

Analiza

Funkcje wielu zmiennych, cz. 2

Wojciech Domitrz

Slajdy: Ewa Stróżyna, Wojciech Domitrz

Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska

2026

Pochodna cząstkowa: definicja

Niech $\Omega \subset \mathbb{R}^n$ będzie zbiorem otwartym, $a \in \Omega$, a $f : \Omega \rightarrow \mathbb{R}$. Oznaczmy przez e_i wektor, którego i -ta współrzędna wynosi 1, a pozostałe są równe 0.

Pochodna względem i -tej zmiennej

Jeśli istnieje skończona granica

$$f_{x_i}(a) = \frac{\partial f}{\partial x_i}(a) = \lim_{h \rightarrow 0} \frac{f(a + he_i) - f(a)}{h},$$

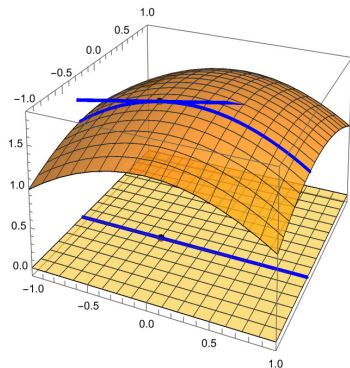
to nazywamy ją pochodną cząstkową f względem x_i w punkcie a .

Zmienia się wyłącznie i -ta współrzędna:

$$a + he_i = (a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n).$$

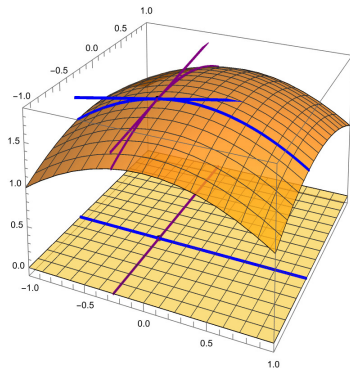
Otwartość dziedziny zapewnia, że te punkty należą do niej dla wszystkich dostatecznie małych $|h|$.

Pochodna cząstkowa: przekrój wykresu



Dla funkcji $z = f(x, y)$ ustalamy $y = b$ i badamy funkcję jednej zmiennej $x \mapsto f(x, b)$. Liczba $f_x(a, b)$ określa nachylenie stycznej do tego przekroju w punkcie $(a, b, f(a, b))$.

Dwa przekroje współrzędnościowe



Pochodne $f_x(a, b)$ i $f_y(a, b)$ opisują nachylenia stycznych w dwóch przekrojach: przy ustalonym $y = b$ oraz przy ustalonym $x = a$. Samo istnienie obu pochodnych nie zapewnia różniczkowalności.

Obliczanie pochodnych cząstkowych

Różniczkując względem jednej zmiennej, pozostałe traktujemy jako stałe. Stosujemy znane reguły sumy, iloczynu, ilorazu i złożenia.

Dla $f(x, y, z) = x^2 y^3 + e^{xy^2} + x \sin z$, określonej na $\mathbb{R}^3$, otrzymujemy

$$f_x = 2xy^3 + y^2 e^{xy^2} + \sin z,$$

$$f_y = 3x^2 y^2 + 2xy e^{xy^2},$$

$$f_z = x \cos z.$$

W pochodnych składnika wykładniczego występują czynniki

$$\frac{\partial(xy^2)}{\partial x} = y^2, \quad \frac{\partial(xy^2)}{\partial y} = 2xy.$$

Składniki niezależne od aktualnej zmiennej mają pochodną równą zero.

Przykład: suma ilorazów

Niech

$$f(x, y, z) = \frac{x}{y} + \frac{y}{z} - \frac{z}{x}, \quad xyz \neq 0.$$

Liczymy po kolei:

$$f_x = \frac{1}{y} + 0 + \frac{z}{x^2},$$

$$f_y = -\frac{x}{y^2} + \frac{1}{z} + 0,$$

$$f_z = 0 - \frac{y}{z^2} - \frac{1}{x}.$$

Na przykład przy różniczkowaniu $-z/x = -zx^{-1}$ względem x liczba z jest stałą, więc otrzymujemy zx^{-2} .

Warunki $x \neq 0$, $y \neq 0$, $z \neq 0$ wynikają już z dziedziny funkcji i obowiązują również w powyższych wzorach.

Przykład: funkcja wykładnicza z ilorazem

Dla

$$f(x, y) = e^{y/x}, \quad x \neq 0,$$

stosujemy regułę łańcuchową. Najpierw różniczkujemy wykładnik:

$$\frac{\partial}{\partial x} \frac{y}{x} = -\frac{y}{x^2}, \quad \frac{\partial}{\partial y} \frac{y}{x} = \frac{1}{x}.$$

Stąd

$$f_x = -\frac{y}{x^2} e^{y/x}, \quad f_y = \frac{1}{x} e^{y/x}.$$

Zmiennej y nie różniczkujemy w pierwszym rachunku, a zmiennej x w drugim.

Przykład: tożsamość dla funkcji logarytmicznej

Niech $z(x, y) = x \ln(y/x)$. Dziedziną jest $xy > 0$: zmienne mogą być obie dodatnie albo obie ujemne.

Bez rozdzielania logarytmu na niedozwolone składniki liczymy

$$z_x = \ln \frac{y}{x} + x \cdot \frac{x}{y} \left(-\frac{y}{x^2} \right) = \ln \frac{y}{x} - 1,$$

$$z_y = x \cdot \frac{x}{y} \cdot \frac{1}{x} = \frac{x}{y}.$$

Wobec tego

$$xz_x + yz_y = x \left(\ln \frac{y}{x} - 1 \right) + y \frac{x}{y} = x \ln \frac{y}{x} = z.$$

Na całej dziedzinie można także użyć tożsamości $\ln(y/x) = \ln |y| - \ln |x|$.
Zapis bez modułów wymagałby $x, y > 0$.

Pochodne cząstkowe drugiego rzędu

Jeśli funkcja f_{x_i} jest określona w otoczeniu punktu a i ma w nim pochodną względem x_j , definiujemy

$$f_{x_i x_j}(a) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Konwencja kolejności: w indeksie $f_{x_i x_j}$ najpierw różniczkujemy po x_i , a następnie po x_j . W zapisie operatorowym jako pierwszy działa operator stojący najbliżej funkcji.

Dla dwóch zmiennych otrzymujemy cztery pochodne:

$$f_{xx}, \quad f_{xy}, \quad f_{yx}, \quad f_{yy}.$$

Pochodne f_{xx}, f_{yy} są czyste, a f_{xy}, f_{yx} mieszane. Równości pochodnych mieszanych nie zakładamy bez uzasadnienia.

Pochodne wyższych rzędów

Pochodna k -tego rzędu powstaje po k kolejnych różniczkowaniach:

$$f_{x_{i_1} \dots x_{i_k}} = \frac{\partial}{\partial x_{i_k}} \dots \frac{\partial}{\partial x_{i_1}} f, \quad i_j \in \{1, \dots, n\}.$$

Jeśli grupujemy różniczkowania w bloki o liczebnościach $j_1, \dots, j_s$, ich łączny rząd wynosi

$$j_1 + \dots + j_s = k.$$

Numery zmiennych nie są wagami w tej sumie. Na przykład dwukrotne różniczkowanie po x_3 i jednokrotne po x_1 daje pochodną trzeciego rzędu.

Przed zastosowaniem twierdzenia o zamianie kolejności trzeba zachować faktyczną kolejność wykonywanych operacji.

Klasa regularności i twierdzenie Schwarza

Niech $\Omega \subset \mathbb{R}^n$ będzie otwarty. Piszemy $f \in C^k(\Omega)$, gdy f oraz wszystkie jej pochodne cząstkowe rzędów od 1 do k istnieją i są ciągłe w Ω .

Twierdzenie Schwarza

Jeżeli $f \in C^k(\Omega)$, to kolejność różniczkowań nie zmienia wartości pochodnej rzędu co najwyżej k , o ile każda zmienna występuje tyle samo razy.

W szczególności dla $f \in C^2(\Omega)$:

$$f_{x_i x_j} = f_{x_j x_i}.$$

Dla $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha_i \geq 0$ całkowitych, rząd pochodnej oznaczamy przez

$$|\alpha| = \alpha_1 + \dots + \alpha_n, \quad D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Klasa C^∞ oznacza istnienie ciągłych pochodnych wszystkich rzędów.

Dlaczego można zamieniać pochodne mieszane?

Dla $f \in C^2$ w otoczeniu (a, b) rozważmy przyrost prostokątny

$$Q(h) = f(a + h, b + h) - f(a + h, b) - f(a, b + h) + f(a, b).$$

Dla małego $h \neq 0$ dwukrotnie stosujemy twierdzenie o wartości średniej, najpierw do zmiennej x , potem do y :

$$\begin{aligned} Q(h) &= h[f_x(a + \theta h, b + h) - f_x(a + \theta h, b)] \\ &= h^2 f_{xy}(a + \theta h, b + \eta h), \end{aligned}$$

gdzie $\theta, \eta \in (0, 1)$ mogą zależeć od h .

W odwrotnej kolejności otrzymujemy

$$Q(h) = h^2 f_{yx}(a + \tilde{\theta}h, b + \tilde{\eta}h).$$

Dzielimy przez h^2 i przechodzimy z h do zera. Ciągłość obu pochodnych daje $f_{xy}(a, b) = f_{yx}(a, b)$.

Pochodne mieszane: rola założeń

Rozważmy funkcję

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

Poza początkiem układu mianownik jest dodatni, więc funkcja jest klasy C^∞ .
Reguła ilorazu daje tam

$$f_x = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}, \quad f_y = \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2}.$$

Ponowne różniczkowanie prowadzi do wspólnej wartości

$$f_{xy} = f_{yx} = \frac{x^6 - y^6 + 9x^4y^2 - 9x^2y^4}{(x^2 + y^2)^3}, \quad (x, y) \neq (0, 0).$$

W punkcie $(0, 0)$ trzeba wrócić do definicji pochodnych.

Brak ciągłości pochodnych mieszanych

Dla $x \neq 0$ oraz $y \neq 0$ wzór z poprzedniej strony daje

$$f_{xy}(x, 0) = f_{yx}(x, 0) = 1, \quad f_{xy}(0, y) = f_{yx}(0, y) = -1.$$

W szczególności dwa ciągi zbieżne do $(0, 0)$ prowadzą do różnych wartości:

$$f_{xy}(1/k, 0) = 1, \quad f_{xy}(0, 1/k) = -1.$$

Zatem ani f_{xy} , ani f_{yx} nie ma granicy w początku układu. Nie mogą być tam ciągłe, niezależnie od swoich wartości w tym punkcie.

Założenie C^2 twierdzenia Schwarz'a nie jest spełnione. Sama ta obserwacja jeszcze nie oblicza pochodnych mieszanych w zerze; wykonamy ten rachunek bezpośrednio.

Obliczenie pochodnych mieszanych w zerze

Ponieważ $f(h, 0) = f(0, h) = f(0, 0) = 0$, mamy

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0)}{h} = 0, \quad f_y(0, 0) = 0.$$

Dla ustalonego $y \neq 0$:

$$f_x(0, y) = \lim_{h \rightarrow 0} \frac{f(h, y) - f(0, y)}{h} = \lim_{h \rightarrow 0} \frac{y(h^2 - y^2)}{h^2 + y^2} = -y.$$

Dla ustalonego $x \neq 0$ analogicznie:

$$f_y(x, 0) = \lim_{h \rightarrow 0} \frac{x(x^2 - h^2)}{x^2 + h^2} = x.$$

Wzory $f_x(0, y) = -y$ oraz $f_y(x, 0) = x$ obejmują także zero, zgodnie z pierwszym rachunkiem.

Nierówność pochodnych mieszanych w zerze

Teraz różniczkujemy funkcje otrzymane na osiach:

$$f_{xy}(0,0) = \lim_{h \rightarrow 0} \frac{f_x(0,h) - f_x(0,0)}{h} = \lim_{h \rightarrow 0} \frac{-h}{h} = -1,$$

$$f_{yx}(0,0) = \lim_{h \rightarrow 0} \frac{f_y(h,0) - f_y(0,0)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1.$$

$$\boxed{f_{xy}(0,0) = -1 \neq 1 = f_{yx}(0,0).}$$

Obie pochodne istnieją, lecz nie są równe. Nie przeczy to twierdzeniu Schwarza: brakuje jego założeń dotyczących ciągłości.

W tym przykładzie pochodne mieszane są równe w każdym punkcie różnym od $(0,0)$, ale nie w samym początku układu.

Drugie pochodne: wielomian

Dla $z(x, y) = x^3 - 2x^2y + 3y^2$ mamy

$$z_x = 3x^2 - 4xy, \quad z_y = -2x^2 + 6y.$$

Różniczkując każdą z tych funkcji po obu zmiennych, otrzymujemy

$$\begin{aligned} z_{xx} &= 6x - 4y, & z_{xy} &= -4x, \\ z_{yx} &= -4x, & z_{yy} &= 6. \end{aligned}$$

Wielomian jest klasy C^∞ na $\mathbb{R}^2$, więc równość $z_{xy} = z_{yx}$ wynika również z twierdzenia Schwarz'a.

Macierz drugich pochodnych, nazywana hesjanem, ma postać

$$H_z(x, y) = \begin{pmatrix} 6x - 4y & -4x \\ -4x & 6 \end{pmatrix}.$$

Jest symetryczna.

Drugie pochodne: trzy zmienne

Niech $u(x, y, t) = e^{xyt}$. Funkcja jest klasy $C^\infty(\mathbb{R}^3)$. Jej pierwsze pochodne to

$$u_x = yt e^{xyt}, \quad u_y = xt e^{xyt}, \quad u_t = xy e^{xyt}.$$

Pochodne czyste drugiego rzędu:

$$u_{xx} = y^2 t^2 e^{xyt}, \quad u_{yy} = x^2 t^2 e^{xyt}, \quad u_{tt} = x^2 y^2 e^{xyt}.$$

W pochodnych mieszanych trzeba różniczkować oba czynniki, np.

$$u_{xy} = te^{xyt} + yt \cdot xte^{xyt} = t(1 + xyt)e^{xyt}.$$

Podobnie

$$u_{xt} = y(1 + xyt)e^{xyt}, \quad u_{yt} = x(1 + xyt)e^{xyt}.$$

Twierdzenie Schwarz'a daje $u_{yx} = u_{xy}$, $u_{tx} = u_{xt}$ i $u_{ty} = u_{yt}$.

Pochodne mieszane: funkcja cosinus

Niech $z(x, y) = \cos(ax - by)$, gdzie $a, b \in \mathbb{R}$ są stałymi. Mamy

$$z_x = -a \sin(ax - by), \quad z_y = b \sin(ax - by).$$

Stąd

$$z_{xy} = -a \cos(ax - by) \cdot (-b) = ab \cos(ax - by),$$

$$z_{yx} = b \cos(ax - by) \cdot a = ab \cos(ax - by).$$

W obu rachunkach uwzględniamy pochodną argumentu funkcji trygonometrycznej. Dwa znaki minus w pierwszym rachunku znoszą się.

Funkcja jest klasy C^∞ na $\mathbb{R}^2$, zatem zgodność wyników jest także konsekwencją twierdzenia Schwarz'a.

Pochodne mieszane: logarytm

Niech $z(x, y) = \ln(1 + x^2 + y^2)$. Argument logarytmu jest dodatni dla każdej pary (x, y) , więc dziedziną jest $\mathbb{R}^2$.

$$z_x = \frac{2x}{1 + x^2 + y^2}, \quad z_y = \frac{2y}{1 + x^2 + y^2}.$$

Przy różniczkowaniu z_x po y licznik $2x$ jest stały:

$$z_{xy} = 2x \left(-\frac{2y}{(1 + x^2 + y^2)^2} \right) = -\frac{4xy}{(1 + x^2 + y^2)^2}.$$

Analogicznie

$$z_{yx} = -\frac{4xy}{(1 + x^2 + y^2)^2} = z_{xy}.$$

Funkcja należy do $C^\infty(\mathbb{R}^2)$.

Sprawdzenie równania różniczkowego

Sprawdźmy, że $z(x, y) = 2 \cos^2(y - x/2)$ spełnia równanie

$$2 \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial x \partial y} = 0.$$

Korzystamy z $2 \cos^2 \alpha = 1 + \cos(2\alpha)$:

$$z = 1 + \cos(2y - x), \quad z_x = \sin(2y - x), \quad z_y = -2 \sin(2y - x).$$

Zatem

$$z_{xx} = -\cos(2y - x), \quad z_{yx} = 2 \cos(2y - x).$$

Po podstawieniu do równania:

$$2z_{xx} + z_{yx} = -2 \cos(2y - x) + 2 \cos(2y - x) = 0.$$

Tożsamość zachodzi na całej $\mathbb{R}^2$; również $z_{xy} = z_{yx}$.

Płaszczyzna styczna do wykresu

Niech f będzie różniczkowalna w punkcie (a, b) . Pojęcie różniczkowalności zdefiniujemy dokładnie za chwilę; oznacza ono możliwość przybliżenia przyrostu funkcji przez funkcję liniową z odpowiednio małym błędem.

Równanie płaszczyzny stycznej

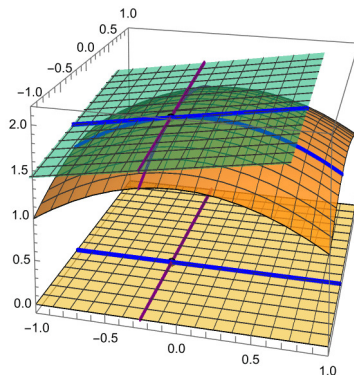
Do wykresu $z = f(x, y)$ w punkcie $P = (a, b, f(a, b))$:

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Płaszczyzna przechodzi przez P . Jej współczynniki opisują nachylenie wykresu w kierunkach obu osi.

Wzór opisuje liniowe przybliżenie całego wykresu w otoczeniu punktu. Istnienie samych dwóch pochodnych cząstkowych nie zapewnia takiego przybliżenia.

Płaszczyzna styczna: interpretacja



Styczne do dwóch przekrojów współrzędnościowych leżą w jednej płaszczyźnie. Przy różniczkowalności jest ona przybliżeniem wykresu we wszystkich kierunkach, a nie tylko wzdłuż tych przekrojów.

Wyprowadzenie równania płaszczyzny

Styczna do przekroju $y = b$ ma równania

$$y = b, \quad z = f(a, b) + f_x(a, b)(x - a),$$

a styczna do przekroju $x = a$:

$$x = a, \quad z = f(a, b) + f_y(a, b)(y - b).$$

Ich wektory kierunkowe

$$u = (1, 0, f_x(a, b)), \quad v = (0, 1, f_y(a, b))$$

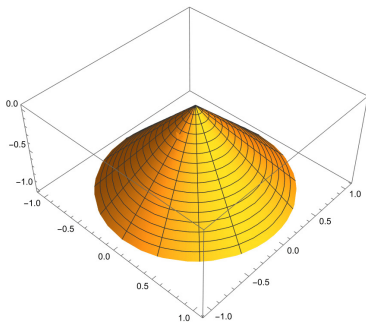
są liniowo niezależne. Wektor $N = (-f_x(a, b), -f_y(a, b), 1)$ jest prostopadły do obu, bo $N \cdot u = N \cdot v = 0$.

Równanie płaszczyzny przez P o wektorze normalnym N to

$$-f_x(a, b)(x - a) - f_y(a, b)(y - b) + z - f(a, b) = 0.$$

Różniczkowalność uzasadnia, że ta płaszczyzna jest styczna do całego wykresu.

Stożek: wierzchołek bez płaszczyzny stycznej



Dla $f(x, y) = -\sqrt{x^2 + y^2}$ przekroje przez początek układu mają ostre wierzchołki. Nie otrzymujemy tam jednej płaszczyzny stycznej przybliżającej wykres.

$$\frac{f(h, 0) - f(0, 0)}{h} = -\frac{|h|}{h} = \begin{cases} -1, & h > 0, \\ 1, & h < 0. \end{cases}$$

Pochodna $f_x(0, 0)$ nie istnieje; tak samo $f_y(0, 0)$.

Różniczkowalność: przybliżenie liniowe

Niech $f : \Omega \rightarrow \mathbb{R}$, gdzie $\Omega \subset \mathbb{R}^n$ jest otwarty, i niech $a \in \Omega$.

Definicja

Funkcja f jest różniczkowalna w a , jeśli istnieje odwzorowanie liniowe $L : \mathbb{R}^n \rightarrow \mathbb{R}$ takie, że

$$f(a + h) = f(a) + L(h) + R(h), \quad \lim_{h \rightarrow 0, h \neq 0} \frac{R(h)}{\|h\|} = 0.$$

Warunek graniczny zapisujemy także $R(h) = o(\|h\|)$. Oznacza on, że błąd jest mały w porównaniu z długością przyrostu, niezależnie od kierunku zbliżania się do zera.

Po podstawieniu $h = te_i$ i podzieleniu przez t otrzymujemy

$$f_{x_i}(a) = L(e_i), \quad i = 1, \dots, n.$$

Zatem różniczkowalność pociąga istnienie wszystkich pochodnych cząstkowych.

Różniczka i równoważny zapis definicji

Liniowość daje $L(h) = \sum_{i=1}^n L(e_i)h_i$. Dlatego

$$f(a+h) = f(a) + \sum_{i=1}^n f_{x_i}(a)h_i + o(\|h\|).$$

Odwzorowanie liniowe L jest jednoznaczne. Nazywamy je **różniczką** i oznaczamy przez $df(a)$ lub $Df(a)$:

$$df(a)[h] = \sum_{i=1}^n f_{x_i}(a)h_i.$$

Zapis $df = \sum_i f_{x_i} dx_i$ wyraża tę samą zależność: dx_i wybiera i -tą współrzędną przyrostu.

Dla dwóch zmiennych błąd przybliżenia płaszczyzną styczną wynosi

$$f(a+u, b+v) - f(a, b) - f_x(a, b)u - f_y(a, b)v = o\left(\sqrt{u^2 + v^2}\right).$$

To właśnie warunek dotyczący całego otoczenia, a nie tylko osi.

Jedna zmienna: zgodność ze znaną definicją

Dla funkcji jednej zmiennej istnienie pochodnej $f'(a)$ oznacza

$$\lim_{h \rightarrow 0} \left(\frac{f(a+h) - f(a)}{h} - f'(a) \right) = 0.$$

Po oznaczeniu $R(h) = f(a+h) - f(a) - f'(a)h$ jest to warunek

$$\frac{R(h)}{h} \rightarrow 0.$$

Ponieważ $|R(h)/h| = |R(h)|/|h|$, warunek ten jest równoważny

$$\boxed{\lim_{h \rightarrow 0} \frac{R(h)}{|h|} = 0.}$$

A więc istnienie skończonej pochodnej jednej zmiennej jest równoważne różniczkowalności, a $df(a)[h] = f'(a)h$.

W wielu wymiarach same pochodne cząstkowe badają tylko wybrane kierunki, więc nie dają takiej równoważności.

Przykład różniczkowalności: rozwinięcie przyrostu

Niech $f(x, y) = x^2 - 2y^2$ i $a = (2, 1)$. Mamy

$$f(2, 1) = 2, \quad f_x(2, 1) = 4, \quad f_y(2, 1) = -4.$$

Dla przyrostu (u, v) obliczamy dokładnie

$$\begin{aligned} f(2 + u, 1 + v) &= (2 + u)^2 - 2(1 + v)^2 \\ &= 4 + 4u + u^2 - 2 - 4v - 2v^2 \\ &= 2 + 4u - 4v + u^2 - 2v^2. \end{aligned}$$

Kandydatem na różniczkę jest $L(u, v) = 4u - 4v$, a reszta wynosi

$$R(u, v) = u^2 - 2v^2.$$

Aby wykazać różniczkowalność, trzeba jeszcze porównać tę resztę z długością przyrostu.

Przykład różniczkowalności: kontrola reszty

Niech $r = \sqrt{u^2 + v^2} > 0$. Ponieważ

$$|u^2 - 2v^2| \leq u^2 + 2v^2 \leq 2(u^2 + v^2),$$

otrzymujemy

$$0 \leq \frac{|R(u, v)|}{\sqrt{u^2 + v^2}} \leq 2r \rightarrow 0.$$

Można też zapisać

$$\frac{R(u, v)}{r} = r - 3\frac{v^2}{r}, \quad 0 \leq \frac{v^2}{r} \leq r,$$

co prowadzi do tego samego wniosku. Funkcja jest różniczkowalna w $(2, 1)$ i

$$df(2, 1) = 4 \, dx - 4 \, dy.$$

Płaszczyzna styczna ma równanie

$$z = 2 + 4(x - 2) - 4(y - 1) = 4x - 4y - 2.$$

Różniczkowalność pociąga ciągłość

Twierdzenie

Jeżeli f jest różniczkowalna w punkcie a , to jest w nim ciągła.

Dowód. Mamy $f(a + h) - f(a) = L(h) + R(h)$. Dla $h \neq 0$ oznaczmy

$$\varepsilon(h) = \frac{R(h)}{\|h\|}, \quad \varepsilon(0) = 0.$$

Wtedy $\varepsilon(h) \rightarrow 0$ oraz

$$f(a + h) - f(a) = \sum_{i=1}^n f_{x_i}(a)h_i + \varepsilon(h)\|h\|.$$

Każda współrzędna h_i dąży do zera, więc skończona suma po prawej stronie dąży do zera. Iloczyn $\varepsilon(h)\|h\|$ również dąży do zera.

Stąd $f(a + h) \rightarrow f(a)$, czyli ciągłość. W szczególności brak ciągłości w punkcie wyklucza różniczkowalność w tym punkcie.

Twierdzenie

Jeżeli wszystkie pochodne cząstkowe pierwszego rzędu istnieją w pewnym otoczeniu punktu a i są ciągłe w punkcie a , to f jest różniczkowalna w a .

W szczególności każda funkcja klasy C^1 na zbiorze otwartym jest w nim różniczkowalna.

Przykład. Dla $f(x, y) = \sin(x^2 + y^3)$:

$$f_x = 2x \cos(x^2 + y^3), \quad f_y = 3y^2 \cos(x^2 + y^3).$$

Obie pochodne są ciągłe na $\mathbb{R}^2$, więc funkcja jest wszędzie różniczkowalna. W istocie należy do $C^\infty(\mathbb{R}^2)$.

Twierdzenie wymaga istnienia pochodnych jako funkcji w otoczeniu. Same wartości pochodnych w pojedynczym punkcie nie mają własności „ciągłości w punkcie”.

Dowód warunku wystarczającego: podział przyrostu

Wyberzmy małą kulę wokół a , w której istnieją wszystkie pochodne cząstkowe. Dla małego $h = (h_1, \dots, h_n)$ wprowadzamy

$$p_0 = a, \quad p_i = a + (h_1, \dots, h_i, 0, \dots, 0), \quad p_n = a + h.$$

Suma teleskopowa daje

$$f(a + h) - f(a) = \sum_{i=1}^n (f(p_i) - f(p_{i-1})).$$

Na odcinku od p_{i-1} do p_i zmienia się tylko i -ta współrzędna. Z jednowymiarowego twierdzenia o wartości średniej:

$$f(p_i) - f(p_{i-1}) = f_{x_i}(\xi_i)h_i,$$

gdzie ξ_i leży na tym odcinku. Jeśli $h_i = 0$, oba wyrazy są równe, a ξ_i można wybrać jako p_i .

Wszystkie odcinki mieszczą się w wybranej kuli i $\|\xi_i - a\| \leq \|h\|$.

Dowód warunku wystarczającego: oszacowanie

Po odjęciu części liniowej pozostaje

$$R(h) = \sum_{i=1}^n (f_{x_i}(\xi_i) - f_{x_i}(a)) h_i.$$

Stąd, z nierówności $\sum_i |h_i| \leq \sqrt{n} \|h\|$,

$$\frac{|R(h)|}{\|h\|} \leq \sqrt{n} \max_{1 \leq i \leq n} |f_{x_i}(\xi_i) - f_{x_i}(a)|.$$

Gdy $h \rightarrow 0$, wszystkie punkty ξ_i dążą do a . Ciągłość każdej pochodnej w a sprawia, że każde wyrażenie pod maksimum dąży do zera.

Jest ich skończenie wiele, więc maksimum także dąży do zera. Otrzymujemy $R(h) = o(\|h\|)$, co kończy dowód.

W dowodzie wystarcza ciągłość pochodnych w samym a ; ich istnienie jest potrzebne w całym małym otoczeniu, aby zastosować twierdzenie o wartości średniej.

Różniczkowalność bez ciągłości pochodnych

Rozważmy

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

Z definicji, korzystając z $|\sin t| \leq 1$:

$$f_x(0, 0) = \lim_{h \rightarrow 0} h \sin(1/h^2) = 0, \quad f_y(0, 0) = 0.$$

Kandydatem na różniczkę w zerze jest odwzorowanie zerowe. Dla $r = \sqrt{x^2 + y^2} > 0$:

$$\frac{|f(x, y) - f(0, 0)|}{r} = r |\sin(1/r^2)| \leq r \longrightarrow 0.$$

Funkcja jest więc różniczkowalna w $(0, 0)$ i $df(0, 0) = 0$.

Pochodne mogą być nawet nieograniczone

Poza zerem oznaczmy $s = x^2 + y^2$. Reguła iloczynu i reguła łańcuchowa dają

$$\begin{aligned}f_x &= 2x \sin(1/s) - \frac{2x}{s} \cos(1/s), \\f_y &= 2y \sin(1/s) - \frac{2y}{s} \cos(1/s).\end{aligned}$$

Dla $x_k = 1/\sqrt{2k\pi}$ oraz $y_k = 0$ mamy

$$1/s_k = 2k\pi, \quad \sin(1/s_k) = 0, \quad \cos(1/s_k) = 1,$$

więc

$$f_x(x_k, 0) = -\frac{2}{x_k} = -2\sqrt{2k\pi} \longrightarrow -\infty.$$

Pochodna f_x nie jest ciągła w zerze. Analogiczny ciąg na osi y pokazuje brak ciągłości f_y .

Ciągłość pochodnych jest zatem warunkiem wystarczającym różniczkowalności, lecz nie jest warunkiem koniecznym.

Gradient i pochodna kierunkowa

Jeśli istnieją wszystkie pochodne cząstkowe w a , ich wektor nazywamy **gradientem**. Przyjmujemy zapis kolumnowy:

$$\nabla f(a) = \text{grad } f(a) = (f_{x_1}(a), \dots, f_{x_n}(a))^T.$$

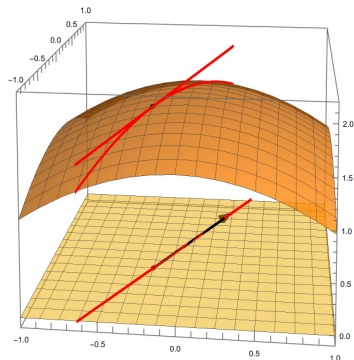
Dla $v \in \mathbb{R}^n$ **pochodną w kierunku wektora** v definiujemy, o ile istnieje skończona granica, przez

$$D_v f(a) = \frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}.$$

Zakładamy, że f jest określona w otoczeniu a . Dla $v = 0$ pochodna wynosi 0.

W tej definicji wektor nie musi być jednostkowy. Gdy $\|v\| = 1$, parametr t mierzy przemieszczenie z jednostkową prędkością. Długość wektora wpływa na wartość pochodnej.

Pochodna kierunkowa: przekrój wzdłuż prostej



W dziedzinie poruszamy się po prostej $a + tv$. Odpowiada jej krzywa na wykresie, a pochodna $D_v f(a)$ określa tempo zmiany wysokości względem parametru t .

Dla wektorów bazowych $D_{e_i} f(a) = f_{x_i}(a)$.

Iloczyn skalarny i prostopadłość

Dla $u, v \in \mathbb{R}^n$ iloczyn skalarny definiujemy wzorem

$$u \cdot v = u^T v = \sum_{i=1}^n u_i v_i.$$

Jest on powiązany z normą: $u \cdot u = \|u\|^2$. Dwa niezerowe wektory są prostopadłe wtedy i tylko wtedy, gdy $u \cdot v = 0$.

Nierówność Cauchy'ego–Schwarza mówi, że

$$|u \cdot v| \leq \|u\| \|v\|.$$

Dla niezerowych wektorów i kąta α między nimi:

$$u \cdot v = \|u\| \|v\| \cos \alpha.$$

Iloczyn skalarny jest największy przy zgodnych zwrotach, jeśli długości wektorów są ustalone. Warunek ustalenia długości będzie istotny przy badaniu największej pochodnej kierunkowej.

Twierdzenie

Jeżeli f jest różniczkowalna w a , to dla każdego $v \in \mathbb{R}^n$ istnieje pochodna kierunkowa i

$$D_v f(a) = \nabla f(a) \cdot v = \sum_{i=1}^n f_{x_i}(a) v_i.$$

Inaczej: $D_v f(a) = df(a)[v]$. W szczególności pochodne cząstkowe są pochodnymi w kierunkach wektorów bazowych.

Dla $v \neq 0$ i $e = v/\|v\|$ mamy

$$D_v f(a) = \|v\| D_e f(a).$$

Wzór z gradientem wynika z różniczkowalności. Samo istnienie pochodnych cząstkowych, a nawet wszystkich pochodnych kierunkowych, nie wystarcza do jego ogólnego zastosowania.

Dowód wzoru na pochodną kierunkową

Z różniczkowalności f w a :

$$f(a+h) = f(a) + df(a)[h] + R(h), \quad R(h) = o(\|h\|).$$

Dla $v \neq 0$ podstawiamy $h = tv$, $t \neq 0$, i dzielimy przez t :

$$\frac{f(a+tv) - f(a)}{t} = df(a)[v] + \frac{R(tv)}{t}.$$

Ponieważ $\|tv\| = |t| \|v\|$, resztę szacujemy następująco:

$$\left| \frac{R(tv)}{t} \right| = \frac{|R(tv)|}{\|tv\|} \|v\| \longrightarrow 0.$$

Granica ilorazu różnicowego wynosi więc $df(a)[v] = \nabla f(a) \cdot v$. Oszacowanie działa zarówno dla $t > 0$, jak i dla $t < 0$.

Dla $v = 0$ licznik jest identycznie zerowy, a prawa strona także wynosi 0.

Kierunek największego wzrostu i spadku

Niech f będzie różniczkowalna w a i niech $g = \nabla f(a) \neq 0$. Rozpatrujemy wyłącznie wektory jednostkowe v .

$$-\|g\| \leq D_v f(a) = g \cdot v \leq \|g\|.$$

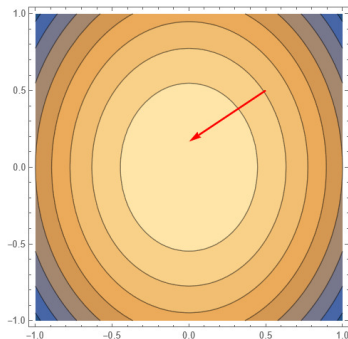
$$\begin{aligned} \max_{\|v\|=1} D_v f(a) &= \|g\|, & v_{\max} &= \frac{g}{\|g\|}, \\ \min_{\|v\|=1} D_v f(a) &= -\|g\|, & v_{\min} &= -\frac{g}{\|g\|}. \end{aligned}$$

Gradient wskazuje kierunek największego wzrostu w pierwszym przybliżeniu. Przeciwny zwrot daje największy spadek.

Jeśli $g = 0$, wszystkie pochodne kierunkowe są równe zero; nie wyróżniają wtedy kierunku wzrostu. Nie rozstrzyga to jeszcze o istnieniu ekstremum.

Bez ograniczenia $\|v\| = 1$ i przy $g \neq 0$ wartości $D_v f(a)$ nie mają maksimum: można dowolnie zwiększać długość v .

Gradient a poziomice



Poziomica ma równanie $f(x, y) = c$. W punkcie regularnym, w którym $\nabla f \neq 0$, gradient jest prostopadły do kierunku stycznej do poziomicy i wskazuje stronę wzrostu wartości funkcji.

Dla funkcji wielu zmiennych analogiczna własność dotyczy zbiorów poziomicowych $\{x : f(x) = c\}$.

Największa pochodna: dowód w dwóch wymiarach

Niech $\nabla f(a) = (A, B)^T$ i $R = \sqrt{A^2 + B^2} > 0$. Każdy wektor jednostkowy w płaszczyźnie ma postać $v(t) = (\cos t, \sin t)$.

$$D_{v(t)}f(a) = A \cos t + B \sin t.$$

Wyberzmy kąt φ tak, aby $A = R \cos \varphi$ i $B = R \sin \varphi$. Wtedy

$$A \cos t + B \sin t = R \cos(t - \varphi).$$

Maksimum wynosi R i występuje dla $t = \varphi$ modulo 2π . Minimum $-R$ odpowiada $t = \varphi + \pi$ modulo 2π .

$$v_{\max} = (\cos \varphi, \sin \varphi) = \frac{(A, B)}{R}.$$

Dowód nie wymaga dzielenia przez A ani B , więc obejmuje także gradient równoległy do jednej z osi. Gdy $R = 0$, rozpatrywana funkcja jest identycznie zerowa.

Prostopadłość gradientu: dowód przez sieczne

Niech f będzie różniczkowalna w a , a $q_k \neq a$ niech spełniają

$$q_k \rightarrow a, \quad f(q_k) = f(a).$$

Z przybliżenia liniowego otrzymujemy

$$0 = \nabla f(a) \cdot (q_k - a) + R(q_k - a).$$

Dzieląc przez $\|q_k - a\|$, mamy

$$\nabla f(a) \cdot \frac{q_k - a}{\|q_k - a\|} = -\frac{R(q_k - a)}{\|q_k - a\|} \rightarrow 0.$$

Jeżeli kierunki siecznych $u_k = (q_k - a)/\|q_k - a\|$ zbiegają do u , to $\nabla f(a) \cdot u = 0$.

Nie zakładamy, że każdy taki ciąg kierunków zbiega. Wniosek dotyczy każdej istniejącej granicy, w tym granicy dowolnego zbieżnego podciągu.

Prostopadłość gradientu: krzywa na poziomicy

Niech różniczkowalna krzywa γ spełnia

$$\gamma(0) = a, \quad f(\gamma(t)) = f(a), \quad \tau = \gamma'(0).$$

Mamy $\gamma(t) - a = t\tau + o(|t|)$, a z różniczkowalności f :

$$0 = f(\gamma(t)) - f(a) = t \nabla f(a) \cdot \tau + o(|t|).$$

Po podzieleniu przez $t \neq 0$ i przejściu do granicy:

$$\nabla f(a) \cdot \gamma'(0) = 0.$$

Zatem gradient jest prostopadły do każdego niezerowego wektora stycznego takiej krzywej.

Jeśli dodatkowo $f \in C^1$ w otoczeniu a i $\nabla f(a) \neq 0$, twierdzenie o funkcji uwikłanej zapewnia lokalnie gładką poziomice (w $\mathbb{R}^n$: powierzchnię wymiaru $n - 1$). Przy gradiencie zerowym równanie prostopadłości nie wyznacza stycznej.

Wszystkie pochodne kierunkowe a ciągłość

Rozważmy

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2 + y^6}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

Dla $p_k = (1/k^3, 1/k)$ mamy

$$f(p_k) = \frac{1/k^6}{1/k^6 + 1/k^6} = \frac{1}{2}.$$

Natomiast dla $q_k = (0, 1/k)$ zachodzi $f(q_k) = 0$. Oba ciągi dążą do $(0, 0)$, więc funkcja nie ma tam granicy i nie jest ciągła.

Z twierdzenia o ciągłości funkcji różniczkowalnej wynika, że f nie jest różniczkowalna w zerze. Mimo to wszystkie jej pochodne kierunkowe w tym punkcie istnieją.

Obliczenie wszystkich pochodnych kierunkowych

Ustalmy $w = (u, v)$. Jeśli $u \neq 0$, to dla $t \neq 0$:

$$\begin{aligned}\frac{f(tu, tv) - f(0, 0)}{t} &= \frac{1}{t} \frac{t^4 uv^3}{t^2 u^2 + t^6 v^6} \\ &= \frac{tuv^3}{u^2 + t^4 v^6} \rightarrow 0.\end{aligned}$$

Mianownik dąży do dodatniej liczby u^2 , a licznik do zera.

Jeśli $u = 0$, to $f(0, tv) = 0$ dla wszystkich t , również wtedy, gdy $v = 0$. Iloraz różnicowy jest zatem identycznie zerowy.

$$D_w f(0, 0) = 0 \quad \text{dla każdego } w \in \mathbb{R}^2.$$

Każdą ustaloną prostą badamy osobno. Brakuje kontroli jednoczesnego zbliżania się po wszystkich możliwych drogach, wymaganej w definicji różniczkowalności.

Dlaczego przybliżenie liniowe nie działa?

W szczególności $f_x(0,0) = f_y(0,0) = 0$, więc jedynym kandydatem na różniczkę jest $L = 0$. Warunek różniczkowalności wymagałby

$$\frac{f(x,y)}{\sqrt{x^2 + y^2}} \longrightarrow 0.$$

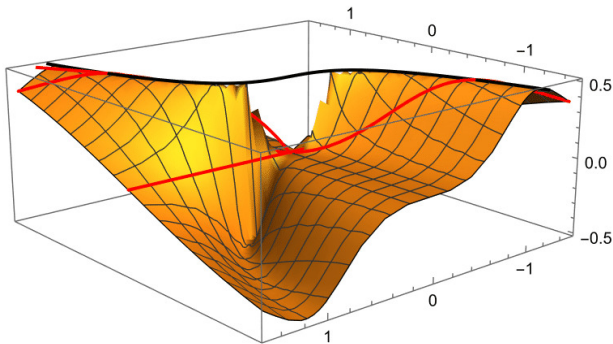
Jednak na ciągu $p_k = (1/k^3, 1/k)$ otrzymujemy

$$\frac{f(p_k)}{\|p_k\|} = \frac{1/2}{\sqrt{k^{-6} + k^{-2}}} = \frac{k}{2\sqrt{1 + k^{-4}}} \longrightarrow +\infty.$$

Błąd nie jest mały względem długości przyrostu. Zerowe pochodne w każdym ustalonym kierunku nie zapewniają warunku $R(h) = o(\|h\|)$.

Ten przykład pokazuje także, że istnienie wszystkich pochodnych kierunkowych nie pociąga nawet ciągłości.

Wykres kontrprzykładu



Na krzywej $x = y^3$, poza zerem, funkcja $xy^3/(x^2 + y^6)$ ma wartość $1/2$. Na osi y ma wartość 0 .

Rysunek ilustruje różnicę między zachowaniem na ustalonych prostych a zachowaniem w pełnym otoczeniu początku układu. Wartość w samym $(0, 0)$ wynosi 0 .

Przykład: pochodna w zadanym kierunku

Dla $f(x, y, z) = x^2 + 3xyz + yz^3$ wyznaczamy gradient:

$$\nabla f = (2x + 3yz, 3xz + z^3, 3xy + 3yz^2)^T.$$

W punkcie $a = (5, 2, 1)$:

$$\nabla f(a) = (16, 16, 36)^T.$$

Funkcja jest wielomianem, a więc jest różniczkowalna. Dla podanego wektora $v = (3, 3, 3)$:

$$D_v f(a) = 16 \cdot 3 + 16 \cdot 3 + 36 \cdot 3 = 204.$$

Jeżeli chodzi o tempo zmiany na jednostkę drogi, normalizujemy wektor:

$$e = \frac{v}{\|v\|} = \frac{(1, 1, 1)}{\sqrt{3}}, \quad D_e f(a) = \frac{68}{\sqrt{3}}.$$

Są to różne parametryzacje tej samej skierowanej prostej.

Przykład: kierunek najszybszego wzrostu

Dla $f(x, y) = x^2 + xy$ oraz $a = (2, 1)$:

$$\nabla f(x, y) = (2x + y, x)^T, \quad \nabla f(2, 1) = (5, 2)^T, \quad \|\nabla f(2, 1)\| = \sqrt{29}.$$

Jednostkowy kierunek najszybszego wzrostu i odpowiadające mu tempo to

$$v_{\max} = \frac{(5, 2)}{\sqrt{29}}, \quad \max_{\|v\|=1} D_v f(2, 1) = \sqrt{29}.$$

Ten sam zwrot opisuje półprosta

$$(x, y) = (2, 1) + t(5, 2), \quad t \geq 0.$$

Wzdłuż tej parametryzacji pochodna w $t = 0$ wynosi 29, ponieważ prędkość ruchu to $\sqrt{29}$, a nie 1.

Kierunek najszybszego spadku to $-(5, 2)/\sqrt{29}$, z pochodną $-\sqrt{29}$.

Reguła łańcuchowa: krzywa na płaszczyźnie

Niech $f : \Omega \rightarrow \mathbb{R}$, gdzie $\Omega \subset \mathbb{R}^2$ jest otwarty. Niech $\gamma(t) = (x(t), y(t))$ będzie określona w otoczeniu t_0 i przyjmowała wartości w Ω .

Założmy, że x, y są różniczkowalne w t_0 , a f jest różniczkowalna w $a = \gamma(t_0)$. Dla $z(t) = f(x(t), y(t))$ zachodzi

Pochodna funkcji złożonej

$$z'(t_0) = f_x(a)x'(t_0) + f_y(a)y'(t_0).$$

W zapisie wektorowym:

$$z'(t_0) = \nabla f(\gamma(t_0)) \cdot \gamma'(t_0).$$

Pochodne f obliczamy w punkcie krzywej, a pochodne x, y w odpowiadającej mu wartości parametru. Oba składniki uwzględniają wpływ zmieniających się współrzędnych na wartość f .

Dowód reguły łańcuchowej: przyrosty

Oznaczmy $a = (x(t_0), y(t_0))$ oraz

$$\Delta x = x(t_0 + \Delta t) - x(t_0), \quad \Delta y = y(t_0 + \Delta t) - y(t_0), \quad h = (\Delta x, \Delta y).$$

Różniczkowalność f daje

$$z(t_0 + \Delta t) - z(t_0) = f_x(a)\Delta x + f_y(a)\Delta y + R(h).$$

Dla $h \neq 0$ niech $\varepsilon(h) = R(h)/\|h\|$, a dla $h = 0$ przyjmijmy $\varepsilon(0) = 0$. Wtedy $\varepsilon(h) \rightarrow 0$.

Po podzieleniu przez $\Delta t \neq 0$:

$$\frac{z(t_0 + \Delta t) - z(t_0)}{\Delta t} = f_x(a)\frac{\Delta x}{\Delta t} + f_y(a)\frac{\Delta y}{\Delta t} + \frac{R(h)}{\Delta t}.$$

Pierwsze dwa ilorazy dążą do $x'(t_0)$ i $y'(t_0)$. Pozostaje wykazać zanikanie ostatniego składnika.

Dowód reguły łańcuchowej: kontrola błędu

Różniczkowalność x, y pociąga ich ciągłość, więc $h \rightarrow 0$, gdy $\Delta t \rightarrow 0$. Mamy

$$\left| \frac{R(h)}{\Delta t} \right| = |\varepsilon(h)| \sqrt{\left(\frac{\Delta x}{\Delta t} \right)^2 + \left(\frac{\Delta y}{\Delta t} \right)^2}.$$

Pierwszy czynnik dąży do zera, a drugi do skończonej liczby

$$\sqrt{x'(t_0)^2 + y'(t_0)^2}.$$

Iloczyn dąży zatem do zera. Równość z modułem pozostaje poprawna także dla ujemnego Δt oraz dla przyrostów $h = 0$.

Przechodząc do granicy w poprzednim wzorze, otrzymujemy

$$z'(t_0) = f_x(a)x'(t_0) + f_y(a)y'(t_0).$$

Nie trzeba zakładać, że prędkość krzywej jest niezerowa.

Reguła łańcuchowa: przykład w punkcie

Niech

$$f(x, y) = x^2 + y \sin x, \quad x(t) = 2t + \sinh t, \quad y(t) = e^t - t^2.$$

Szukamy $z'(0)$ dla $z(t) = f(x(t), y(t))$. Ponieważ $x(0) = 0$, $y(0) = 1$, obliczamy pochodne f w $(0, 1)$:

$$f_x = 2x + y \cos x,$$

$$f_x(0, 1) = 1,$$

$$f_y = \sin x,$$

$$f_y(0, 1) = 0.$$

Pochodne współrzędnych wynoszą

$$x'(t) = 2 + \cosh t, \quad x'(0) = 3, \quad y'(t) = e^t - 2t, \quad y'(0) = 1.$$

Zatem

$$z'(0) = 1 \cdot 3 + 0 \cdot 1 = 3.$$

Wszystkie użyte funkcje są gładkie, więc założenia reguły łańcuchowej są spełnione.

Ten sam przykład: bezpośrednie sprawdzenie

Po wstawieniu $x(t)$ i $y(t)$ do wzoru na f :

$$z(t) = (2t + \sinh t)^2 + (e^t - t^2) \sin(2t + \sinh t).$$

Różniczkujemy sumę oraz iloczyn:

$$\begin{aligned} z'(t) &= 2(2t + \sinh t)(2 + \cosh t) \\ &\quad + (e^t - 2t) \sin(2t + \sinh t) \\ &\quad + (e^t - t^2) \cos(2t + \sinh t)(2 + \cosh t). \end{aligned}$$

Podstawiając $t = 0$, korzystamy z

$$\sinh 0 = 0, \quad \cosh 0 = 1, \quad e^0 = 1, \quad \sin 0 = 0, \quad \cos 0 = 1.$$

Otrzymujemy

$$z'(0) = 0 + 0 + 1 \cdot 1 \cdot 3 = 3.$$

Wynik jest zgodny z obliczeniem przy użyciu pochodnych cząstkowych.

Przykład: spirala i suma kwadratów

Niech $f(x, y) = x^2 + y^2$, $x(t) = e^t \cos t$, $y(t) = e^t \sin t$. Wtedy

$$f_x = 2x, \quad f_y = 2y, \quad x' = e^t(\cos t - \sin t), \quad y' = e^t(\sin t + \cos t).$$

Z reguły łańcuchowej:

$$\begin{aligned} z' &= 2e^{2t} \cos t(\cos t - \sin t) + 2e^{2t} \sin t(\sin t + \cos t) \\ &= 2e^{2t}(\cos^2 t + \sin^2 t) = 2e^{2t}. \end{aligned}$$

Można też najpierw uprościć funkcję złożoną:

$$z(t) = e^{2t} \cos^2 t + e^{2t} \sin^2 t = e^{2t},$$

więc od razu $z'(t) = 2e^{2t}$.

Oba sposoby uwzględniają zmianę obu współrzędnych; składniki mieszane w pierwszym rachunku wzajemnie się redukują.

Druga pochodna funkcji wzdłuż krzywej

Założmy, że f jest klasy C^2 w otoczeniu krzywej, a $x(t), y(t)$ są klasy C^2 .
Dla $z(t) = f(x(t), y(t))$ mamy

$$z' = f_x x' + f_y y'.$$

Pochodne f_x, f_y są tu funkcjami złożonymi. Dlatego

$$\frac{d}{dt} f_x(x(t), y(t)) = f_{xx} x' + f_{xy} y',$$

$$\frac{d}{dt} f_y(x(t), y(t)) = f_{yx} x' + f_{yy} y'.$$

Stosujemy dodatkowo regułę iloczynu:

$$\begin{aligned} z'' &= (f_{xx} x' + f_{xy} y') x' + f_x x'' \\ &\quad + (f_{yx} x' + f_{yy} y') y' + f_y y''. \end{aligned}$$

Wszystkie pochodne f są obliczane w $(x(t), y(t))$, a pochodne współrzędnych w t .

Druga pochodna: wzór końcowy i hesjan

Ponieważ $f \in C^2$, mamy $f_{xy} = f_{yx}$. Po zebraniu składników:

$$z'' = f_{xx}(x')^2 + 2f_{xy}x'y' + f_{yy}(y')^2 + f_x x'' + f_y y''.$$

Pierwsze trzy składniki uwzględniają drugie pochodne funkcji f , a dwa ostatnie przyspieszenie ruchu po krzywej.

Dla $\gamma = (x, y)^T$ i

$$H_f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

możemy zapisać ten sam wzór jako

$$z''(t) = \gamma'(t)^T H_f(\gamma(t)) \gamma'(t) + \nabla f(\gamma(t)) \cdot \gamma''(t).$$

Jeśli krzywa jest prostą parametryzowaną ze stałą prędkością, to $\gamma'' = 0$ i drugi składnik znika.

Krzywa w przestrzeni wielowymiarowej

Dla $\gamma(t) = (x_1(t), \dots, x_n(t))^T$ różniczkowalność oznacza istnienie granicy wektorowej

$$\gamma'(t_0) = \lim_{h \rightarrow 0} \frac{\gamma(t_0 + h) - \gamma(t_0)}{h}.$$

Jest ona równoważna istnieniu wszystkich pochodnych współrzędnych. Wówczas

$$\gamma'(t_0) = (x'_1(t_0), \dots, x'_n(t_0))^T.$$

Jeżeli $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ jest różniczkowalna w $\gamma(t_0)$, to

$$(f \circ \gamma)'(t_0) = \sum_{i=1}^n f_{x_i}(\gamma(t_0)) x'_i(t_0) = \nabla f(\gamma(t_0)) \cdot \gamma'(t_0).$$

Dowód jest taki sam jak w dwóch wymiarach: zamiast dwóch składników występuje skończona suma n składników.

Złożenie zależne od dwóch parametrów

Niech $D \subset \mathbb{R}^2$ i $\Omega \subset \mathbb{R}^2$ będą otwarte, $X(u, v) = (x(u, v), y(u, v)) \in \Omega$ dla $(u, v) \in D$, oraz

$$Z(u, v) = f(x(u, v), y(u, v)).$$

Jeżeli X jest różniczkowalne w (u, v) , a f w $X(u, v)$, to

$$Z_u = f_x x_u + f_y y_u, \quad Z_v = f_x x_v + f_y y_v.$$

Przy obliczaniu Z_u ustalamy v i stosujemy regułę łańcuchową do krzywej $u \mapsto X(u, v)$. Dla Z_v postępujemy analogicznie.

Pochodne f_x, f_y w powyższych wzorach oznaczają wartości w punkcie $(x(u, v), y(u, v))$. Pochodne x_u, x_v, y_u, y_v są obliczane w punkcie (u, v) .

Różne litery f i Z pozwalają rozróżnić funkcję w starych zmiennych od jej złożenia w nowych zmiennych.

Reguła łańcuchowa: zapis macierzowy

Macierz Jacobiego odwzorowania $X = (x, y)^T$ ma postać

$$J_X(u, v) = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix}.$$

Przy przyjętej konwencji gradientów kolumnowych:

$$\begin{pmatrix} Z_u \\ Z_v \end{pmatrix} = \begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix} \begin{pmatrix} f_x \\ f_y \end{pmatrix},$$

czyli

$$\nabla Z(u, v) = J_X(u, v)^T \nabla f(X(u, v)).$$

Transpozycja jest istotna: wiersz odpowiadający parametrowi u zawiera pochodne obu współrzędnych względem u .

Różniczka funkcji skalarnej jest w zapisie macierzowym wierszem $Df = (\nabla f)^T$; stąd również $D(f \circ X) = Df DX$.

Ogólne odwzorowanie i jego pochodna

Niech $D \subset \mathbb{R}^m$ będzie otwarty, a $X : D \rightarrow \mathbb{R}^n$. Mówimy, że X jest różniczkowalne w a , gdy istnieje macierz A rozmiaru $n \times m$ taka, że

$$X(a+h) = X(a) + Ah + r(h), \quad \frac{\|r(h)\|}{\|h\|} \rightarrow 0.$$

Równoważnie każda funkcja składowa X_i jest różniczkowalna. Macierz $A = DX(a) = J_X(a)$ składa się z pochodnych

$$J_X(a) = \left(\frac{\partial X_i}{\partial u_j}(a) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}.$$

Wiersze odpowiadają funkcjom składowym, a kolumny zmiennym wejściowym. Przyrost $h \in \mathbb{R}^m$ jest przekształcany w przyrost liniowy $Ah \in \mathbb{R}^n$.

Równoważność z różniczkowalnością składowych wynika z porównania normy reszty wektorowej z modułami jej składowych.

Ogólna reguła łańcuchowa

Niech $X : D \subset \mathbb{R}^m \rightarrow \Omega \subset \mathbb{R}^n$ i $f : \Omega \rightarrow \mathbb{R}$, gdzie D, Ω są otwarte. Jeśli X jest różniczkowalne w a , a f w $X(a)$, to $Z = f \circ X$ jest różniczkowalna w a i

$$DZ(a) = Df(X(a)) \circ DX(a).$$

Współrzędnościowo, dla $j = 1, \dots, m$:

$$\frac{\partial Z}{\partial u_j}(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X(a)) \frac{\partial X_i}{\partial u_j}(a).$$

W zapisie gradientowym:

$$\nabla Z(a) = J_X(a)^T \nabla f(X(a)).$$

Wymiary są zgodne: macierz $m \times n$ mnożymy przez wektor $n \times 1$, otrzymując gradient $m \times 1$.

Dowód ogólnej reguły łańcuchowej

Oznaczmy $b = X(a)$, $A = DX(a)$ i $L = Df(b)$. Mamy

$$X(a+h) = b + Ah + r(h), \quad r(h) = o(\|h\|),$$

$$f(b+k) = f(b) + L(k) + s(k), \quad s(k) = o(\|k\|).$$

Podstawiamy $k = Ah + r(h)$. Wtedy $\|k\| = O(\|h\|)$, czyli dla małego h zachodzi $\|k\| \leq C\|h\|$ z pewną stałą C .

$$Z(a+h) - Z(a) = L(Ah) + L(r(h)) + s(k).$$

Liniowe odwzorowanie L jest ograniczone, więc $L(r(h)) = o(\|h\|)$. Dla $k \neq 0$:

$$\frac{|s(k)|}{\|h\|} = \frac{|s(k)|}{\|k\|} \frac{\|k\|}{\|h\|} \rightarrow 0.$$

Dla $k = 0$ reszta wynosi zero. Pozostaje część liniowa $L \circ A$, a więc $DZ(a) = Df(b) \circ DX(a)$.

Drugie pochodne złożenia: pochodna czysta

Niech $Z(u, v) = f(x(u, v), y(u, v))$, gdzie wszystkie funkcje są klasy C^2 na odpowiednich otwartych dziedzinach. Zaczynamy od

$$Z_u = f_x x_u + f_y y_u.$$

Różniczkując jeszcze raz po u , stosujemy reguły iloczynu i łańcuchową:

$$\begin{aligned} Z_{uu} &= (f_{xx} x_u + f_{xy} y_u) x_u + f_x x_{uu} \\ &\quad + (f_{yx} x_u + f_{yy} y_u) y_u + f_y y_{uu}. \end{aligned}$$

Ponieważ $f_{xy} = f_{yx}$:

$$Z_{uu} = f_{xx} x_u^2 + 2f_{xy} x_u y_u + f_{yy} y_u^2 + f_x x_{uu} + f_y y_{uu}.$$

Pochodne f są obliczane w $(x(u, v), y(u, v))$. Nie są stałymi podczas kolejnego różniczkowania.

Drugie pochodne złożenia: druga zmienna

Z wzoru $Z_v = f_x x_v + f_y y_v$ otrzymujemy analogicznie

$$\begin{aligned} Z_{vv} &= (f_{xx}x_v + f_{xy}y_v)x_v + f_x x_{vv} \\ &\quad + (f_{yx}x_v + f_{yy}y_v)y_v + f_y y_{vv}. \end{aligned}$$

Po zebraniu wyrazów:

$$Z_{vv} = f_{xx}x_v^2 + 2f_{xy}x_v y_v + f_{yy}y_v^2 + f_x x_{vv} + f_y y_{vv}.$$

W obu wzorach występują dwie grupy składników: drugie pochodne f pomnożone przez pierwsze pochodne współrzędnych oraz pierwsze pochodne f pomnożone przez drugie pochodne współrzędnych.

Jeżeli zmiana zmiennych jest afiniczna, to $x_{uu}, x_{vv}, y_{uu}, y_{vv}$ są zerowe. Wtedy pozostają wyłącznie składniki pierwszej grupy.

Drugie pochodne złożenia: pochodna mieszana

Różniczkujemy $Z_u = f_x x_u + f_y y_u$ względem v :

$$\begin{aligned} Z_{uv} &= (f_{xx}x_v + f_{xy}y_v)x_u + f_x x_{uv} \\ &\quad + (f_{yx}x_v + f_{yy}y_v)y_u + f_y y_{uv}. \end{aligned}$$

Stąd

$$\begin{aligned} Z_{uv} &= f_{xx}x_u x_v + f_{xy}(x_u y_v + y_u x_v) \\ &\quad + f_{yy}y_u y_v + f_x x_{uv} + f_y y_{uv}. \end{aligned}$$

Założenia C^2 zapewniają zarówno $f_{xy} = f_{yx}$, jak i równość mieszanych pochodnych x, y, Z .

W przeciwieństwie do pochodnej czystej składnik przy f_{xy} jest sumą dwóch różnych iloczynów. Nie można go bez dodatkowego powodu zastąpić dwukrotnością jednego z nich.

Współrzędne biegunowe: rozróżnienie funkcji

Niech $z = z(x, y)$ i wprowadźmy

$$x = r \cos \theta, \quad y = r \sin \theta, \quad Z(r, \theta) = z(r \cos \theta, r \sin \theta).$$

Dla funkcji różniczkowalnej reguła łańcuchowa daje

$$\begin{aligned} Z_r &= z_x \cos \theta + z_y \sin \theta, \\ Z_\theta &= -rz_x \sin \theta + rz_y \cos \theta. \end{aligned}$$

Pochodne z_x, z_y są liczone w $(r \cos \theta, r \sin \theta)$.

Macierz Jacobiego tej parametryzacji i jej wyznacznik:

$$J = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}, \quad \det J = r.$$

Dla $r > 0$ można lokalnie odwrócić zmianę zmiennych. W $r = 0$ wyznacznik znika, a kąt nie jest wyznaczony jednoznacznie.

Współrzędne biegunowe: wzory odwrotne

Dla $r > 0$ dzielimy drugie równanie przez r :

$$Z_r = z_x \cos \theta + z_y \sin \theta, \quad \frac{Z_\theta}{r} = -z_x \sin \theta + z_y \cos \theta.$$

Pierwsze równanie mnożymy przez $\cos \theta$, drugie przez $-\sin \theta$ i dodajemy. Składniki z z_y się redukują:

$$z_x = \cos \theta Z_r - \frac{\sin \theta}{r} Z_\theta.$$

Analogicznie, mnożąc przez $\sin \theta$ oraz $\cos \theta$:

$$z_y = \sin \theta Z_r + \frac{\cos \theta}{r} Z_\theta.$$

Wzory wykorzystują $\cos^2 \theta + \sin^2 \theta = 1$. Nie wolno stosować ich w $r = 0$, ponieważ zawierają dzielenie przez r .

Przekształcenie wyrażenia z pochodnymi

Wyrażmy $xz_x + yz_y$ we współrzędnych biegunowych. Dla $r > 0$ podstawiamy wzory odwrotne:

$$\begin{aligned}xz_x + yz_y &= r \cos \theta \left(\cos \theta Z_r - \frac{\sin \theta}{r} Z_\theta \right) \\&\quad + r \sin \theta \left(\sin \theta Z_r + \frac{\cos \theta}{r} Z_\theta \right) \\&= r(\cos^2 \theta + \sin^2 \theta) Z_r \\&\quad + (-\cos \theta \sin \theta + \sin \theta \cos \theta) Z_\theta \\&= r Z_r.\end{aligned}$$

Otrzymujemy

$$xz_x + yz_y = r Z_r.$$

Krótsze sprawdzenie polega na pomnożeniu przez r wzoru $Z_r = z_x \cos \theta + z_y \sin \theta$. Wyrażenie opisuje zmianę w kierunku promieniowym.

Liniowa zmiana zmiennych

Wprowadźmy $u = x + y$, $v = x + 2y$. Wyznacznik tej zmiany wynosi 1, więc jest ona odwracalna:

$$x = 2u - v, \quad y = v - u.$$

Dla $z = z(x, y)$ oznaczmy

$$Z(u, v) = z(2u - v, v - u), \quad z(x, y) = Z(x + y, x + 2y).$$

Z reguły łańcuchowej:

$$z_x = Z_u + Z_v, \quad z_y = Z_u + 2Z_v.$$

Symbolicznie, przy działaniu na funkcje zapisane w nowych zmiennych:

$$\partial_x = \partial_u + \partial_v, \quad \partial_y = \partial_u + 2\partial_v.$$

Pochodne Z w pierwszych wzorach są obliczane w $(u, v) = (x + y, x + 2y)$.

Liniowa zmiana: pochodna mieszana

Założmy, że z jest klasy C^2 . Różniczkujemy $z_x = Z_u + Z_v$ względem y :

$$\begin{aligned} z_{xy} &= \partial_y(Z_u + Z_v) \\ &= (Z_{uu}u_y + Z_{uv}v_y) + (Z_{vu}u_y + Z_{vv}v_y) \\ &= Z_{uu} + 2Z_{uv} + Z_{vu} + 2Z_{vv}. \end{aligned}$$

Ponieważ $Z \in C^2$, twierdzenie Schwarz'a daje $Z_{vu} = Z_{uv}$. W rezultacie

$$z_{xy} = Z_{uu} + 3Z_{uv} + 2Z_{vv}.$$

To samo wynika z rozwinięcia iloczynu operatorów

$$(\partial_u + 2\partial_v)(\partial_u + \partial_v).$$

Współczynniki tej zmiany są stałe, więc ich pochodne nie wnoszą dodatkowych składników.

Przesunięte współrzędne biegunowe

Dla koła otwartego o środku (a, b) i promieniu $R > 0$:

$$D = \{(x, y) : (x - a)^2 + (y - b)^2 < R^2\}$$

wygodna jest parametryzacja

$$\boxed{x = a + r \cos \theta, \quad y = b + r \sin \theta,}$$

$$0 \leq r < R, \quad 0 \leq \theta < 2\pi.$$

Rzeczywiście,

$$(x - a)^2 + (y - b)^2 = r^2(\cos^2 \theta + \sin^2 \theta) = r^2 < R^2.$$

Odwrotnie, każdy punkt koła ma odległość $r < R$ od środka i może być tak przedstawiony. Wartość $r = R$ odpowiada brzegowi, który nie należy do koła otwartego.

Parametryzacja koła a lokalne współrzędne

W środku koła $r = 0$ wszystkie kąty opisują ten sam punkt (a, b) .

Parametryzacja całego koła nie jest więc różnowartościowa.

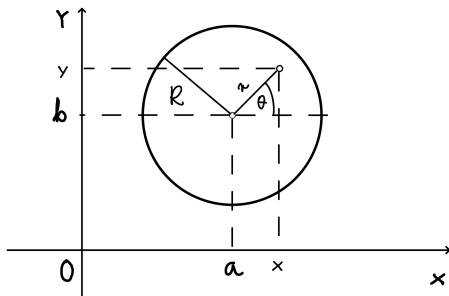
Dla $r > 0$ przedział $0 \leq \theta < 2\pi$ wybiera po jednym reprezentancie kąta. Jednak przy przejściu przez promień odpowiadający $\theta = 0$ reprezentant przeskakuje pomiędzy wartościami bliskimi 2π i 0 .

Dlatego do różniczkowania używamy **lokalnego** przedziału kątów wokół badanego punktu, bez takiego skoku. Wyznacznik Jacobiego nadal wynosi r i jest niezerowy poza środkiem.

Warunek $r \geq 0$ w opisie parametrów nie oznacza, że koło jest domknięte: jego środek jest punktem wewnętrznym. O brzegu decyduje równanie $r = R$.

Wzory zawierające $1/r$ obowiązują tylko dla $r > 0$; zachowanie funkcji w środku trzeba badać oddzielnie.

Przesunięte współrzędne: interpretacja geometryczna



Promień r jest odległością punktu (x, y) od środka (a, b) , a kąt θ mierzymy względem dodatniego kierunku osi poziomej wychodzącej ze środka.

$$x - a = r \cos \theta, \quad y - b = r \sin \theta.$$

Na schemacie okrąg o promieniu R wyznacza brzeg koła; punkty dziedziny opisane są warunkiem $r < R$.