Fizyka dla firm — Matematyka

28. Funkcje wielu zmiennych

P. F. Góra

http://th-www.if.uj.edu.pl/zfs/gora/

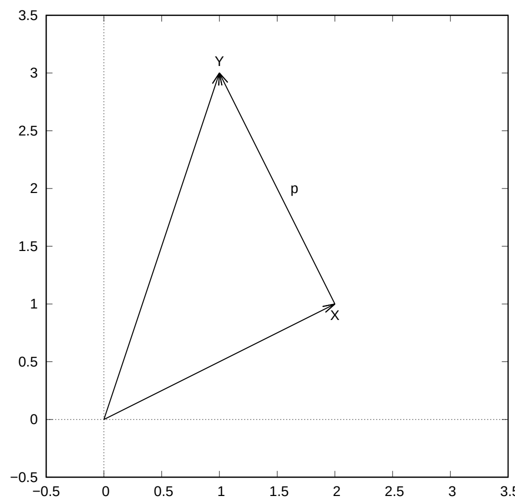
14 kwietnia 2021

Dodawanie punktów i wektorów ☺

Punkt w przestrzeni $X \in \mathbb{R}^N$ można utożsamiać z jego wektorem wodzącym $\vec{X}$. Jeśli $\mathbf{p} \in \mathbb{R}^N$ jest wektorem, napis

$$\vec{Y} = \vec{X} + \mathbf{p} \tag{1}$$

będzie oznaczać *punkt* o wektorze wodzącym $\vec{Y}$.



Funkcje wielu zmiennych

Rozważąmy przestrzeń $\mathbb{R}^N$ z zadaną metryką $\varrho(\cdot,\cdot)$. Nas będzie interesować wyłącznie metryka euklidesowa, ale w rozmaitych zastosowaniach bardziej użyteczne mogą okazać się inne metryki.

Funkcją rzeczywistą N zmiennych rzeczywistych nazywam **relację jednoznaczną** $f : \mathbb{R}^N \to \mathbb{R}$: Pewnemu punktowi przestrzeni $\mathbb{R}^N$ — lub pewnemu zestawowi zmiennych $x_1, x_2, \dots, x_N$, będącymi współrzędnymi tego punktu — przyporządkowuję w sposób **jednoznaczny** pewną liczbę rzeczywistą, będącą wartością funkcji w tym punkcie.

Jednoznaczność ma zasadnicze znaczenie: temu samemu punktowi może być przypisana jedna i tylko jedna wartość funkcji, ale, w ogólności, wartość ta może być przypisana wielu punktom. Oczywiście *inne* funkcje mogą w tym punkcie przybierać inne wartości.

Przykład 1

Przykłady funkcji rzeczywistych dwu, trzech, czterech zmiennych rzeczywistych:

$$f_1(x,y) = x^2 + xy + y^2 \tag{2a}$$

$$f_2(x,y) = x \cdot \sin(y) - y \cdot \cos(x) \tag{2b}$$

$$f_3(x,y,z) = \frac{xy + xz + yz}{x^2 + y^2 + z^2} \tag{2c}$$

$$f_4(x_1,x_2,x_3,x_4) = \frac{\ln\left(4x_1^2 + 3x_2^2 + 2x_3^2 + x_1^2\right)}{x_1^2 + x_2^2 + x_3^2 + x_4^2} \tag{2d}$$

Zauważmy, że $f_1, f_2 : \mathbb{R}^2 \to \mathbb{R}$, $f_3 : \mathbb{R}^3 \to \mathbb{R}$, $f_4 : \mathbb{R}^4 \to \mathbb{R}$.

Granica funkcji wielu zmiennych

W przestrzeni metrycznej $\left(\mathbb{R}^N, \varrho\right)$ możemy zdefiniować granicę funkcji w punkcie, uogólniając znaną z teorii funkcji jednej zmiennej definicję Cauchy'ego:

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = g \Leftrightarrow \forall\, \varepsilon > 0\, \exists\, \delta > 0\, \forall\, \mathbf{x} : (\varrho(\mathbf{x},\mathbf{x}_0) < \delta \Rightarrow |f(\mathbf{x}) - g| < \varepsilon) \tag{3}$$

Można też sformułować równoważną definicję Heinego, wykorzystującą zbieżność ciągów:

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = g \Leftrightarrow \forall \{\mathbf{x}_n\}_{n=1}^{\infty} : \left(\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}_0 \Rightarrow \lim_{n\to\infty} f(\mathbf{x}_n) = g\right) \tag{4}$$

Można także definiować granice niewłaściwe.

Uwaga!

Granica funkcji wielu zmiennych może nie istnieć, podobnie jak nie musi istnieć granica funkcji jednej zmiennej. W przypadku wielowymiarowym pojawia się jednak nowy problem, niewystęujący w przypadku jednowymiarowym: Granica funkcji wielu zmiennych nie może zależeć od drogi, wzdłuż której zmierzamy, czy też od sposobu, w jaki zmierzamy do granicy.

Przykład 2

Przyjrzyjmy się funkcji

$$f(x,y) = \frac{xy}{x^2 + y^2} \tag{5}$$

i rozpatrzmy granicę

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2 + y^2} \tag{6}$$

1. Niech $y = x$. Wówczas granica (6) przybiera postać

$$\lim_{x\to 0} \frac{x^2}{x^2 + x^2} = \lim_{x\to 0} \frac{1}{2} = \frac{1}{2} \tag{7a}$$

2. Niech $y = \sqrt{|x|}$. W tym wypadku otrzymujemy

$$\lim_{x\to 0} \frac{x\sqrt{|x|}}{x^2 + |x|} = \lim_{x\to 0} \frac{(\mathrm{sgn}(x))\sqrt{|x|}}{x\,\mathrm{sgn}(x) + 1} = 0 \tag{7b}$$

3. Niech teraz $x = r\cos\varphi, y = r\sin\varphi$. W tym wypadku otrzymujemy

$$\lim_{r\to 0^+} \frac{r^2 \sin\varphi\ \cos\varphi}{r^2\cos^2\varphi + r^2\sin^2\varphi} = \frac{1}{2}\sin(2\varphi) \tag{7c}$$

Jak widzimy, granica (6) przybiera różne wartości w zależności od sposobu podążania do punktu $(0, 0)$. Wnioskujemy stąd, że granica (6) nie istnieje.

Granice iterowane

Jeżeli granica (3) istnieje, istnieją też granice iterowane, czyli granice, gdy poszczególne zmienne “po kolei” zmierzają do punktu granicznego.

Dla przykładu funkcji dwóch zmiennych

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = g \;\Rightarrow\; \lim_{x\to x_0}\left(\lim_{y\to y_0} f(x,y)\right) = g \;\wedge\; \lim_{y\to y_0}\left(\lim_{x\to x_0} f(x,y)\right) = g \tag{8}$$

Istnienie i równość granic iterowanych jest warunkiem koniecznym istnienia granicy, ale *nie* jest warunkiem wystarczającym.

Przykład 3

Po raz kolejny rozpatrzmy funkcję (5). Granice iterowane spełniają

$$\lim_{x\to 0}\left(\lim_{y\to 0}\frac{xy}{x^2+y^2}\right) = \lim_{x\to 0}\frac{x\cdot 0}{x^2+0^2} = \lim_{x\to 0} 0 = 0 \tag{9a}$$

$$\lim_{y\to 0}\left(\lim_{x\to 0}\frac{xy}{x^2+y^2}\right) = \lim_{y\to 0}\frac{0\cdot y}{0^2+y^2} = \lim_{y\to 0} 0 = 0 \tag{9b}$$

a jednak granica (6) nie istnieje.

Zdefiniowawszy pojęcie granicy w przestrzeni $\mathbb{R}^N$, możemy teraz zdefiniować ciągłość funkcji, pojęcie ciągu Cauchy'ego, a także zastanowić się, w jaki sposób można uogólnić pojęcie pochodnej. Dla funkcji jednej zmiennej x pochodna określa prędkość zmian, gdy wartość zmiennej niezależnej zmienia się o infinitezymalnie małą wielkość. W przypadku funkcji wielu zmiennych każda z nich może się zmieniać niezależnie od pozostałych.

Pochodna kierunkowa

Niech $f : \mathbb{R}^N \to \mathbb{R}$ będzie pewną funkcją, $\mathbf{x}_0 \in \mathbb{R}^N$ pewnym punktem, $\mathbf{p} \in \mathbb{R}^N$ — pewnym wektorem unormowanym do jedności, $\|\mathbf{p}\| = 1$. Zdefiniujmy

$$g(\alpha) = f(\mathbf{x}_0 + \alpha\mathbf{p}) \tag{10}$$

Zauważmy, że funkcja (10) *jest funkcją jednej zmiennej, α*. Możemy zatem rozważać jej pochodną

$$\left.\frac{dg}{d\alpha}\right|_{\alpha=0} = \left.\left(\frac{d}{d\alpha} f(\mathbf{x}_0 + \alpha\mathbf{p})\right)\right|_{\alpha=0} \tag{11}$$

Jeżeli pochodna (11) istnieje, nazywam ją **pochodną kierunkową funkcji f** w punkcie $\mathbf{x}_0$ w kierunku wektora $\mathbf{p}$.

Przykład 4

$$f(x, y) = (x + y)^2 \cdot \sin(x - y) \tag{12}$$

Niech $\mathbf{x}_0 = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$, $\mathbf{p} = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$. Wówczas funkcja (10) przybiera postać

$$\begin{aligned} g(\alpha) &= \left(\frac{\pi}{2} + \frac{\alpha}{\sqrt{2}} + \frac{\pi}{2} + \frac{\alpha}{\sqrt{2}}\right)^2 \sin\left(\frac{\pi}{2} + \frac{\alpha}{\sqrt{2}} - \frac{\pi}{2} - \frac{\alpha}{\sqrt{2}}\right) \\ &= \left(\pi + \sqrt{2}\,\alpha\right)^2 \cdot \sin 0 = 0 \end{aligned} \tag{13}$$

$\frac{d0}{d\alpha} = 0$ i widzimy, że pochodna kierunkowa funkcji (12) w podanym punkcie i w kierunku podanego wektora wynosi 0.

Przykład 5

Niech $f(x,y)$ będzie określona wzorem (12), $\mathbf{x}_0 = \left(\frac{\pi}{2}, \frac{\pi}{2}\right)$ jak poprzednio, ale $\mathbf{p} = \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right]$. Wówczas funkcja (10) przybiera postać

$$\begin{aligned} g(\alpha) &= \left(\frac{\pi}{2} + \frac{\alpha}{\sqrt{2}} - \frac{\pi}{2} + \frac{\alpha}{\sqrt{2}}\right)^2 \sin\left(\frac{\pi}{2} + \frac{\alpha}{\sqrt{2}} - \frac{\pi}{2} + \frac{\alpha}{\sqrt{2}}\right) \\ &= \pi^2 \cdot \sin\left(\sqrt{2}\,\alpha\right) \end{aligned} \tag{14}$$

Zatem pochodna kierunkowa funkcji (12) w tym samym punkcie, co poprzednio, ale w kierunku drugiego podanego wektora wynosi

$$\begin{aligned} \left.\frac{dg}{d\alpha}\right|_{\alpha=0} &= \left.\frac{d}{d\alpha}\pi^2 \cdot \sin\left(\sqrt{2}\,\alpha\right)\right|_{\alpha=0} \\ &= \left.\pi^2\sqrt{2}\cos\left(\sqrt{2}\alpha\right)\right|_{\alpha=0} = \sqrt{2}\,\pi^2 \end{aligned} \tag{15}$$

Pochodne cząstkowe

Pochodne kierunkowe w kierunku kolejnych wektorów osi kartezjańskiego układu współrzędnych nazywam pochodnymi cząstkowymi i oznaczam $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots$, gdzie $x_1, x_2, \dots$ są kolejnymi zmiennymi.

Tak więc

$$\frac{\partial f}{\partial x_1} = \left.\frac{d}{d\alpha} f(\mathbf{x} + \alpha \cdot [1, 0, 0, \dots])\right|_{\alpha=0} = \left.\frac{d}{d\alpha} f(x_1 + \alpha, x_2, x_3, \dots)\right|_{\alpha=0} \tag{16a}$$

$$\frac{\partial f}{\partial x_2} = \left.\frac{d}{d\alpha} f(\mathbf{x} + \alpha \cdot [0, 1, 0, \dots])\right|_{\alpha=0} = \left.\frac{d}{d\alpha} f(x_1, x_2 + \alpha, x_3, \dots)\right|_{\alpha=0} \tag{16b}$$

$$\dots$$

Operacyjnie pochodne cząstkowe oblicza się jak “zwykłe” pochodne, traktując pozostałe zmienne jak parametry o ustalonych wartościach.

Przykład 6

$$f(x,y) = xy\,,\ \frac{\partial f}{\partial x} = y\,,\ \frac{\partial f}{\partial y} = x \tag{17a}$$

$$f(x,y) = \frac{x+y}{x^2+y^2}\,,\ \frac{\partial f}{\partial x} = \frac{-x^2+y^2-2xy}{(x^2+y^2)^2}\,,\ \frac{\partial f}{\partial y} = \frac{x^2-y^2-2xy}{(x^2+y^2)^2} \tag{17b}$$

$$f(x,y) = \sin(2x+3y)\,,\ \frac{\partial f}{\partial x} = 2\cos(2x+3y)\,,\ \frac{\partial f}{\partial y} = 3\cos(2x+3y) \tag{17c}$$

Przykład 7

Pochodne cząstkowe funkcji (12) wynoszą

$$\frac{\partial f}{\partial x} = 2(x+y)\cdot\sin(x-y) + (x+y)^2\cos(x-y) \tag{18a}$$

$$\frac{\partial f}{\partial y} = 2(x+y)\cdot\sin(x-y) - (x+y)^2\cos(x-y) \tag{18b}$$

Pochodna zupełna

Dla uproszczenia notacji ograniczę się tu do funkcji dwu zmiennych, ale wynik można łatwo uogólnić na przypadki więcejwymiarowe.

Niech $f(u, v)$ będzie różniczkowalną funkcją dwu zmiennych. Załóżmy teraz, że $u = u(t), v = v(t)$ same są funkcjami jakiejś innej zmiennej, t. Wówczas

$$z(t) = f(u(t), v(t)) \tag{19}$$

Jak obliczyć pochodną $z(t)$? Jeśli znamy *jawną* postać funkcji f, możemy podstawić i różniczkować jak w przypadku jednowymiarowym. Okazuje się jednak, że — uogólniając pojęcie pochodnej funkcji złożonej — można pokazać, że

$$\frac{dz}{dt} = \frac{\partial f}{\partial u} \cdot \frac{du}{dt} + \frac{\partial f}{\partial v} \cdot \frac{dv}{dt} \tag{20}$$

lub, uogólniając na N zmiennych,

$$\frac{dz}{dt} = \frac{\partial f}{\partial x_1} \cdot \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_N} \cdot \frac{dx_N}{dt} \,. \tag{21}$$

Wzór (21) nazywa się wzorem na pochodną zupełną lub całkowitą.

Przykład 8

Rozpatrzmy funkcję (12) zakładając, że x, y zależą od zmiennej t w ten sposób: $x = at, y = bt$. Niech $z(t) = f(at, bt)$. Korzystając ze wzoru na pochodną zupełną i obliczonych wcześniej pochodnych cząstkowych, mamy

$$
\begin{aligned}
\frac{dz}{dt} &= \frac{\partial f}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dt} \\
&= \Big(2(at+bt)\cdot\sin(at-bt) + (at+bt)^2\cos(at-bt)\Big)\cdot a \\
&+ \Big(2(at+bt)\cdot\sin(at-bt) - (at+bt)^2\cos(at-bt)\Big)\cdot b \\
&= 2(a+b)^2 t\sin((a-b)t) + (a-b)(a+b)^2 t^2\cos((a-b)t)
\end{aligned}
\tag{22}
$$

Gradient funkcji

Wektor złożony z pochodnych cząstkowych (ze *wszystkich* pochodnych cząstkowych) pewnej funkcji w danym punkcie nazywam gradientem funkcji i oznaczam

$$\left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \dots\right] = \mathsf{grad}\, f = \nabla f \tag{23}$$

gdzie symbol ∇, "nabla", oznacza formalny operator różniczkowania po kolejnych współrzędnych:

$$\nabla = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \dots\right] \tag{24}$$

Geometryczna interpretacja gradientu

Niech $f : \mathbb{R}^N \to \mathbb{R}$ będzie pewną funkcją (co najmniej) jednokrotnie różniczkowalną w sposób ciągły. Przypuśćmy, że znajdujemy się w punkcie $\mathbf{x}$ i przesuwamy się o niewielką wartość α, $|\alpha| \ll 1$, w kierunku pewnego wektora $\mathbf{p}$:

$$\mathbf{x} \to \mathbf{x} + \alpha\mathbf{p} \,. \tag{25}$$

Pytanie, jak bardzo na skutek tego przesunięcia zmieni się wartość funkcji f?

Zauważmy, że funkcja

$$g(\alpha) = f(\mathbf{x} + \alpha\mathbf{p}) \tag{26}$$

podobnie jak wcześniej, jest, przy ustalonych wartościach $\mathbf{x}$ i $\mathbf{p}$, *funkcją jednej zmiennej*, α. Rozwijając ją w szereg Taylora do pierwszego rzędu dostajemy

$$\begin{aligned} g(\alpha) &\simeq g(0) + \alpha \left.\frac{dg}{d\alpha}\right|_{\alpha=0} && \text{(27a)} \\ \Delta f \equiv \Delta g = g(\alpha) - g(0) &\simeq \alpha \left.\frac{dg}{d\alpha}\right|_{\alpha=0} && \text{(27b)} \end{aligned}$$

gdzie "Δf" oznacza zmianę funkcji f.

Możemy napisać

$$g(\alpha) = f(\mathbf{x} + \alpha\mathbf{p}) = f(x_1 + \alpha p_1, x_2 + \alpha p_2, \dots, x_N + \alpha p_N) \quad (28)$$

zatem, zgodnie ze wzorem na pochodną zupełną (21),

$$\left.\frac{dg}{d\alpha}\right|_{\alpha=0} = \sum_{i=1}^{N} \left. \left(\frac{\partial f(\mathbf{x}+\alpha\mathbf{p})}{\partial(x_i+\alpha p_i)} \cdot \frac{d(x_i+\alpha p_i)}{d\alpha} \right) \right|_{\alpha=0}$$

$$= \sum_{i=1}^{N} \frac{\partial f}{\partial x_i} p_i = (\nabla f) \circ \mathbf{p} \tag{29}$$

gdzie p_i są składowymi wektora $\mathbf{p}$. Innymi słowy,

$$\Delta f \simeq \alpha \cdot (\nabla f) \circ \mathbf{p} \tag{30}$$

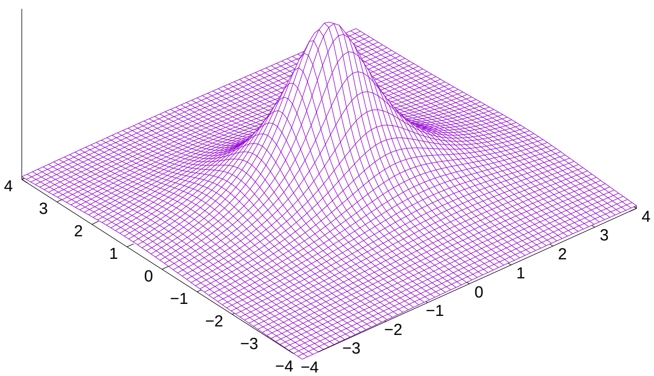
Zatem przy ustalonych wartościach α, $\|\mathbf{p}\|$, zmiana funkcji (30) jest największa, gdy $\measuredangle(\nabla f, \mathbf{p}) = 0$. **Gradient wskazuje kierunek najszybszego wzrostu wartości funkcji.**

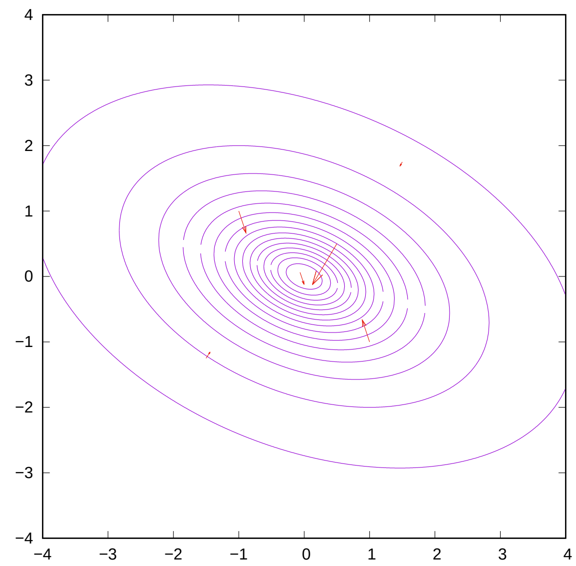
Poziomnice

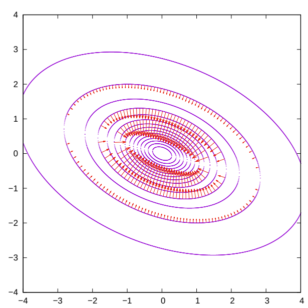
Poziomnice to krzywe, wzdłuż których wartość funkcji wielu zmiennych jest stała.

Z wyrażenia (30) wynika, że jeżeli $(\nabla f) \circ \mathbf{p} = 0$, czyli gdy kierunek gradientu i kierunek lokalnego przesunięcia są prostopadłe, $\Delta f = 0$: Bardzo małe ($|\alpha| \ll 1$) przesunięcie w kierunku prostopadłym do gradientu nie prowadzi do zmiany wartości funkcji. Wynika stąd, że gradient jest (lokalnie) prostopadły do poziomnic funkcji.

Przykład 9







Pochodne cząstkowe wyższych rzędów

Pochodne cząstkowe funkcji N zmiennych same są funkcjami N zmiennych. Można więc definiować *ich* pochodne cząstkowe, czyli pochodne pochodnych. Zmienne mogą się mieszać! Dla funkcji dwóch zmiennych $f(x,y)$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \tag{31a}$$

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \tag{31b}$$

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \tag{31c}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{31d}$$

W sposób oczywisty można to uogólnić na funkcje trzech i więcej zmiennych, a także na pochodne wyższych rzędów.

Przykład 10

Niech

$$f(x,y) = \frac{x-y}{x+y} \tag{32a}$$

Wówczas

$$\frac{\partial f}{\partial x} = \frac{2y}{(x+y)^2} \tag{32b}$$

$$\frac{\partial f}{\partial y} = \frac{-2x}{(x+y)^2} \tag{32c}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{-4y}{(x+y)^3} \tag{32d}$$

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{2(x-y)}{(x+y)^3} \tag{32e}$$

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{2(x-y)}{(x+y)^3} \tag{32f}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{4x}{(x+y)^3} \tag{32g}$$

Funkcje klasy C_2

Jeżeli funkcja dwu zmiennych jest odpowiednio gładka, to
1° jej drugie pochodne cząstkowe są ciągłe
2° pochodne mieszane są sobie równe

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x} \tag{33}$$

i podobnie dla większej liczby zmiennych:

$$\frac{\partial^2 f}{\partial x_i\,\partial x_j} = \frac{\partial^2 f}{\partial x_j\,\partial x_i}\,,\ i,j = 1,\dots,N \tag{34}$$

O funkcjach, których drugie pochodne są ciągłe i które spełniają warunek (34) mówimy, że są klasy C_2.

W ogólności funkcja N zmiennych ma N^2 drugich pochodnych cząstkowych. Jeżeli funkcja jest klasy co najmniej C_2, liczba *niezależnych* drugich pochodnych cząstkowych redukuje się do $N(N-1)/2$.

Różniczka zupełna

Niech $f : \mathbb{R}^N \to \mathbb{R}$ będzie funkcją klasy (co najmniej) C_2. Wyrażenie

$$df = \frac{\partial f}{\partial x_1} \cdot dx_1 + \frac{\partial f}{\partial x_2} \cdot dx_2 + \cdots + \frac{\partial f}{\partial x_N} \cdot dx_N \tag{35}$$

nazywam różniczką zupełną funkcji f w pewnym punkcie, mianowicie w tym, w którym obliczane są wszystkie pochodne cząstkowe. Różniczka zupełna opisuje zmianę funkcji, gdy poszczególne zmnienne *niezależnie* zmieniają się o infinitezymalnie małe wielkości, odpowiednio, $dx_1, dx_2, \ldots, dx_N$.

Zauważmy, że ponieważ df jest różniczką funkcji klasy C_2, pochodne mieszane z założenia są sobie równe.

Różniczka zupełna określa geometrię hiperpłaszczyzny lokalnie stycznej do N wymiarowego “wykresu” funkcji f.

Forma różniczkowa

Uogólnieniem wyrażenia (35) jest wyrażenie

$$\begin{aligned} DQ &= u_1(x_1, x_2, \ldots, x_N)dx_1 + u_2(x_1, x_2, \ldots, x_N)dx_2 + \ldots \\ &+ u_N(x_1, x_2, \ldots, x_N)dx_N \end{aligned} \tag{36}$$

gdzie $u_1, u_2, \ldots, u_N$ są pewnymi funkcjami klasy (co najmniej) C_1. Jeżeli dla każdej pary i, j zachodzi

$$\forall\, i, j = 1, \ldots, N : \frac{\partial u_i}{\partial x_j} = \frac{\partial u_j}{\partial x_i} \tag{37}$$

wyrażenie (36) możemy traktować jako różniczkę zupełną pewnej funkcji; warunek (37) jest wówczas warunkiem równości pochodnych mieszanych (34).

Jeżeli dla choć jednej pary i, j warunek (37) *nie* zachodzi, wyrażenie (36) *nie* jest różniczką zupełną żadnej funkcji. Formę różniczkową (36) nazywam wówczas formą Pfaffa.

Przykład 11

Forma różniczkowa

$$df = \frac{1}{y}dx - \frac{x}{y^2}dy \tag{38a}$$

jest różniczką zupełną, gdyż

$$\frac{\partial}{\partial y}\left(\frac{1}{y}\right) = -\frac{1}{y^2} = \frac{\partial}{\partial x}\left(-\frac{x}{y^2}\right) \tag{38b}$$

$$d\left(\frac{x}{y}\right) = \frac{1}{y}dx - \frac{x}{y^2}dy \tag{39}$$

Proszę to porównać ze wzorem na pochodną zupełną (20).

Czynnik całkujący

Niech (36) będzie formą Pfaffa. Jeżeli istnieje taka funkcja $\mu = \mu(x_1, x_2, \ldots, x_N)$, że wyrażenie

$$\mu(x_1,\ldots,x_N)u_1(x_1,\ldots,x_N)dx_1+\cdots+\mu(x_1,\ldots,x_N)u_N(x_1,\ldots,x_N)dx_N \tag{40}$$

jest różniczką zupełną, to znaczy

$$\forall\, i,j = 1,\ldots,N : \frac{\partial(\mu\cdot u_i)}{\partial x_j} = \frac{\partial(\mu\cdot u_j)}{\partial x_i} \tag{41}$$

formę Pfaffa (36) nazywam całkowalną, a funkcję μ nazywam czynnikiem całkującym.

Twierdzenie: Każda forma Pfaffa dwu zmiennych posiada czynnik całkujący.

Nie każda forma trzech i więcej zmiennych jest całkowalna.

Przykład 12

Forma różniczkowa

$$DQ = \frac{1}{y} dx + \frac{x}{y^2} dy \tag{42a}$$

nie jest różniczką zupełną, gdyż

$$\frac{\partial}{\partial y}\left(\frac{1}{y}\right) = -\frac{1}{y^2} \neq \frac{\partial}{\partial x}\left(\frac{x}{y^2}\right) = \frac{1}{y^2} \tag{42b}$$

Jednak forma Pfaffa (42a) posiada czynnik całkujący $\mu = y^2$. Istotnie, wyrażenie

$$y^2 \cdot \frac{1}{y} dx + y^2 \cdot \frac{x}{y^2} dy = y\,dx + x\,dy \tag{42c}$$

jest różniczką zupełną:

$$d(xy) = y\,dx + x\,dy \tag{42d}$$

Uwaga! Dla form różniczkowych dwu zmiennych

$$DQ = u(x,y)\,dx + v(x,y)\,dy \tag{43}$$

czynnik całkujący *często* (choć nie zawsze!) można znaleźć zakładając, że jest on funkcją tylko jednej zmiennej: $\mu = \mu(x)$ lub $\mu = \mu(y)$.

Twierdzenie:

1° Jeżeli wyrażenie

$$\frac{1}{v}\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right) \tag{44a}$$

jest funkcją wyłącznie zmiennej x, czynnik całkujący $\mu = \mu(x)$ i jest on postaci

$$\mu(x) = \exp\left(\int \frac{1}{v}\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right) dx\right) \tag{44b}$$

2° Jeżeli wyrażenie

$$\frac{1}{u}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right) \tag{45a}$$

jest funkcją wyłącznie zmiennej y, czynnik całkujący $\mu = \mu(y)$ i jest on postaci

$$\mu(y) = \exp\left(\int \frac{1}{u}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right) dy\right) \tag{45b}$$

Istnieją także warunki istnienia bardziej ogólnych form czynników całkujących.

Przykład 13

Czy forma różniczkowa

$$DQ = \frac{y^2}{x}\,dx + y\,dy \tag{46a}$$

jest różniczką zupełną? Liczymy

$$\frac{\partial}{\partial y}\left(\frac{y^2}{x}\right) - \frac{\partial}{\partial x}y = \frac{2y}{x} - 0 = \frac{2y}{x} \neq 0 \tag{46b}$$

a więc podana forma różniczkowa nie jest różniczką zupełną. Jednak

$$\frac{1}{y}\left(\frac{\partial}{\partial y}\left(\frac{y^2}{x}\right) - \frac{\partial}{\partial x}y\right) = \frac{1}{y}\cdot\frac{2y}{x} = \frac{2}{x} \tag{46c}$$

jest funkcją tylko zmiennej x, istnieje czynnik całkujący mający postać

$$\mu = \exp\left(\int \frac{2}{x}\,dx\right) = \exp(2\ln x + \tilde{C}) = C\,\exp\left(\ln x^2\right) = C\,x^2\,, \tag{46d}$$

gdzie $\tilde{C}, C = \exp(\tilde{C})$ są stałymi. Przyjmijmy $C = 2$. Po przemnożeniu przez czynnik całkujący $\mu = 2x^2$, podana forma różniczkowa przybiera postać

$$
\begin{aligned}
df &= 2xy^2\,dx + 2x^2y\,dy && \text{(46e)}\\
f &= x^2y^2 && \text{(46f)}
\end{aligned}
$$

Przykład 14

Czy forma różniczkowa

$$DQ = (x^2 + y)\,dx - x\,dy \tag{47a}$$

jest różniczką zupełną?

Sprawdzamy:

$$\frac{\partial}{\partial y}(x^2 + y) - \frac{\partial}{\partial x}(-x) = 1 - (-1) = 2 \neq 0 \tag{47b}$$

Ponieważ jednak

$$\frac{1}{-x}\left(\frac{\partial}{\partial y}(x^2 + y) - \frac{\partial}{\partial x}(-x)\right) = -\frac{2}{x} \tag{47c}$$

jest funkcją wyłącznie zmiennej x, $\mu = \mu(x)$ oraz

$$\mu = \mathrm{exp}\left(-2\int\frac{dx}{x}\right) = \frac{C}{x^2} \tag{47d}$$

Wyrażenie

$$df = \frac{1}{x^2}\cdot(x^2+y)\,dx + \frac{1}{x^2}\cdot(-x)\,dy = \left(1+\frac{y}{x^2}\right)dx - \frac{1}{x}\,dy \tag{47e}$$

jest różniczką zupełną.

$$f(x,y) = x - \frac{y}{x} \tag{47f}$$

Zadanie

Czy forma różniczkowa

$$DQ = y\,dx - \frac{2x+y}{y^2}\,dy \tag{48}$$

jest różniczką zupełną? Jeśli nie, znajdź jej czynnik całkujący. Bonus: Znajdź funckcję, której różniczka ma postać (48), po ewentualnym przemnożeniu prawej strony przez czynnik całkujący.

Szereg Taylora

Odpowiednio gładkie funkcje wielu zmiennych, podobnie jak funkcje jednej zmiennej, można rozwijać w szereg Taylora. Ze względów praktycznych ograniczymy się do funkcji klasy C_2 i do rozwinięcia do drugiego rzędu.

Niech $f : \mathbb{R}^N \to \mathbb{R}$ będzie funkcją klasy C_2 w otoczeniu pewnego punktu $\mathbf{x}_0 \in \mathbb{R}^N$. Przypuśćmy, że oddalamy się od tego punktu o ε_i w kierunku i-tej współrzędnej:

$$
\begin{aligned}
(x_0)_1 &\to (x_0)_1 + \varepsilon_1 && \text{(49a)}\\
(x_0)_2 &\to (x_0)_2 + \varepsilon_2 && \text{(49b)}\\
&\dots\\
(x_0)_N &\to (x_0)_N + \varepsilon_N && \text{(49c)}
\end{aligned}
$$

$(x_0)_i$ oznacza i-tą współrzędną punktu $\mathbf{x}_0$. Wektor przesunięć $[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]$ oznaczam $\boldsymbol{\varepsilon}$. Wówczas

$$f(\mathbf{x}_0+\boldsymbol{\varepsilon}) \simeq f(\mathbf{x}_0)+\left(\nabla f|_{\mathbf{x}_0}\right)^T \boldsymbol{\varepsilon}+\frac{1}{2}\boldsymbol{\varepsilon}^T\mathbf{A}\boldsymbol{\varepsilon}+O\left(\|\boldsymbol{\varepsilon}\|^3\right) \quad (50)$$

gdzie $\mathbf{A} \in \mathbb{R}^{N\times N}$ jest macierzą drugich pochodnych cząstkowych funkcji f obliczanych w $\mathbf{x}_0$:

$$A_{ij} = \left.\frac{\partial^2 f}{\partial x_i\,\partial x_j}\right|_{\mathbf{x}_0} \quad (51)$$

Macierz tę nazywamy hesjanem lub macierzą Hessego. Ponieważ funkcja jest klasy C_2, hesjan jest macierzą symetryczną.

Minima i maksima funkcji wielu zmiennych

Gradient pokazuje kierunek najszybszego wzrostu funcji. Jeśli $\|\nabla f|_{\mathbf{x}}\| = 0$, funkcja lokalnie, w przybliżeniu liniowych, nie zmienia wartości. W takich punktach funkcja może mieć minimum lub maksimum.

Warunkiem koniecznym istnienia minimum lub maksimum funkcji dwu zmiennych jest

$$\nabla f|_{\mathbf{x}} = 0 \tag{52}$$

Warunki istnienia ekstremów

W otoczeniu piunktu, w którym gradient znika, o zachowaniu funkcji decydują algebraiczne własności hesjanu: Jeżeli $\nabla f|_{\mathbf{x}_0} = 0$,

$$f(\mathbf{x}_0 + \boldsymbol{\varepsilon}) \simeq f(\mathbf{x}_0) + \frac{1}{2}\boldsymbol{\varepsilon}^T \mathbf{A} \boldsymbol{\varepsilon} \tag{53}$$

gdzie $\mathbf{A}$ jest hesjanem w $\mathbf{x_0}$.

Minimum

Jeżeli w punkcie, w którym $\nabla f|_{\mathbf{x}_0} = 0$, hesjan jest dodatnio określony, z warunku (53) wynika, że małe odchylenie się od tego punktu w dowolnym kierunku powoduje, że wartości funkcji rosną: funkcja ma zatem w tym punkcie minimum.

Ponieważ hesjan jest macierzą symetryczną, rzeczywistą, warunek dodatniej określoności jest rónoważny temu, że wszystkie wartości własne hesjanu są większe od zera.

Zauważmy, że jeśli funkcja jest klasy C_2, a więc jej drugie pochodne cząstkowe są ciągłe, to jeżeli hesjan jest dodatnio określony w pewnym punkcie $\mathbf{x}$ (niekoniecznie w minimum), istnieje otoczenie tego punktu, w którym hesjan także jest dodatnio określony.

Przykład 15

Funkcja Rosenbrocka:

$$f(x,y) = (1-x)^2 + 100(y-x^2)^2 \tag{54a}$$

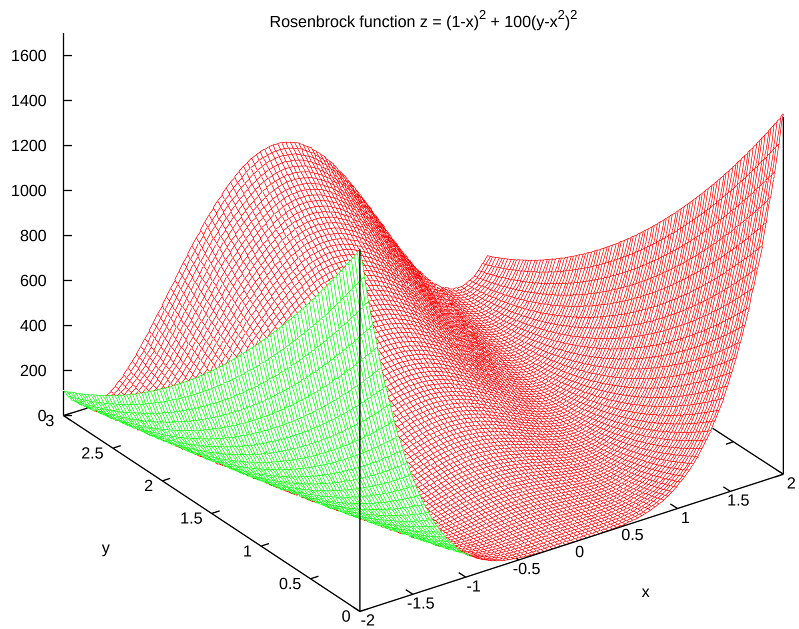
$$\nabla f = [-2(1-x) - 400x(y-x^2), 200(y-x^2)] \tag{54b}$$

Jak łatwo się przekonać, punkt $(x,y) = (1,1)$ jest jedynym punktem, w którym ∇f znika.

Hesjan ma postać

$$\mathbf{A} = \begin{bmatrix} 2+800x & -400x \\ -400x & 200 \end{bmatrix}_{(1,1)} = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix} \tag{54c}$$

a jego wartościami własnymi są $\lambda_1 = 501 + \sqrt{250\,601} \simeq 1001.6$, $\lambda_2 = 501 - \sqrt{250\,601} \simeq 0.4$. Obie wartości włąsne są dodatnie, zatem hesjan jest dodatnio określony, zatem funkcja Rosenbrocka osiąga w punkcie $(1,1)$ minimum.


Rosenbrock function z = (1-x)^2 + 100(y-x^2)^2
1600
1400
1200
1000
800
600
400
200
0
3
2.5
2
1.5
1
0.5
0
y
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
x

Maksimum

Jeżeli w punkcie, w którym $\nabla f|_{\mathbf{x}_0} = 0$, hesjan jest ujemnie określony, czyli wszystkie jego wartości własne są ujemne, małe odchylenie się od tego punktu w dowolnym kierunku powoduje, że wartości funkcji maleją: funkcja ma zatem w tym punkcie maksimum.

W przypadku funkcji wielu zmiennych możliwe są sytuacje nie mające swoich odpowiedników dla funkcji jednej zmiennej.

Punkt siodłowy

Jeżeli hesjan ma zarówno dodatnie, jak i ujemne wartości własne, o charakterze zmienności funkcji w otoczeniu punktu będącego potencjalnym ekstremum, decyduje wektor ε: w kierunkach odpowiadających wektorom własnym hesjanu do dodatnich wartości własnych funkcja rośnie, w kierunkach odpowiadających wektorom własnym hesjanu do ujemnych wartości własnych funkcja maleje. Punkt taki nazywamy punktem siodłowym.

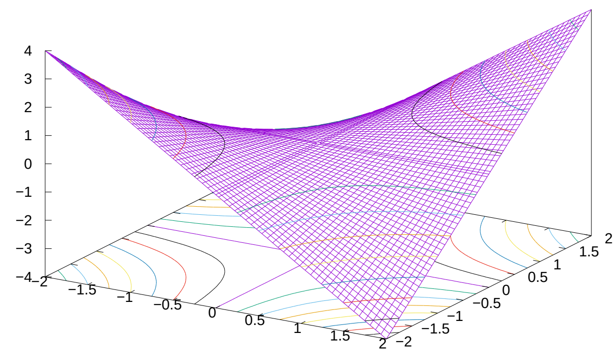
Przykład 16

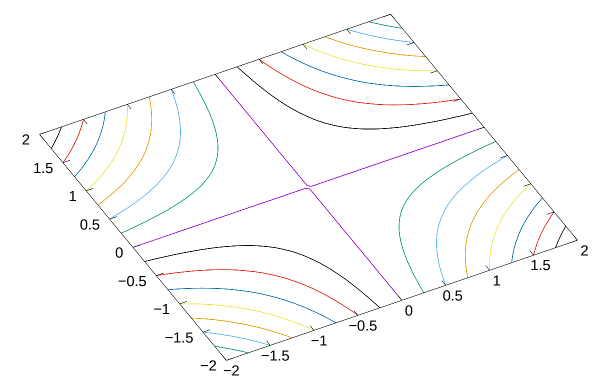
$$f(x,y) = xy\,, \quad \nabla f = [y,x] \tag{55a}$$

$\nabla f = 0 \Leftrightarrow (x,y) = (0,0)$. Hesjan ma postać

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{55b}$$

a jego wartości własne wynoszą $\lambda_{1,2} = \pm 1$. Punkt $(0,0)$ jest punktem siodłowym


4
3
2
1
0
−1
−2
−3
−4
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2


2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2

Mod zerowy

Jeżli w jakimś punkcie $\nabla f|_{\mathbf{x}_0} = 0$, a przy tym hesjan ma zerową wartość własną, oznacza to, że istnieje kierunek — odpowiadający kierunkowi wektora własnego hesjanu do zerowej wartości własnej — wzdłuż którego funkcja nie zmienia swojej wartości. W żargonie kierunek taki nazywa się “modem zerowym”.

Przykład 17

Rozpatrzmy funkcję nazywaną “potencjałem dna butelki”:

$$f(x, y) = a(x^2 + y^2 - 1)^2\,, \; a > 0 \tag{56a}$$

$$\nabla f = [4ax(x^2 + y^2 - 1), 4ay(x^2 + y^2 - 1)] \tag{56b}$$

Gradient znika w punkcie $(0, 0)$ *oraz* na okręgu $x^2 + y^2 = 1$. Hesjan ma postać

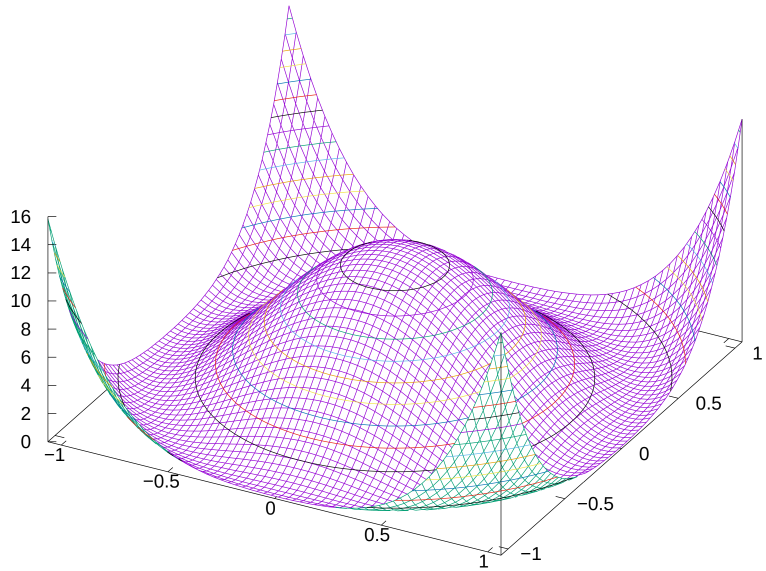
$$\mathbf{A} = 4a \begin{bmatrix} 3x^2 + y^2 - 1 & 2xy \\ 2xy & x^2 + 3y^2 - 1 \end{bmatrix} \tag{56c}$$

a jego wartości własne wynoszą

$$\lambda_1 = 4a(x^2 + y^2 - 1) \tag{56d}$$

$$\lambda_2 = 4a(3(x^2 + y^2) - 1) \tag{56e}$$

W punkcie $(0,0)$ obie wartości własne $\lambda_{1,2} = -4a < 0$ i punkt ten odpowiada maksimum. Natomiast w dowolnym punkcie okręgu $x^2 + y^2 = 1$ $\lambda_1 = 0, \lambda_2 = 8a$. Gdyby funkcja (56a) była potencjałem, cząstka w tym potencjale mogłaby poruszać się po okręgu $x^2 + y^2 = 1$ bez straty (i bez zyskiwania) energii.



Potencjał (56a) z $a = 10$

Podsumowanie

	funkcja jednej zmiennej	funkcja wielu zmiennych
Ekstremum	$\dfrac{df}{dx} = 0$	$\nabla f = 0$
Minimum	$\dfrac{d^2 f}{dx^2} > 0$	hesjan dodatnio określony
Maksimum	$\dfrac{d^2 f}{dx^2} < 0$	hesjan ujemnie określony
Punkt siodłowy	—	dodatnie i ujemne wartości własne
Mod zerowy	—	zerowa wartość własna

Ekstrema wyższych rzędów

Podobnie jak w przypadku funkcji jednej zmiennej, może się zdarzyć, że badanie macierzy drugich pochodnych cząstkowych nie rozstrzyga, czy funkcja przybiera maksimum bądź minimum. Jest to jednak przypadek rzadki i, formalnie, dość trudny do zapisania w sposób zwarty, dlatego nie będziemy o tym szczegółowo mówić.

Przykład 18

Funkcja

$$f(x, y) = x^4 + y^4 \tag{57}$$

ma w punkcie $(0, 0)$ minimum, o czym decyduje znak pochodnych $\partial^4 f/\partial x^4$, $\partial^4 f/\partial y^4$. Wszystkie pochodne cząstkowe niższego rzędu oraz wszystkie pochodne mieszane czwartego rzędu w tym punkcie znikają.