Fizyka dla firm — Matematyka

28. Ekstrema funkcji wielu zmiennych

Funkcje uwikłane i ekstrema warunkowe

P. F. Góra

https://zfs.fais.uj.edu.pl/pawel_gora

22 marca 2023

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:

$$
\begin{aligned}
(x_0)_1 &\to (x_0)_1 + \varepsilon_1 && \text{(1a)}\\
(x_0)_2 &\to (x_0)_2 + \varepsilon_2 && \text{(1b)}\\
&\ldots \\
(x_0)_N &\to (x_0)_N + \varepsilon_N && \text{(1c)}
\end{aligned}
$$

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

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

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

$$A_{ij} = \left.\frac{\partial^2 f}{\partial x_i \, \partial x_j}\right|_{\mathbf{x}_0} \quad (3)$$

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

$$\nabla f|_{\mathbf{x}} = 0 \tag{4}$$

Warunki istnienia ekstremów

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

$$f(\mathbf{x}_0 + \boldsymbol{\varepsilon}) \simeq f(\mathbf{x}_0) + \frac{1}{2}\boldsymbol{\varepsilon}^T \mathbf{A}\boldsymbol{\varepsilon} \tag{5}$$

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 (5) 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 1

Funkcja Rosenbrocka:

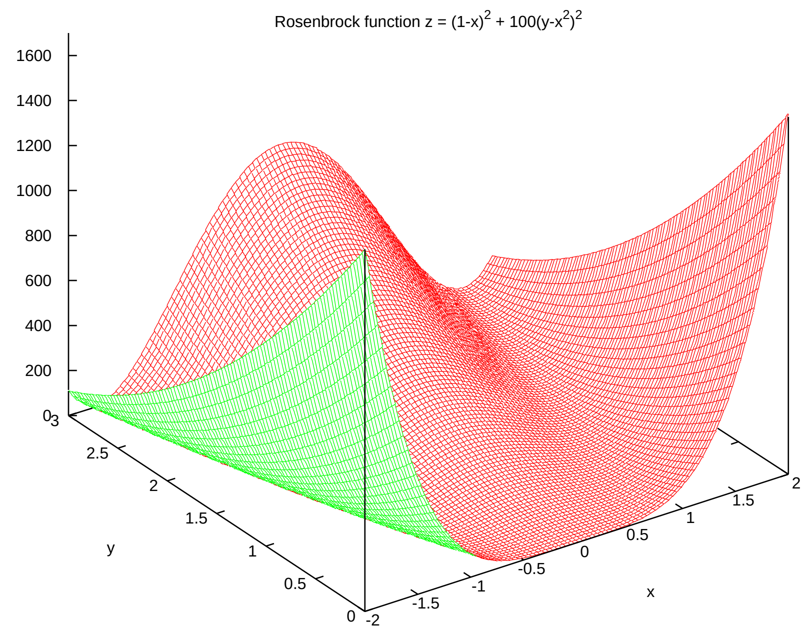
$$f(x,y) = (1-x)^2 + 100(y-x^2)^2 \tag{6a}$$
$$\nabla f = [-2(1-x) - 400x(y-x^2), 200(y-x^2)] \tag{6b}$$

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

Hesjan ma postać

$$\mathbf{A} = \left[\begin{array}{cc} 2+800x & -400x \\ -400x & 200 \end{array}\right]_{(1,1)} = \left[\begin{array}{rr} 802 & -400 \\ -400 & 200 \end{array}\right] \tag{6c}$$

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 własne 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.

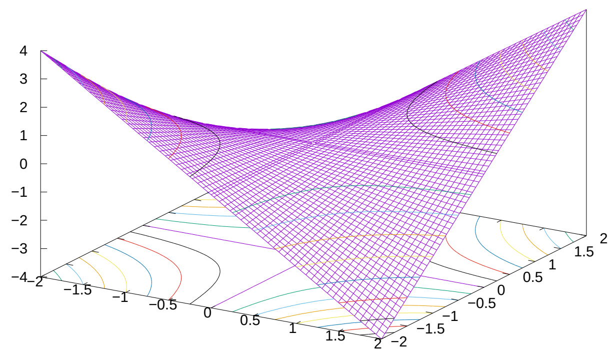
Przykład 2

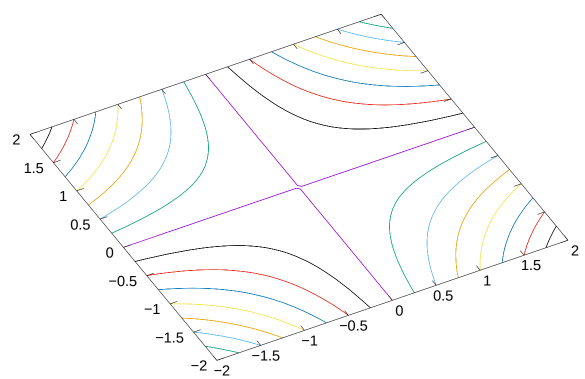
$$f(x,y) = xy\,, \quad \nabla f = [y,x] \tag{7a}$$

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

$$\mathbf{A} = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right] \tag{7b}$$

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 3

Rozpatrzmy funkcję nazywaną “potencjałem dna butelki”:

$$f(x,y) = a(x^2+y^2-1)^2\,,\ a>0 \tag{8a}$$

$$\nabla f = [4ax(x^2+y^2-1), 4ay(x^2+y^2-1)] \tag{8b}$$

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

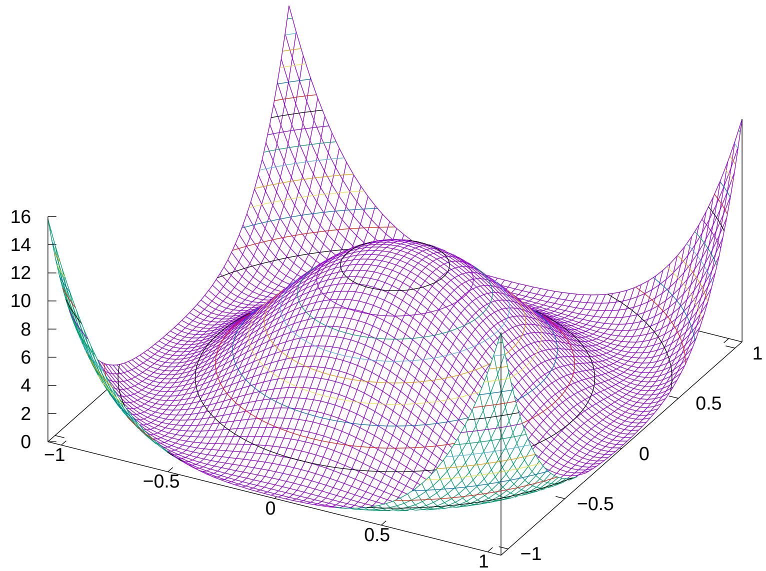
$$\mathbf{A} = 4a\begin{bmatrix} 3x^2+y^2-1 & 2xy \\ 2xy & x^2+3y^2-1 \end{bmatrix} \tag{8c}$$

a jego wartości własne wynoszą

$$\lambda_1 = 4a(x^2+y^2-1) \tag{8d}$$

$$\lambda_2 = 4a(3(x^2+y^2)-1) \tag{8e}$$

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 (8a) była potencjałem, cząstka w tym potencjale mogłaby poruszać się po okręgu $x^2 + y^2 = 1$ bez straty (i bez zyskiwania) energii.



Potencjał (8a) 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 4

Funkcja

$$f(x,y) = x^4 + y^4 \tag{9}$$

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.

Przekształcenia płaszczyzny w siebie

Przypuśćmy, że dane jest przekształcenie $\mathbb{R}^2 \to \mathbb{R}^2$

$$x = f(u,v) \tag{10a}$$

$$y = g(u,v) \tag{10b}$$

przy czym funkcje f, g są określone i ciągłe w pewnym obszarze W. Jeśli punkt $P(u,v) \in W$, to punkt $Q(f(u,v), g(u,v))$ nazywam obrazem punktu P w tym przekształceniu.

Przykład 5

Niech

$$x = u^2 , \quad y = v^2 \tag{11a}$$

Obrazem koła

$$u^2 + v^2 \leqslant 4 \tag{11b}$$

w przekształceniu (11a) jest trójkąt

$$0 \leqslant x + y \leqslant 4 , \; x \geqslant 0 , \; y \geqslant 0 . \tag{11c}$$

Jakobian

Weźmy przekształcenie (10) i załóżmy, że funkcje f, g są (co najmniej) klasy C_1. Wyznacznik

$$\frac{D(f,g)}{D(u,v)} = \begin{vmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{vmatrix} \tag{12}$$

nazywam jakobianem odwzorowania (10).

Homeomorfizm i dyfeomorfizm

Twierdzenie: Jeżeli w pewnym punkcie $(u_0, v_0) \in W$ jakobian (12) jest różny od zera

$$\left.\frac{D(f,g)}{D(u,v)}\right|_{(u_0,v_0)} \neq 0 \tag{13}$$

to istnieje otoczenie U punktu $Q_0(x_0, y_0) = Q(f(u_0, v_0), g(u_0, v_0))$, w którym odwzorowanie (10) jest odwracalne, to znaczy istnieją funkcje jednoznaczne h, k takie, że

$$u = h(x, y)\,, \quad v = k(x, y)\,. \tag{14}$$

Jeżeli odwzorowanie (10) jest odwracalne na całej płaszczyźnie $\mathbb{R}^2$ i odwzorowanie odwrotne jest ciągłe, odwzorowanie to nazywam homeomorfizmem. Jeżeli ponadto odwzorowanie odwrotne jest co najmniej klasy C_1, odwzorowanie (10) nazywam dyfeomorfizmem.

Uogólnienie

Powyższe rozważania można uogólnić i zapisać w bardziej zwartej formie. Niech $\mathbf{f} : \mathbb{R}^N \to \mathbb{R}^N$ będzie funkcją klasy co najmniej C_1. Dla $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^N$

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_N) \\ f_2(x_1, x_2, \dots, x_N) \\ \vdots \\ f_N(x_1, x_2, \dots, x_N) \end{bmatrix} \in \mathbb{R}^N \tag{15}$$

i funkcje $f_1, f_2, \dots, f_N$ są klasy co najmniej C_1 jako funkcje N zmiennych rzeczywistych ($\forall i = 1, \dots, N : f_i : \mathbb{R}^N \to \mathbb{R}$).

Macierz pochodnych czątkowych

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_N} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1} & \frac{\partial f_N}{\partial x_2} & \dots & \frac{\partial f_N}{\partial x_N} \end{bmatrix} = \left[\frac{\partial f_i}{\partial x_j}\right]_{i,j=1,\dots,N} \tag{16}$$

nazywam macierzą Jacobiego funkcji (15). Wówczas jakobian jest wyznacznikiem macierzy (16), $J = \det \mathbf{J}$.

Pojęcia homeomorfizmu i dyfeomorfizmu uogólniają się odpowiednio: Jeśli funkcja $\mathbf{f}$ (15) jest odwracalna na całej $\mathbb{R}^N$, a jej odwrotność jest ciągła, nazywamy ją homeomorfizmem. Jeśli funkcja (15) jest (co najmniej) klasy C_1, jest odwracalna, a jej odwrotność też jest klasy C_1, nazywamy ją dyfeomorfizmem.

Przykład 6

Współrzędne biegunowe i ich jakobian:

$$\begin{cases} x & = r\cos\varphi \\ y & = r\sin\varphi \end{cases} \tag{17a}$$

$$J = \frac{D(x,y)}{D(r,\varphi)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r \tag{17b}$$

Przykład 7

Współrzędne cylindryczne i ich jakobian:

$$\begin{cases} x &= r\cos\varphi \\ y &= r\sin\varphi \\ z &= z \end{cases} \tag{18a}$$

$$J = \frac{D(x,y,z)}{D(r,\varphi,z)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r \tag{18b}$$

Przykład 8

Współrzędne sferyczne i ich jakobian:

$$\begin{cases} x &= r\cos\theta\cos\phi \\ y &= r\cos\theta\sin\phi \\ z &= r\sin\theta \end{cases} \tag{19a}$$

$$
\begin{aligned}
J = \frac{D(x,y,z)}{D(r,\theta,\phi)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} \\
&= \begin{vmatrix} \cos\theta\cos\phi & -r\sin\theta\cos\phi & -r\cos\theta\sin\phi \\ \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta & r\cos\theta & 0 \end{vmatrix} \\
&= 0 - r^2\sin^2\theta\cos\theta\cos^2\phi - r^2\cos^3\theta\sin^2\phi \\
&\quad - r^2\sin^2\theta\cos\theta\sin^2\phi - r^2\cos^3\theta\cos^2\phi - 0 \\
&= -r^2\cos\theta \qquad (19b)
\end{aligned}
$$

Funkcje uwikłane

Niech dane będzie wyrażenie

$$F(x, y) = 0 \tag{20}$$

Jeżeli daje się ono “rozwiązać” w sposób **jednoznaczny** ze względu na y, to znaczy przekształcić do postaci

$$y = f(x) \tag{21}$$

przy czym równość

$$F(x, f(x)) \equiv 0 \tag{22}$$

zachodzi w sposób tożsamościowy, mówimy, że wyrażenie (20) zadaje funkcję $y = y(x)$ w sposób niejawny, a wyrażenie (21) jest jego *jawną* reprezentacją.

Przykład 9

Wyrażenia

$$2x + 3y - 1 = 0 \tag{23a}$$

$$y = \frac{1}{3} - \frac{2}{3}x \tag{23b}$$

są niejawną i jawną reprezentacją tej samej funkcji.

Podobnie

$$\exp\left(\frac{x-y}{x+y}\right) = 5 \tag{24a}$$

$$y = \frac{1-\ln 5}{1+\ln 5}\,x \tag{24b}$$

Rozwiązanie równania (20) ze względu na y niekiedy nie istnieje (np. $x^2 + y^2 + 1 = 0$), niekiedy rozwiązanie nie jest jednoznaczne, niekiedy zaś nie da się go przedstawić w postaci analitycznej (jawnej). W tych dwóch ostatnich przypadkach mówimy, że wyrażenie (20) zadaje funkcję $y = y(x)$ w sposób uwikłany.

Przykład 10

$$x^2 + y^2 - 1 = 0 \;\Rightarrow\; y = \pm\sqrt{1 - x^2} \tag{25}$$

Przykład 11

$$e^{x+y} - \sin(x \cdot y) - 1 = 0 \;\Rightarrow\; \text{nie ma rozwiązania w postaci jawnej} \tag{26}$$

Twierdzenie o funkcjach uwikłanych

Twierdzenie: Jeżeli istnieje punkt (x_0, y_0) taki, że $F(x_0, y_0) = 0$ i funkcja $F(x, y)$ jest w otoczeniu tego punktu ciągła i ma w otoczeniu tego punktu ciągłą pochodną $\frac{\partial F}{\partial y}$, przy czym $\left.\frac{\partial F}{\partial y}\right|_{x_0, y_0} \neq 0$, to w otoczeniu punktu (x_0, y_0) istnieje funkcja $y = f(x)$ taka, że $F(x, f(x)) \equiv 0$ oraz $y_0 = f(x_0)$.
Ponadto jeśli funkcja $F(x, y)$ ma w otoczeniu punktu (x_0, y_0) ciągłą pochodną $\frac{\partial F}{\partial x}$, to

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \tag{27}$$

Przykład 12

Rozważmy wyrażenie

$$F(x,y) = x^2 + y^2 - 1 = 0 \tag{28a}$$

Punkt $\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$ spełnia równanie (28a). Ponadto pochodna cząstkowa $\frac{\partial F}{\partial y} = 2y$ przybiera w tym punkcie wartość $-\sqrt{3} \neq 0$, a zatem wyrażenie (28a) można w otoczeniu tego punktu rozwikłać ze względu na y:

$$y = -\sqrt{1-x^2} \tag{28b}$$

Wybór punktu leżącego na konkretnym półokręgu ustalił znak w wyrażeniu (25).

Przykład 13

Nadal rozważamy wyrażenie (28a), ale tym razem bierzemy punkt $(1, 0)$. Punkt ten leży na okręgu, ale w tym punkcie pochodna cząstkowa $\frac{\partial F}{\partial y} = 0$ i wyrażenia (28a) nie można w otoczeniu tego punktu rozwikłać ze względu na y.

Przykład 14

Obliczmy pochodną $\frac{dy}{dx}$, jeżeli $y = y(x)$ jest dana w postaci uwikłanej za pomocą wyrażenia

$$F(x, y) = x\,e^y - y + 1 = 0 \tag{29a}$$

Przede wszystkim

$$\frac{\partial F}{\partial y} = x\,e^y - 1 = y - 2 \tag{29b}$$

gdyż z równania (29a) wynika, że $x\,e^y = y - 1$. Stąd wniosek, że wyrażenie (29a) daje się rozwikłać ze względu na y poza prostą $y = 2$. Dalej,

$$\frac{\partial F}{\partial x} = e^y \tag{29c}$$

zatem

$$\frac{dy}{dx} = -\frac{e^y}{y - 2} = \frac{e^y}{2 - y} \tag{29d}$$

Ten sam wynik uzyskalibyśmy różniczkując (29a) wyraz po wyrazie:

$$\frac{d}{dx}\left(x\,e^y - y + 1\right) = \frac{d}{dx}0 \tag{29e}$$

$$\frac{d}{dx}\left(x\,e^y\right) - \frac{dy}{dx} = 0 \tag{29f}$$

$$e^y + x\,e^y\frac{dy}{dx} - \frac{dy}{dx} = 0 \tag{29g}$$

$$-\frac{dy}{dx}\left(1 - x\,e^y\right) + e^y = 0 \tag{29h}$$

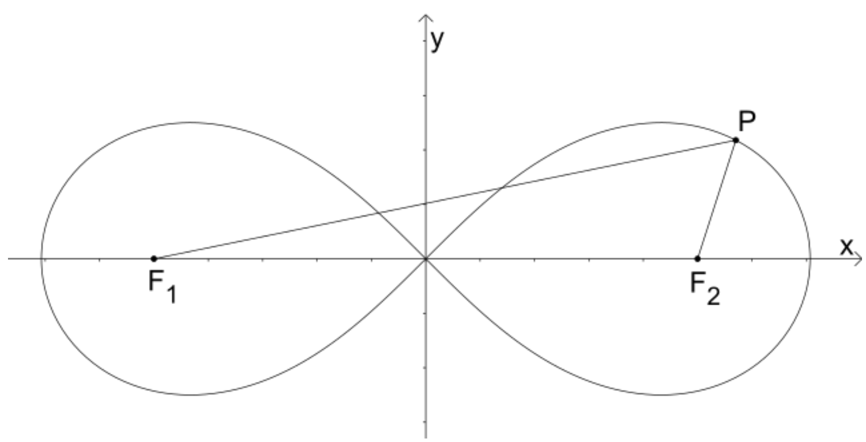
$$\frac{dy}{dx} = \frac{e^y}{1 - x\,e^y} = \frac{e^y}{1-(y-1)} = \frac{e^y}{2-y} \tag{29i}$$

Przykład 15

Lemniskata opisana jest równaniem

$$F(x, y) = (x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0 \tag{30a}$$

Czy z tego wyrażenia daje się odwikłać $y = y(x)$? Ile wynosi pochodna $\frac{dy}{dx}$?



$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}(x^4+2x^2y^2+y^4-2a^2x^2+2a^2y^2) = 4y(x^2+y^2+a^2) \quad \text{(30b)}$$

Widzimy, że $\partial F/\partial y \neq 0$ dla $y \neq 0$, gdyż wyrażenie w nawiasie jest dodatnie. Wyrażenia (30a) nie daje się odwikłać ze względu na y w punktach $y = 0$.

Różniczkując (30a) po x otrzymujemy

$$2(x^2+y^2)\left(2x+2y\frac{dy}{dx}\right) - 2x^2\left(2x-2y\frac{dy}{dx}\right) \quad \text{(30c)}$$

$$\frac{dy}{dx} = \frac{x}{y}\cdot\frac{a^2-x^2-y^2}{a^2+x^2+y^2} \quad \text{(30d)}$$

Pochodna (30d) **nie istnieje** dla $y = 0$, czyli dla $x = 0$, $x = \pm a\sqrt{2}$. Można pokazać, że w punktach $x = \pm a\sqrt{2}$ pochodna (30d) jest rozbieżna. Możemy natomiast zbadać zachowanie pochodnej w pobliżu $x = 0$.

Zauważmy, że

$$\lim_{(x,y)\to(0,0)} \frac{a^2 - x^2 - y^2}{a^2 + x^2 + y^2} = 1 \tag{30e}$$

a zatem w otoczeniu punktu $(0, 0)$ pochodna (30d) zachowuje się jak

$$\frac{dy}{dx} \simeq \frac{x}{y} \tag{30f}$$

co jest spełnione dla

$$y^2 = x^2 + C\,, \tag{30g}$$

jednak ponieważ $y(0) = 0$, musi zachodzić $C = 0$. Wybór znaku pierwiastka przy rozwikłaniu wyrażenia (30g) ze względu na y odpowiada wyborowi odpowiedniej gałęzi lemniskaty. Jeśli wybierzemy znak $+$, czyli gałąź $y > 0$ dla $x > 0$, otrzymujemy

$$\lim_{x \to 0^+, y > 0} \frac{dy}{dx} = 1 \tag{30h}$$

Dla drugiej gałęzi otrzymalibyśmy

$$\lim_{x \to 0^+, y < 0} \frac{dy}{dx} = -1 \tag{30i}$$

Uogólnienie na więcej zmiennych

Rozważmy funkcję $F : \mathbb{R}^N \to \mathbb{R}$ i niech

$$F(x_1, x_2, \dots, x_N) = 0 \tag{31}$$

Jeżeli istnieje punkt $(x_{1_0}, x_{2_0}, \dots, x_{N_0})$ spełniający równanie (31), przy czym funkcja F jest w otoczeniu tego punktu ciągła oraz w tym punkcie $\frac{\partial F}{\partial x_N} \neq 0$, wyrażenie (31) daje się rozwikłać ze względu na x_n w otoczeniu tego punktu, to znaczy istnieje funkcja $f : \mathbb{R}^{N-1} \to \mathbb{R}$ taka, że

$$x_N = f(x_1, x_2, \dots, x_{N-1}) \tag{32}$$

oraz

$$F(x_1, x_2, \dots, x_{N-1}, f(x_1, x_2, \dots, x_{N-1})) \equiv 0 \tag{33}$$

Ponadto jeśli pozostałe pochodne cząstkowe istnieją i są ciągłe,

$$\frac{\partial f}{\partial x_i} = -\frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial x_N}} \tag{34}$$

.

Przykład 16

Niech

$$F(x,y,z) = xy + xz + yz - 3 = 0 \tag{35a}$$

Punkt $(1,1,1)$ spełnia równanie (35a). W tym punkcie $\frac{\partial F}{\partial z} = 2 \neq 0$, a więc równanie (35a) daje się rozwikłać ze względu na z. Różniczkując wyrażenie (35a), odpowiednio, po x i po y otrzymujemy

$$y + z + x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial x} = 0 \;\Rightarrow\; \frac{\partial z}{\partial x} = -\frac{y+z}{x+y} \tag{35b}$$

$$x + x\frac{\partial z}{\partial y} + z + y\frac{\partial z}{\partial y} = 0 \;\Rightarrow\; \frac{\partial z}{\partial y} = -\frac{x+z}{x+y} \tag{35c}$$

co jest zgodne z (34).

Ekstrema funkcji uwikłanej

Niech będzie spełnione równanie (20) oraz założenia twierdzenia o funkcjach uwikłanych. Wówczas funkcję $y = y(x)$ daje się rozwikłać i zachodzi równanie (27). Aby funkcja $y = y(x)$ mogło mieć ekstremum, muszą być spełnione oba warunki

$$F(x, y) = 0 \tag{36a}$$

$$\frac{\partial F}{\partial x} = 0 \tag{36b}$$

Aby zbadać, czy rozwiązania układu równań (36) w istocie są ekstremami, a jeśli tak, to czy minimami, czy maksimami, musimy obliczyć drugą pochodną, $\frac{d^2y}{dx^2}$. Formalnie dla funkcji dwóch zmiennych, x, y, przy czym $y = y(x)$, operator różniczkowania

$$\frac{d}{dx} = \frac{\partial}{\partial x} + \frac{dy}{dx} \cdot \frac{\partial}{\partial y} \tag{37}$$

Różniczkując wyrażenie (20), otrzymujemy

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx} = 0 \tag{38}$$

równoważne warunkowi (27). Różniczkując jeszcze raz

$$\frac{d}{dx}\left(\frac{\partial F}{\partial x}+\frac{\partial F}{\partial y}\cdot\frac{dy}{dx}\right)=0 \tag{39a}$$

$$\frac{d}{dx}\frac{\partial F}{\partial x}+\frac{d}{dx}\left(\frac{\partial F}{\partial y}\cdot\frac{dy}{dx}\right)=0 \tag{39b}$$

$$\frac{d}{dx}\frac{\partial F}{\partial x}+\frac{d}{dx}\left(\frac{\partial F}{\partial y}\right)\cdot\frac{dy}{dx}+\frac{\partial F}{\partial y}\cdot\frac{d^2y}{dx^2}=0 \tag{39c}$$

$$\left(\frac{\partial}{\partial x}+\frac{dy}{dx}\frac{\partial}{\partial y}\right)\frac{\partial F}{\partial x}+\left[\left(\frac{\partial}{\partial x}+\frac{dy}{dx}\frac{\partial}{\partial y}\right)\frac{\partial F}{\partial y}\right]\frac{dy}{dx}+\frac{\partial F}{\partial y}\cdot\frac{d^2y}{dx^2}=0 \tag{39d}$$

$$\frac{\partial^2 F}{\partial x^2}+\frac{\partial^2 F}{\partial y\partial x}\cdot\frac{dy}{dx}+\left(\frac{\partial^2 F}{\partial x\partial y}+\frac{\partial^2 F}{\partial y^2}\cdot\frac{dy}{dx}\right)\frac{dy}{dx}+\frac{\partial F}{\partial y}\cdot\frac{d^2y}{dx^2}=0 \tag{39e}$$

$$\frac{\partial^2 F}{\partial x^2}+2\frac{\partial^2 F}{\partial x\partial y}\cdot\frac{dy}{dx}+\frac{\partial^2 F}{\partial y^2}\left(\frac{dy}{dx}\right)^2+\frac{\partial F}{\partial y}\cdot\frac{d^2y}{dx^2}=0 \tag{39f}$$

gdzie skorzystaliśmy z równości pochodnych mieszanych. Zatem

$$\frac{d^2y}{dx^2} = -\frac{1}{\frac{\partial F}{\partial y}} \left(\frac{\partial^2 F}{\partial x^2} + 2\frac{\partial^2 F}{\partial x \partial y} \cdot \frac{dy}{dx} + \frac{\partial^2 F}{\partial y^2} \left(\frac{dy}{dx} \right)^2 \right) \tag{40}$$

Jeżeli $\frac{dy}{dx} = 0$, wyrażenie to upraszcza się do

$$\left.\frac{d^2y}{dx^2}\right|_{\frac{dy}{dx}=0} = -\frac{\frac{\partial^2 F}{\partial x^2}}{\frac{\partial F}{\partial y}} \tag{41}$$

Przykład 17

Zbadajmy ekstrema funkcji uwikłanej $y = y(x)$ danej w sposób niejawny za pomocą równania

$$F(x, y) = y^4 - 8xy - 4y + 8x^2 = 0 \tag{42a}$$

Musimy rozwiązać układ równań składający się z (42a) oraz równania

$$\frac{\partial F}{\partial x} = -8y + 16x = -8(y - 2x) = 0 \tag{42b}$$

skąd $x = \frac{1}{2}y$. Podstawiając do (42a) otrzymujemy

$$y^4 - 8 \cdot \frac{1}{2} \cdot y^2 - 4y + 8\left(\frac{1}{2}\right)^2 y^2 = 0 \tag{42c}$$

$$y^4 - 4y^2 - 4y + 2y^2 = 0 \tag{42d}$$

$$y^4 - 2y^2 - 4y = 0 \tag{42e}$$

$$y(y^3 - 2y - 4) = 0 \tag{42f}$$

Rzeczywistymi pierwiastkami tego równania są $y = 0 \;\Rightarrow\; x = 0$ oraz $y = 2 \;\Rightarrow\; x = 1$.

Musimy jeszcze obliczyć

$$\frac{\partial^2 F}{\partial x^2} = 16 \tag{42g}$$

$$\frac{\partial F}{\partial y} = 4y^3 - 8x - 4 = 4(y^3 - 2x - 1) \tag{42h}$$

W punkcie $(0,0)$ $\partial F/\partial y = -4 \neq 0$, $\partial^2 F/\partial x^2 = 16$, $\frac{d^2y}{dx^2} = -\frac{16}{-4} = 4 > 0$ i funkcja uwikłana ma tam minimum.

W punkcie $(1,2)$ $\partial F/\partial y = 4(8-2-1) = 20 \neq 0$, $\partial^2 F/\partial x^2 = 16$, $\frac{d^2y}{dx^2} = -\frac{16}{20} = -\frac{4}{5} < 0$ i funkcja uwikłana ma tam maksimum.

Ekstrema warunkowe

Niech funkcje $f(x,y), g(x,y)$ będą określone i ciągłe w pewnym obszarze D. Mówimy, że funkcja $f(x,y)$ osiąga w puncie (x_0,y_0) ekstremum warunkowe przy warunku $g(x,y) = 0$, jeżeli $g(x_0,y_0) = 0$ oraz istnieje takie otoczenie punktu (x_0,y_0), że dla każdego punktu (x,y) należącego do tego otoczenia i spełniającego warunek $g(x,y) = 0$ zachodzi $f(x,y) \leqslant f(x_0,y_0)$ (maksimum) lub $f(x,y) \geqslant f(x_0,y_0)$ (minimum).

Twierdzenie: Jeżeli funkcje f, g są klasy (co najmniej) C_1, warunkiem koniecznym istnienia ekstremum w punkcie (x_0,y_0) jest aby jakobian

$$\left.\frac{D(f,g)}{D(x,y)}\right|_{x_0,y_0} = 0 \tag{43}$$

Z powyższego twierdzenia wynika, że punkty, w których mogą istnieć ekstrema warunkowe, znajdujemy rozwiązując układ równań

$$
\begin{aligned}
g(x,y) &= 0 && \text{(44a)}\\
\frac{D(f,g)}{D(x,y)} &= 0 && \text{(44b)}
\end{aligned}
$$

Metoda mnożników Lagrange'a

Wprowadzam funkcję pomocniczą

$$F(x,y) = f(x,y) + \lambda g(x,y) \tag{45}$$

znajduję warunki konieczne na istnienie ekstremum funkcji $F(x,y)$, czyli zerowanie się gradientu ∇F, a następnie eliminuję czynnik nieoznaczony λ. Widać, że jest to równoważne warunkom (44).

Uogólnienie na N wymiarów jest oczywiste: Jeżeli szukam ekstremum funkcji $f(x_1,\dots,x_N)$ przy warunku $g(x_1,\dots,x_N) = 0$, wprowadzam funkcję

$$F(x_1,\dots,x_N) = f(x_1,\dots,x_N) + \lambda g(x_1,\dots,x_N) \tag{46}$$

i eliminuję λ z układu równań

$$\nabla F(x_1,\dots,x_N) = 0 \tag{47a}$$

$$g(x_1,\dots,x_N) = 0 \tag{47b}$$

Przykład 18

Znajdźmy ekstrema funkcji $f(x, y) = xy$ przy warunku $x^2 + y^2 = 1$.

W tym celu tworzę funkcję

$$F(x, y) = xy + \lambda(x^2 + y^2 - 2) \tag{48a}$$

a następnie rozwiązuję układ równań

$$\begin{cases} \frac{\partial F}{\partial x} = y + 2\lambda x & = 0 \\ \frac{\partial F}{\partial y} = x + 2\lambda y & = 0 \\ x^2 + y^2 & = 1 \end{cases} \tag{48b}$$

Z pierwszego z tych równań otrzymuję $\lambda = -\frac{y}{2x}$. Po podstawieniu do drugiego równania daje $x^2 - y^2 = 0$. Ostatecznie otzrymuję układ dwu równań z dwiema niewiadomymi

$$\begin{cases} x^2 - y^2 & = 0 \\ x^2 + y^2 & = 1 \end{cases} \tag{48c}$$

którego rozwiązaniami są *cztery* punkty $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$. Łatwo sprawdzić, że $f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{1}{2}$, co odpowiada maksimum, natomiast $f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -\frac{1}{2}$, co odpowiada minimum.