zera przy $\|h\| \to 0$- dzięki ciągłości pochodnych cząstkowych, gdyż odległości między punktami Q_k oraz P, to $\|Q_k - P\| = (\sum_{j=1}^{k} h_j^2)^{\frac{1}{2}} \leqslant \|h\|$ -więc zmierzją do zera. □

(Uwaga: występujący w dowodzie symbol θ, to nie jest "stylizowane zero", tylko grecka litera theta.) Powyższe twierdzenie daje tylko warunek wystarczający. Jest on daleki od koniecznego -bo już nawet dla funkcji 1 zmiennej różniczkowalność w punkcie (np. w $t = 0$) nie implikuje ciągłości w żadnym innym punkcie, np. gdy ψ jest funkcją Dirichleta (lub jakąkolwiek inną funkcją ograniczoną), to funkcja $t \mapsto t^2\psi(t)$ ma pochodną równą 0 w punkcie $t = 0$, który może być jedynym jej punktem ciągłości.

Zaletą (zwłaszcza zapisanej bezargumentowo) różniczki w stosunku do pochodnych cząstkowych jest unikanie żmudnych (często -wielowskaźnikowych) sumowań. Następne twierdzenie jest tego wyraźnym przykładem. Chodzi o różniczkę złożenia.

Twierdzenie (Reguła łańcucha). Jeżeli odwzorowanie $f : D \subset \mathbb{R}^d \to \mathbb{R}^k$ jest różniczkowalne w punkcie wewnętrznym x zbioru D, $y = f(x)$ oraz $g : D_1 \subset \mathbb{R}^k \to \mathbb{R}^m$ określone w otoczeniu D_1 punktu y jest w tym punkcie różniczkowalne, to złożenie $g \circ f$ jest różniczkowalne w punkcie x, a jego różniczka jest złożeniem odpowiednich różniczek:

$$d_x(g \circ f) = d_{f(x)}g \circ d_x f. \tag{1}$$

Dowód.

Dziedziną złożenia $g \circ f$ jest $D \cap f^{-1}[D_1]$. Z warunku ciągłości f w punkcie x wiemy, że ten przeciwobraz, $f^{-1}[D_1]$ jest otoczeniem punktu x, przecięcie dwu otoczeń punktu x jest również jego otoczeniem. Możemy więc przystąpić do badania różniczkowalności złożenia. Z definicji (w jej drugiej wersji) istnieją odwzorowania α : (otoczenie 0 w $\mathbb{R}^d$) $\to \mathbb{R}^k$ oraz β : (otoczenie 0 w $\mathbb{R}^k$) $\to \mathbb{R}^m$ -odpowiednio, zmierzajace w tych punktach do zera, dla których

$$f(x+h) = f(x) + d_x f(h) + \|h\|\alpha(h), \quad gdzie \quad \|\alpha(h)\| \to 0 \quad przy\ h \to 0$$
$$g(y+k) = g(y) + d_y g(k) + \|k\|\beta(k), \quad gdzie \quad \|\beta(k)\| \to 0 \quad przy\ k \to 0 \quad (2)$$

Aby poskładać te wzory, przyjmijmy $k := f(x+h) - f(x)$ -wyrażone przez pierwszy z tych wzorów i niech $y = f(x)$. Z ciągłości w punktach, w których f jest różniczkowalna widzimy, że $k \to 0$ gdy $h \to 0$. Takie k możemy więc wstawić do drugiego wzoru (2), pamiętając, że $y+k = f(x)+f(x+h)-f(x) = f(x+h)$, a stąd wynika, że $g(y+k) = (g \circ f)(x+h)$.Przydatne będzie też ograniczenie ilorazu $\frac{\|k\|}{\|h\|}$. Z warunku ciągłości operatora liniowego $d_x f$ wiemy (por. wzór (1) z poprzedniego wykładu). że dla pewnej stalej M mamy oszacowania

$$\|(d_x f)(h)\| \leqslant M\|h\|$$

i wykorzystując teraz nierówność trójkąta i jednorodność dla normy widzimy, że

$$\frac{\|k\|}{\|h\|} \leqslant \frac{\|d_x f(h)\|}{\|h\|} + \frac{\big\|\|h\|\alpha(h)\big\|}{\|h\|} \leqslant M + \|\alpha(h)\|. \quad (3)$$

Możemy też wykorzystać liniowość $d_y g$, dzięki czemu

$$d_y g(k) = d_y g(d_x f(h) + \|h\|\alpha(h)) = \big((d_y g) \circ (d_x f)\big)(h) + \|h\| d_y g(\alpha(h)).$$

Tak więc, po podstawieniach, drugi ze wzorów (2) przyjmie postać

$$(g \circ f)(x+h) = (g \circ f)(x) + \big((d_y g) \circ (d_x f)\big)(h) + \|h\| d_y g(\alpha(h)) + \big\|k\big\|\beta(k) \quad (4)$$

Funkcja $h \mapsto \|\alpha(h)\|$ zmierzająca do zera gdy $h \to 0$ –jest w pewnym otoczeniu zera ograniczona. Zastępując w ostatnim wzorze $\|k\|$ przez $\|h\|\frac{\|k\|}{\|h\|}$, możemy więc skorzystać dzięki nierówności (3) z ograniczoności jej lewej strony w pewnym otoczeniu zera, zapisując wzór (4) w postaci

$$(g \circ f)(x+h) = (g \circ f)(x) + \big((d_y g) \circ (d_x f)\big)(h) + \|h\|\big\{d_y g(\alpha(h)) + \frac{\|k\|}{\|h\|}\beta(k)\big\},$$

w którym wyrażenie w nawiasie { } zmierza do zera przy $\|h\| \to 0$. Dowodzi to naszej tezy. □

Zauważmy, że wynika stąd, w szczególności, zaanonsowany przed tygodniem wzór na pochodną ϕ' z funkcji postaci $t \mapsto \phi(t) = f(x(t), y(t))$. Faktycznie, jedno-elementowa baza kanoniczna przestrzeni wektorowej $\mathbb{R}$, to liczba 1. Dla wygody oznaczmy przez P_0 punkt przestrzeni $\mathbb{R}^2$ o współrzędnych $(x(t_0), y(t_0))$. Różniczka odwzorowania tak złożonego ϕ w punkcie t_0, to odwzorowanie liniowe $\mathbb{R} \ni s \to \phi'(t_0)s$, więc z różniczki odzyskamy pochodną licząc wartość $d_{t_0}\phi(1)$. Odwzorowanie wewnętrzne przypisuje liczbie t wektor $(x(t), y(t)) \in \mathbb{R}^2$ i jego różniczka w punkcie t_0, to odwzorowanie liniowe: $\mathbb{R} \ni s \to (sx'(t_0), sy'(t_0))$ Z kolei, różniczka f w jednym z zapisów, to $\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$, gdzie dx, dy - to operatory liniowe rzutów na oś OX, odpowiednio OY. Różniczka tego złożenia w punkcie t_0, to złożenie powyższych różniczek, czyli odwzorowanie

$$\mathbb{R} \ni s \to \frac{\partial f}{\partial x}(P_0)x'(t)s + \frac{\partial f}{\partial y}(P_0)y'(t_0)s.$$

Wstawiając teraz $s = 1$ otrzymamy wzór na pochodną $\phi'(t_0)$. Wcześniej uzyskaliśmy go inną metodą -ale przy silniejszych założeniach o ciągłości pochodnych cząstkowych. Obecnie wystarczy założyć różniczkowalności: funkcji x, y -w punkcie t_0 oraz funkcji f w punkcie P_0.

Analogiczną metodę możemy stosować do liczenia pochodnych cząstkowych. Mając różniczkę $d_P f$ funkcji wielu zmiennych f w punkcie P otrzymujemy dla wektora e_k z kanonicznej bazy 0-1 kowej, że

$$\frac{\partial f}{\partial x_k}(P) = d_P f(e_k).$$

Dla funkcji f o wartościach skalarnych otrzymamy następujące twierdzenie o wartości średniej: Oznaczmy przez $[P_0 : P_1]$ odcinek w przestrzeni $\mathbb{R}^d$ o końcach w punktach P_0, P_1. Naturalną parametryzacją dla tego odcinka jest odwzorowanie afiniczne $[0,1] \ni t \to P_t := P_0 + t(P_1 - P_0)$. Punkty postaci P_t, gdzie $0 < t < 1$ będziemy nazywali punktami wewnętrznymi tego odcinka. (Odwzorowanie to określone jest, oczywiśce, nawet dla dowolnych $t \in \mathbb{R}$.) Różniczką jest tu część liniowa tego odwzorowana, czyli odwzorowanie $\mathbb{R} \ni t \to t(P_1 - P_0)$. Jego wartość w punkcie $t = 1$, to pochodna tego odwzorowania, jet nią (w każdym punkcie) wektor (stały) $P_1 - P_0$. Stosując więc regułę łańcucha do funkcji złożonej $t \mapsto f(P_t)$, dzięki Twierdzeniu Lagrange'a otrzymujemy jego wersję dla funkcji wielu miennych:

Twierdzenie (o wartości średniej). Jeżeli funkcja $f : D \subset \mathbb{R}^d \to \mathbb{R}$ jest różniczkowalna w punktch wewnętrznych odcinka $[P_0 : P_1]$ i ciągła na tym odcinku, to dla pewnego $\theta \in (0,1)$ jej przyrost wyraża się wzorem:

$$f(P_1) - f(P_0) = d_{P_\theta} f(P_1 - P_0).$$

Gdy wektor $P_1 - P_0$ jest wektorem $h = (h_1, h_2, \ldots, h_d)$, to przyrost ten możemy wyrazić przy użyciu pochodnych cząstkowych: $f(P_1) - f(P_0) = \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}(P_\theta) h_j$.

Podobnie, jak w przypadku odwzorowań $\gamma : [a, b] \to \mathbb{R}^k$, analogiczna teza przestaje być prawdziwa dla odwzorowań o wartościach wektorowych (dla $k > 1$). Mamy wówczas tylko możliwość oszacowania normy przyrostu. Dowód (analogiczny, jak w przypadku odwzorowań zależnych od 1 zmiennej -polegający na pomnożeniu skalarnym F przez wektor w będący wektorem przyrostu podzielonym przez jego normę) -pomińmy. Samo sformułowanie warto jednak odnotować:

Twierdzenie (o przyrostach). Jeżeli odwzorowanie $F : D \subset \mathbb{R}^d \to \mathbb{R}^k$ jest różniczkowalne w punktch wewnętrznych odcinka $[P_0 : P_1]$ i ciągłe na tym odcinku, to

$$\|F(P_1) - F(P_0)\| \leqslant \|P_1 - P_0\| \sup_{t \in (0,1)} \|d_{P_t}\|.$$

Tu norma z różniczki, to norma odwzorowania liniowego. Przypomnnijmy, że dla $T : \mathbb{R}^d \to \mathbb{R}^k$ taka norma, $\|T\|$, to najmniejsza stała M, dla której na każdym z wektorów $v \in \mathbb{R}^d$ zachodzą nierówności: $\|T(v)\| \leqslant M\|v\|$. Ponadto $\|T\| = \sup\{\|T(w)\| : \|w\| = 1\}$. W przypadku macierzy symetrycznych $d \times d$ taka norma (zwana normą spektralną), to maksimum z modułów jej wartości własnych. Zbiór wartości własnych -to widmo (spectrum) macierzy.

Innym zastosowaniem reguły łańcucha jest wzór dla tzw. pochodnej kierunkowej. W niektórych źródłach przyjmuje się jej definicję "obustronną" =a w innych "jednostronną" -która jest nieco bardziej "elastyczna" i z niej skorzystamy.

Definicja. Jeżeli $w \in \mathbb{R}^d$ jest wektorem niezerowym, $f : D \subset \mathbb{R}^d \to \mathbb{R}$ jest pewną funkcją oraz $P_0 \in D$ jest takim punktem, że zbiór $\{t \in \mathbb{R} : P_0 + tw \in D\}$ jest prawostronnym otoczeniem zera, to przez pochodną $\partial^w f(P_0)$ w kierunku wektora w w punkcie P_0 rozumiemy granicę

$$\partial^w f(P_0) := \lim_{h \to 0^+} \frac{f(P_0 + hw) - f(P_0)}{h}.$$

Ściślej- można mówić o pochodnej kierunkowej jednostronnej- stronę wyznacza to zwrot wektora, bo kierunek - to cała prosta $\{P_0 + tw : t \in \mathbb{R}\}$. W przypadku $P_0 = 0$ odwzorowanie $P \mapsto \|P\|$ ma w każdym kierunku pochodną 1-stronną równą 1, ale brak tu granicy obustronnej. Jest to więc pochodna 1-stronna w punkcie $t = 0$ z funkcji złożonej $f(P_0 + tw)$ i dzięki regule łańcucha (pochodną odzyskujemy z różniczki licząc jej wartość na 1-wymiarowym wektorze 1), mamy następujący

Wniosek. Funkcja różniczkowalna w danym punkcie ma również pochodne kierunkowe w kierunku każdego wektora $w \in \mathbb{R}^d$, przy czym

$$\partial^w f(P_0) = d_{P_0} f(w).$$

W szczególności, gdy pochodne kierunkowe istnieją, ale nie zależą w sposób liniowy od wektora kierunku, to dana funkcja nie jest różniczkowalna. Następujący przykład pokazuje, że z istnienia pochodnych kierunkwych $\partial^w f$ zależnych w sposób liniowy od w jeszcze nie wynika różniczkowalność f. Niech mianowicie

$$f(x,y) = \begin{cases} \frac{x^3 y}{x^4+y^2}, & \text{gdy } x^2+y^2>0 \\ 0, & \text{w przeciwnym przypadku.} \end{cases}$$

Wówczas dla wektora $w = (s,t)$ i dla $h > 0$ mamy $hw = (hs, ht)$ oraz $\frac{1}{h}(f(0+hw) - f(0)) = \frac{1}{h}\frac{(hs)^3 ht}{h^2(h^2s^4+t^2)}$. To wyrażenie w przypadku $t = 0, s \neq 0$ jest stale równe 0 , a gdy $t \neq 0$ -zmierza do zera, więc $\partial^w f(0,0) = 0$ dla każdego wektora w. Jednak różniczkowalności w zerze nie mamy. Pochodne w kierunkach osi, to pochodne cząstkowe i są one równe zero. Gdyby zachodziła różniczkowalność, mielibyśmy zmierzanie do zera przy $(x,y) \to (0,0)$ ułamka, którego licznikiem jest $f(x,y) - f(0,0) - d_0 f(x,y) = f(x,y)$, a mianownikiem $\sqrt{x^2+y^2}$ Ze względu na wyższe potęgi, szybciej do zera zmierza mianownik ze wzoru na $f(x,y)$ i to ten mianownik spóbujmy "ujednorodnić" -czyli wyrównać potęgi jego składników. Podstawmy $y = x^2$ Zerem powinna być więc granica z wyrażenia

$$\frac{x^3x^2}{\sqrt{x^2+x^4}(2x^4)} = \frac{x^5}{2|x|x^4\sqrt{1+x^2}} = \frac{signum(x)}{2\sqrt{1+x^2}}.$$

Ale dla $x > 0$ przy $x \to 0^+$ to wyrażenie zmierza do $\frac{1}{2}$, zamiast do zera.

Jednym z wniosków z twierdzenia o przyrostach jest uogólnienie zasady, że funkcja jednej zmiennej, której pochodna jest równa zero na prewnym przedziale -musi tam być stała. Problem geometryczny polega na tym, że dziedziną f jest teraz jakiś zbiór otwarty D.Co należy założyć o D, by z relacji $\forall_{P \in D}\ d_p f = 0$ wynikało, że f jest stała na zbiorze (otwartym) D? Dzięki twierdzeniu o przyrostach wiemy, że f ma zerowe przyrosty na odcinkach zawartych w D (więc jest na nich stała). Jeśli założyć wypukłość zbioru D, to otrzymamy stałość f na całym zbiorze. Słabszym założeniem będzie istnienie punktu $P_* \in D$ o tej własności, że odcinki łączące dowolne punkty $P \in D$ z tym punktem P_* są zawarte w D. Mówimy wtedy, że D jest obszarem gwiaździstym.

Jeszcze słabsze - a w przypadku zbiorów otwartych D -najsłabsze z możliwych jest założenie spójności. W przypadku zbiorów otwartych w $\mathbb{R}^d$ to jest pojęcie równoważne temu, że każde dwa punkty zbioru D można połączyć krzywą w całości zawartą w zbiorze D. Zamiast krzywej -wystarczy użyć linii łamanej. Topologiczna definicja spójności (tu w charakterze ciekawostki) jest następująca: Każda funkcja ciągła na zbiorze D przyjmująca jedynie wartości 0 lub 1 -musi być stała. Inna równoważna definicja: nie istnieją zbiory otwarte U, W w przestrzeni $\mathbb{R}^d$ takie, dla których $U_1 := U \cap D$ oraz $W_1 := W \cap D$ są niepuste, rozłączne, o sumie równej D. Np. jedynymi spójnymi podzbiorami prostej $\mathbb{R}$ są odcinki (z końcami lub bez), półproste i cała $\mathbb{R}$. Własność Bolzano-Darboux ma następujące uogólnienie: "Obraz zbioru spójnego przez odwzorowanie ciągłe jest zbiorem spójnym".