
ROZDZIAŁ 3

Funkcje wielu zmiennych

3.1 Wprowadzenie

Zajmiemy się teraz uogólnieniem pojęcia funkcji na przypadek, kiedy jest kilka zmiennych niezależnych. Na przykład, objętość walca $V = \pi R^2 H$ jest funkcją dwóch zmiennych, a stożka ściętego $V = \frac{1}{3}\pi H(R^2 + Rr + r^2)$ nawet trzech.

Nasze rozważania rozpoczniemy od najprostszego przypadku funkcji dwóch zmiennych rzeczywistych $f(x, y)$. Dziedziną D będzie zbiór par (x, y) będący podzbiorem $\mathbb{R} \times \mathbb{R}$, na którym funkcja $f(x, y)$ jest określona, to znaczy jest pewnym sposobem przyporządkowania każdej parze $(x, y) \in D$ pewnej wartości $z = f(x, y)$. W dalszym ciągu ograniczymy się do przypadku funkcji jednoznacznych o wartościach w zbiorze liczb rzeczywistych.

Przykłady:

1. $z = \arcsin(\frac{x}{a}) + \arcsin(\frac{y}{b})$. Wówczas $D = \{(x, y) : -a \leq x \leq a, -b \leq y \leq b\}$
2. Rozważmy trójkąty o długościach dwóch bokach x, y oraz o długości obwodu równej $2p$. Ze szkoły wiemy, że pole trójkąta wyraża się wzorem $\mathcal{P} = \sqrt{p(p-x)(p-y)(x+y-p)}$, bo trzeci bok wynosi $z = 2p - x - y$. Powyższy wzór pozwala łatwo zdefiniować obszar zmienności argumentów (dziedzinę): $0 < x < p, 0 < y < p, x + y > p$.

W przypadku funkcji n zmiennych dziedziną M będzie zbiór n -wymiarowych punktów $x = (x_1, \dots, x_n)$ będący podzbiorem $X = \mathbb{R}^n$. Wprowadzimy w tej przestrzeni miarę (patrz przestrzenie metryczne - wcześniej). Odległość dwóch punktów x i x' określimy zgodnie z metryką euklidesowską, która w bazie ortonormalnej ma postać

$$d(x, x') = \sqrt{\sum_{i=1}^n (x'_i - x_i)^2}.$$

Funkcja ta jest dodatnio określona, spełnia nierówność trójkąta i wynosi zero jedynie dla $x = x'$, a więc spełnia wszelkie własności wymagane od metryki.

W przestrzeni metrycznej (X, d) możemy wprowadzić pojęcie otoczenia-kuli $K(x, r)$ o promieniu r i środku x jako zbioru punktów y takich, że $d(x, y) < r$. Przypomnijmy sobie kilka definicji.

Podzbiór $\mathcal{O}$ przestrzeni metrycznej nazywamy *otwartym* jeśli jest pusty albo gdy dla każdego $x \in \mathcal{O}$ istnieje kula o promieniu ϵ i środku w x zawarta w $\mathcal{O}$, tzn. $K(x, \epsilon) \subset \mathcal{O}$.

Zbiór D jest *domknięty*, gdy jego dopełnienie $X - D$ jest zbiorem otwartym.

Punkt $y \in X$ jest *punktem skupienia* zbioru A , jeśli każde otoczenie punktu y zawiera co najmniej jeden punkt zawarty w A różny od y . Innymi słowy, punkt y jest granicą ciągu elementów z A różnych od y .

Punkt skupienia zbioru A nie musi należeć do tego zbioru. Jeśli A zawiera wszystkie swe punkty skupienia, to jest domknięty.

Zbiór nazywamy *spójnym*, jeśli dwa dowolne punkty tego zbioru można połączyć krzywą, której wszystkie punkty należą do tego zbioru.

3.2 Granica funkcji wielu zmiennych

Definicja granicy funkcji:

Niech funkcja f jest określona na zbiorze punktów $x \in M$ i $x^0 = (x_1^0, \dots, x_n^0)$ jest punktem skupienia M . W analogii do definicji granicy funkcji jednej zmiennej, mówimy, że liczba A jest granicą funkcji f , gdy zmienne $x_1 \dots x_n$ dążą do $x_1^0 \dots x_n^0$, jeżeli

$$\epsilon > 0 \quad \forall \quad \delta > 0 \quad d(x, x_0) < \delta \quad |f(x_1, \dots, x_n) - A| < \epsilon.$$

Granice tę będziemy oznaczali w następujący sposób

$$\lim_{\substack{x_1 \rightarrow x_1^0 \\ \vdots \\ x_n \rightarrow x_n^0}} f(x_1, \dots, x_n) = A$$

lub prosto $\lim_{x \rightarrow x^0} f(x) = A$ pamiętając, że x oznacza zbiór n zmiennych.

Analogicznie wprowadzamy pojęcie nieskończonej granicy funkcji. W wypadku $A = \infty$ lub $A = -\infty$ nierówność $|f(x_1, \dots, x_n) - A| < \epsilon$ zastępujemy przez $|f(x_1, \dots, x_n)| > E$ lub $|f(x_1, \dots, x_n)| < -E$, gdzie E jest dowolną, zadaną liczbą dodatnią.

Tę definicję można również uogólnić na przypadek punktu skupienia, którego niektóre lub wszystkie współrzędne są nieskończone, analogicznie do przypadku funkcji jednej zmiennej.

Przykłady:

1. Rozpatrzmy

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \frac{xy}{x^2 + y^2}.$$

Funkcja ta jest określona na całej płaszczyźnie z wyjątkiem punktu $(0,0)$. Jeżeli weźmiemy dwa ciągi punktów $(x_k, y_k) = (\frac{1}{k}, \frac{1}{k})$ oraz $(x_k, y_k) = (\frac{2}{k}, \frac{1}{k})$, oba zbiegające do $(0,0)$, to dla pierwszego

ciągu dostajemy

$$f(x_k, y_k) = \frac{\frac{1}{k^2}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2},$$

natomiast dla drugiego

$$f(x_k, y_k) = \frac{\frac{2}{k^2}}{\frac{4}{k^2} + \frac{1}{k^2}} = \frac{2}{5}.$$

Jak widać, wynik zależy od sposobu dochodzenia do punktu $(0,0)$. Dlatego granica tej funkcji w $(0,0)$ nie istnieje!

2. Rozpatrzmy

$$f(x, y) = \frac{x^2 y}{x^2 + y^2}.$$

Teraz mamy

$$\lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} f(x, y) = 0$$

i granica istnieje. Aby to zobaczyć zapiszemy tę funkcję w postaci

$$\frac{x^2 y}{x^2 + y^2} = \frac{x}{\frac{x}{y} + \frac{y}{x}},$$

Oznaczając $x/y = a$, łatwo sprawdzić, że funkcja w mianowniku $f(a) = |\frac{1}{a} + a|$ ma minimum w punkcie $a = 1$ równe 2. Zatem $|\frac{1}{a} + a| \geq 2$. Stąd dostajemy

$$\left| \frac{x^2 y}{x^2 + y^2} \right| \leq \frac{1}{2} |x|,$$

i z tej nierówności widać już, że granica jest równa zero.

3.2.1 Granice iterowane

W powyższej definicji granicy funkcji $f(x_1, \dots, x_n)$ wszystkie zmienne $x_1, \dots, x_n$ dążyły do granicy $x_1^0, \dots, x_n^0$. Takie granice będziemy też nazywać n -krotnymi (podwójnymi, potrójnymi itd. – dla $n = 2, 3, \dots$). Możemy postawić pytanie, czy wykonując przejścia do granicy oddzielnie dla poszczególnych zmiennych w tej lub innej kolejności dostalibyśmy ten sam wynik. Dla prostoty ograniczymy się do funkcji dwóch zmiennych.

Wprowadzimy pojęcie *granicy iterowanej*, gdy obliczamy najpierw granicę w jednej zmiennej, a następnie w drugiej. Dla dwóch zmiennych mamy dwie możliwości: $\lim_{x \rightarrow a} f(x, y) = f(a, y)$ a następnie przechodząc do granicy $\lim_{y \rightarrow b} f(a, y)$. Albo odwrotnie, najpierw zbiegamy z $y \rightarrow b$, albo potem $x \rightarrow a$. Okazuje się, że dwie granice iterowane $\lim_{x \rightarrow a} \lim_{y \rightarrow b} f(x, y)$ i $\lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x, y)$ nie muszą być sobie równe. Może też się zdarzyć, że jedna granica iterowana istnieje, a druga nie. Ilustrują to następujące

Przykłady:

1. Niech

$$f(x, y) = \frac{x - y + x^2 + y^2}{x + y}.$$

Obliczamy granice iterowane w $(x_0, y_0) = (0, 0)$:

$$f(0, y) = \lim_{x \rightarrow 0} f(x, y) = y - 1, \quad \lim_{y \rightarrow 0} f(0, y) = -1$$

oraz

$$f(x, 0) = \lim_{y \rightarrow 0} f(x, y) = x + 1, \quad \lim_{x \rightarrow 0} f(x, 0) = 1.$$

2. Niech teraz

$$f(x, y) = \frac{x \sin \frac{1}{x} + y}{x + y}.$$

Tym razem granica iterowana w $(0, 0)$

$$\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} f(x, y) = \lim_{y \rightarrow 0} \frac{y}{y} = 1,$$

istnieje, natomiast druga granica iterowana

$$\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y) = \lim_{x \rightarrow 0} \sin \frac{1}{x}$$

nie istnieje.

3. Weźmy funkcję

$$f(x, y) = x \sin \frac{1}{y}.$$

Funkcja ta jest określona na całej płaszczyźnie z wyjątkiem prostej $y = 0$. Funkcja ta ma granicę podwójną w punkcie $(0, 0)$ równą zero, co widać z nierówności

$$\left| x \sin \frac{1}{y} \right| \leq |x|.$$

Istnieje również granica iterowana $\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} f(x, y)$, ale druga granica iterowana nie istnieje, bo nie istnieje nawet granica pojedyncza $\lim_{y \rightarrow 0} f(x, y)$.

Jak widać, musimy zachować ostrożność przy obliczaniu granic iterowanych. Nawet istnienie granicy n -krotnej funkcji (w przykładzie 3 powyżej – podwójnej) nie musi gwarantować istnienia granic iterowanych. Pomocne tutaj jest

Twierdzenie:

Jeśli istnieje granica podwójna

$$\lim_{\substack{x \rightarrow a \\ y \rightarrow b}} f(x, y) = A$$

oraz istnieje granica

$$\phi(y) = \lim_{x \rightarrow a} f(x, y),$$

to istnieje również granica iterowana

$$\lim_{y \rightarrow b} \phi(y)$$

i jest równa A .

Dowód: Ponieważ istnieje granica podwójna, to dla każdego $\epsilon > 0$ istnieje δ , taka że $|f(x, y) - A| < \epsilon$ o ile $|x - a| < \delta$ oraz $|y - b| < \delta$. Ustalimy teraz y tak, by spełniona była nierówność $|y - b| < \delta$ i przejdziemy teraz z $x \rightarrow a$. Ponieważ przy tym przejściu $f(x, y)$ dąży do $\phi(y)$, to

$$|\phi(y) - A| < \epsilon,$$

co oznacza, że

$$\lim_{y \rightarrow b} \phi(y) = \lim_{y \rightarrow b} \lim_{x \rightarrow a} f(x, y) = A.$$

cbdo

Jeżeli spełnione są założenia powyższego twierdzenia oraz istnieje granica $f(x, b) = \lim_{y \rightarrow b} f(x, y)$, to obie granice iterowane istnieją i są sobie równe. W przykładach podanych powyżej, z wyjątkiem przykładu 3, nie istnieje granica podwójna. W przykładzie 3 granica podwójna istnieje, ale nie istnieje $\lim_{y \rightarrow 0} f(x, y)$ dla $x \neq 0$, więc i granica iterowana nie istnieje.

3.3 Funkcje ciągłe

Definicja ciągłości w punkcie:

Niech funkcja $f(x_1, \dots, x_n)$ będzie określona w pewnym zbiorze D punktów przestrzeni n -wymiarowej i punkt $x^0 = (x_1^0, \dots, x_n^0)$ będzie punktem skupienia D i należącym do D . Mówimy, że funkcja jest ciągła w punkcie $x^0 = (x_1^0, \dots, x_n^0)$, jeśli zachodzi równość

$$\lim_{\substack{x_1 \rightarrow x_1^0 \\ \vdots \\ x_n \rightarrow x_n^0}} f(x) = f(x^0).$$

W przeciwnym razie mówimy, że funkcja f ma w punkcie x^0 nieciągłość.

Mówimy, że funkcja jest ciągła w pewnym zbiorze D , jeśli jest ciągła w każdym punkcie tego zbioru, który jest punktem skupienia D .

Podaliśmy już powyżej przykład funkcji nieciągłej w $(0, 0)$: $f(x, y) = \frac{xy}{x^2 + y^2}$. Istnieją oczywiście granice wzdłuż osi x oraz y : $\lim_{x \rightarrow 0} f(x, y) = \lim_{y \rightarrow 0} f(x, y) = 0$, ale istnieje nieskończenie wiele sposobów zbliżania się do zera i dla tej funkcji wynik zależy od sposobu podchodzenia do $(0, 0)$.

Zbiór punktów nieciągłości nie musi być zbiorem izolowanych punktów, lecz może wypełniać linie, powierzchnie itp. Na przykład, dla funkcji $f(x) = \frac{x^2 + y^2}{x^2 - y^2}$ punkty nieciągłości leżą na prostych $x = \pm y$, a dla funkcji $f(x) = \frac{1}{x^2 + y^2 - 1}$ na okręgu $x^2 + y^2 = 1$.

Na funkcje wielu zmiennych przenoszą się w naturalny sposób poznane wcześniej twierdzenia o ciągłości sumy, różnicy, iloczynie, ilorazie i superpozycji funkcji ciągłych.

Jeżeli funkcja jest ciągła w pewnym spójnym obszarze D i w dwóch różnych punktach x i y tego obszaru przyjmuje różne wartości $f(x) \neq f(y)$, np. niech $f(x) < f(y)$, to dla każdej

wartości pośredniej c , $f(x) < c < f(y)$, istnieje punkt $z \in D$, taki że $f(z) = c$. Innymi słowy, funkcja ciągła w obszarze spójnym przyjmuje wszystkie wartości pośrednie.

Na koniec podamy jeszcze bez dowodu **twierdzenie Weierstrassa**: Jeśli funkcja $f(x)$ jest określona i ciągła w określonym obszarze domkniętym D , to jest ona ograniczona.

3.4 Pochodne cząstkowe

Przejdziemy teraz do badania możliwości różniczkowania funkcji wielu zmiennych. Zmieniając tylko jedną zmienną i traktując $n-1$ pozostałych zmiennych jako ustalone, możemy wykorzystać definicję pochodnej funkcji jednej zmiennej do zdefiniowania pochodnej cząstkowej. Wystarczy rozpatrzyć przypadek funkcji dwóch zmiennych.

Definicja pochodnej cząstkowej:

Niech $f(x, y)$ będzie określona na otoczeniu otwartym punktu (x_0, y_0) . Funkcja zdefiniowana relacją $x \mapsto f(x, y_0)$ jest funkcją jednej zmiennej x w otoczeniu punktu $x_0 \in \mathbb{R}^1$. Możemy zatem obliczyć pochodną względem x :

$$\lim_{x \rightarrow x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = \left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)}$$

Jeżeli ta granica (pochodna) istnieje, to nazywamy ją pochodną cząstkową w punkcie (x_0, y_0) względem x . Używane są następujące oznaczenia pochodnych cząstkowych

$$\frac{\partial f}{\partial x}, \quad \frac{\partial}{\partial x} f, \quad \partial_x f, \quad f_{,x}$$

Odwzorowanie $(x, y) \rightarrow \frac{\partial f}{\partial x}(x, y)$, jeżeli istnieje w pewnym otoczeniu $\mathcal{O}$, nazywamy pochodną cząstkową względem x . Podobnie definiujemy

$$\left. \frac{\partial f}{\partial y} \right|_{(x_0, y_0)} = \lim_{y \rightarrow y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}.$$

Analogicznie definiujemy pochodne cząstkowe funkcji większej liczby zmiennych.

Przykłady:

1. Niech $u = x^y$ ($x > 0$). Pochodne cząstkowe tej funkcji są następujące

$$\frac{\partial f}{\partial x} = yx^{y-1}, \quad \frac{\partial f}{\partial y} = x^y \ln x.$$

2. Niech $z = yf(x^2 - y^2)$, gdzie $f(u)$ jest dowolną funkcją różniczkowalną. Pokazać, że spełniona jest równość

$$\frac{1}{x} \frac{\partial z}{\partial x} + \frac{1}{y} \frac{\partial z}{\partial y} = \frac{z}{y^2}.$$

Obliczamy pochodne cząstkowe traktując $f(x^2 - y^2)$ jako funkcję złożoną:

$$\frac{\partial z}{\partial x} = 2xyf', \quad \frac{\partial z}{\partial y} = f - 2yf',$$

gdzie f' oznacza pochodną f względem $u = x^2 - y^2$. Wstawiając te pochodne cząstkowe do powyższego równania przekonujemy się, że jest ono prawdziwe.

3. Znany z fizyki wzór Clapeyrona $pV = RT$ wyraża związek między ciśnieniem p , objętością V i temperaturą absolutną jednej gramocząsteczki gazu doskonałego (R jest uniwersalną stałą gazową). Wzór ten wyraża jedną wielkość fizyczną jako funkcję dwóch pozostałych. Jeśli potraktujemy p, V jako niezależne, to $T = pV/R$ i wtedy

$$\frac{\partial T}{\partial p} = \frac{V}{R}, \quad \frac{\partial T}{\partial V} = \frac{p}{R}.$$

Gdy traktujemy p i T jako zmienne niezależne, to $V = RT/p$ i

$$\frac{\partial V}{\partial p} = -\frac{RT}{p^2}, \quad \frac{\partial V}{\partial T} = \frac{R}{p}.$$

Wreszcie, gdy V i T są zmiennymi niezależnymi, to $p = RT/V$ i

$$\frac{\partial p}{\partial V} = -\frac{RT}{V^2}, \quad \frac{\partial p}{\partial T} = \frac{R}{V}.$$

Stąd dostajemy ważną dla termodynamiki zależność

$$\frac{\partial p}{\partial V} \frac{\partial V}{\partial T} \frac{\partial T}{\partial p} = -1$$

Przykład ten pokazuje, że oznaczenia pochodnych cząstkowych należy traktować symbolicznie, a nie jak wielkości, które można “upraszczać”.

Definicje pochodnej cząstkowej dla funkcji dwóch zmiennych łatwo jest uogólnić na przypadek funkcji n -zmiennych, na przykład:

$$\frac{\partial f}{\partial x_i} = \lim_{h \rightarrow 0} \frac{f(x_1 \dots x_i + h \dots x_n) - f(x_1 \dots x_i \dots x_n)}{h}.$$

3.5 Ciągłość funkcji różniczkowalnej cząstkowo

Dla funkcji jednej zmiennej różniczkowalność w x_0 wymagała określoności funkcji w pewnym otoczeniu otwartym punktu x_0 i ciągłości w x_0 . W przypadku funkcji wielu zmiennych pochodne cząstkowe mogą istnieć nawet, gdy funkcja nie jest określona, co ilustruje poniższy

Przykład:

Funkcja

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

jak wiemy jest nieciągła w $(0,0)$. Łatwo to widać we współrzędnych biegunowych: $x = r \cos \varphi$ oraz $y = r \sin \varphi$. Wtedy $f(r, \varphi) = \frac{1}{2} \sin 2\varphi$ nie zależy od r i przyjmuje różne wartości w zależności pod jakim kątem zbliżamy się do zera. Jeśli tę funkcję dookreślimy w $(0,0)$, jak powyżej, to funkcja ta jest różniczkowalna cząstkowo, bo

$$\frac{\partial f}{\partial x} = \frac{f(x, 0) - f(0, 0)}{x} = 0, \quad \frac{\partial f}{\partial y} = \frac{f(0, y) - f(0, 0)}{y} = 0.$$

Istnienie pochodnych cząstkowych gwarantuje nam jedynie, że funkcja $f(x, y)$ zachowuje się regularnie wzdłuż osi x lub osi y .

3.6 Pochodne cząstkowe wyższych rzędów

Podobnie jak w przypadku pochodnych zwyczajnych funkcji jednej zmiennej, także w przypadku pochodnych cząstkowych możemy obliczać pochodne wyższych rzędów, na przykład pochodne cząstkowe drugiego rzędu po x i y definiujemy jako

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} = f_{,xx}, \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2} = f_{,yy}$$

oraz tzw. pochodne mieszane drugiego rzędu

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y} = f_{,yx}, \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x} = f_{,xy}$$

Powstaje pytanie, czy pochodna mieszana nie zależy od kolejności obliczania pochodnych cząstkowych. Rozpatrzmy dwa

Przykłady :

1. Obliczmy pochodne cząstkowe do drugiego rzędu włącznie z funkcji $f = x^3 y^2 + x^4 \sin y - \sin xy$

$$\begin{aligned} \frac{\partial f}{\partial x} &= 3x^2 y^2 + 4x^3 \sin y - y \cos xy, & \frac{\partial f}{\partial y} &= 2x^3 y + x^4 \cos y - x \cos xy, \\ \frac{\partial^2 f}{\partial x^2} &= 6xy^2 + 12x^2 \sin y + y^2 \sin xy, & \frac{\partial^2 f}{\partial y^2} &= 2x^3 - x^4 \sin y + x^2 \sin xy \end{aligned}$$

oraz pochodne mieszane

$$\frac{\partial^2 f}{\partial x \partial y} = 6x^2 y + 4x^3 \cos y + xy \cos xy - \sin xy = \frac{\partial^2 f}{\partial y \partial x}.$$

W tym przykładzie wynik nie zależy od kolejności obliczania pochodnych cząstkowych. Ale tak nie musi być, o czym przekonuje następny przykład.

2. Rozważmy funkcję

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

Funkcja ta ma granicę w punkcie $(0, 0)$. Łatwo się o tym przekonać przechodząc do współrzędnych biegunowych $x = r \cos \varphi$, $y = r \sin \varphi$, gdyż wtedy widać, że licznik jest rzędu 4 a mianownik rzędu 2 w r i funkcja dąży do zera wraz z $r \rightarrow 0$. Funkcja jest więc ciągła w tym punkcie i możemy ją dookreślić zgodnie z podanym wyżej wzorem. Obliczmy teraz pochodne cząstkowe

$$\begin{aligned}\left. \frac{\partial f}{\partial x} \right|_{0,y} &= \lim_{x \rightarrow 0} \frac{f(x, y) - f(0, y)}{x} = \lim_{x \rightarrow 0} \frac{y(x^2 - y^2)}{x^2 + y^2} = -y \\ \left. \frac{\partial f}{\partial y} \right|_{x,0} &= \lim_{y \rightarrow 0} \frac{f(x, y) - f(x, 0)}{y} = \lim_{y \rightarrow 0} \frac{x(x^2 - y^2)}{x^2 + y^2} = x.\end{aligned}$$

Stąd dostajemy

$$\frac{\partial^2 f}{\partial x \partial y} = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}.$$

Zatem nie zawsze pochodne mieszane muszą być sobie równe. Warto też sprawdzić, że poza punktem $(0, 0)$ pochodne mieszane są równe.

Jeżeli funkcja zachowuje się “regularnie” – kolejność liczenia pochodnych nie gra roli. O tym traktuje następujące

Twierdzenie:

Jeżeli funkcja $f(x, y)$ posiada pochodne cząstkowe pierwszego rzędu na pewnym obszarze otwartym $\mathcal{O} \subset \mathcal{R}^2$ oraz pochodną mieszaną $f_{,xy}$ ciągłą w tym obszarze, to istnieje pochodna $f_{,yx}$ w tym obszarze i obie pochodne są sobie równe.

Dowód: Niech $P_0 = (x_0, y_0) \in \mathcal{O}$. Badamy istnienie granicy

$$\begin{aligned}\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{x_0, y_0} &= \left. \frac{\partial}{\partial x} \left(\frac{\partial}{\partial y} f(x, y_0) \right) \right|_{x_0} = \left. \frac{\partial}{\partial x} \lim_{h \rightarrow 0} \frac{f(x, y_0 + h) - f(x, y_0)}{h} \right|_{x_0} \\ &= \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{1}{hk} (f(x_0 + k, y_0 + h) - f(x_0 + k, y_0) - f(x_0, y_0 + h) + f(x_0, y_0)) \\ &= \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{1}{hk} (\phi(x_0 + k) - \phi(x_0)),\end{aligned}$$

gdzie wprowadziliśmy funkcję $\phi(x) = f(x, y_0 + h) - f(x, y_0)$. Funkcja ta jest różniczkowalna na $[x_0, x_0 + k]$ względem x oraz $\phi'(x) = f_{,x}(x, y_0 + h) - f_{,x}(x, y_0)$ na mocy założenia twierdzenia. Z twierdzenia o wartości średniej Lagrange’a

$$\phi(x_0 + k) - \phi(x_0) = k\phi'(x_0 + \Theta k) = k(f_{,x}(x_0 + k\Theta, y_0 + h) - f_{,x}(x_0 + k\Theta, y_0)).$$

A zatem

$$f_{,yx}(x_0, y_0) = \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{1}{h} (f_{,x}(x_0 + \Theta k, y_0 + h) - f_{,x}(x_0 + \Theta k, y_0)).$$

Dalej wiemy, że $\Psi(y) = f_{,x}(x_0 + \Theta k, y)$ jest różniczkowalna na $[y_0, y_0 + h]$ i z twierdzenia Lagrange’a istnieje λ takie że

$$\Psi(y_0 + h) - \Psi(y_0) = h\Psi'(y_0 + \lambda h) = hf_{,xy}(x_0 + \Theta k, y_0 + \lambda h).$$

Ostatecznie

$$f_{,yx}(x_0, y_0) = \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} f_{,xy}(x_0 + \Theta k, y_0 + \lambda h) = f_{,xy}(x_0, y_0).$$

Ta ostatnia równość wynika z ciągłości pochodnej cząstkowej mieszanej $f_{,xy}$.

cbdo

Uwaga

Korzystaliśmy tutaj z pojęcia granicy iterowanej, to znaczy $\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y)$. Jeżeli funkcja jest ciągła – ma zatem granicę w (x_0, y_0) , to wartość tej granicy nie zależy od drogi, a więc także od kolejności obliczania granic.

Powyższe twierdzenie można uogólnić na wyższe rzędy pochodnej. Jeżeli $f(x_1 \dots x_n)$ jest określona na otwartym zbiorze w $\mathcal{R}^n$ i posiada pochodne cząstkowe rzędu $p - 1$ oraz wszystkie pochodne cząstkowe rzędu p są ciągłe na tym zbiorze, to wszystkie pochodne cząstkowe rzędu p nie zależą od kolejności obliczania pochodnych cząstkowych.

3.7 Pochodne cząstkowe funkcji złożonej

Definicja:

Niech funkcje $u = \varphi(x, y)$ i $v = \psi(x, y)$ będą określone w pewnym obszarze D zmiennych x, y , i niech zbiór wartości u, v należy do obszaru E , w którym określona jest funkcja $z = f(u, v)$. Wtedy

$$z = f(\varphi(x, y), \psi(x, y))$$

określa w obszarze D funkcję złożoną zmiennych x, y .

Posługując się definicją pochodnej cząstkowej łatwo wykazać następujące

Twierdzenie:

Jeśli funkcje $\varphi(x, y)$ i $\psi(x, y)$ mają pochodne cząstkowe w punkcie x_0, y_0 , a funkcja $f(u, v)$ ma pochodne cząstkowe *ciągłe* w punkcie $u_0 = \varphi(x_0, y_0)$, $v_0 = \psi(x_0, y_0)$, to funkcja złożona z ma pochodne cząstkowe wyrażone wzorami

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x}, \quad \frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y}$$

Dowód tego twierdzenia pominiemy. Podkreślimy jedynie istotną rolę założenia ciągłości pochodnych cząstkowych, o czym przekonamy się w przykładzie 3 poniżej.

Przykłady:

1. Funkcja $z = f(x, y)$ jest funkcją złożoną współrzędnych biegunowych r, φ punktu x, y . Obliczamy pochodne cząstkowe względem r i φ korzystając z $x = r \cos \varphi$, $y = r \sin \varphi$

$$\frac{\partial z}{\partial r} = f_{,x} \cos \varphi + f_{,y} \sin \varphi, \quad \frac{\partial z}{\partial \varphi} = -f_{,x} r \sin \varphi + f_{,y} r \cos \varphi.$$

2. Równanie falowe dla funkcji $u(x, t)$ w jednej zmiennej przestrzennej x i czasie t ma postać

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}.$$

Wprowadzimy nowe zmienne “na stożku” zdefiniowane $\xi = x + ct$, $\eta = x - ct$. Obliczając drugie pochodne i wstawiając dostajemy

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0$$

skąd widać, że ogólnym rozwiązaniem tego równania jest superpozycja fal (płaskich) $u = f(\xi) + g(\eta)$, gdzie f i g są dowolnymi funkcjami różniczkowalnymi swoich argumentów.

3. Rozpatrzmy jeszcze raz funkcję

$$f(x, y) = \frac{x^2 y}{x^2 + y^2} \quad (\text{dla } x^2 + y^2 > 0), \quad f(0, 0) = 0.$$

Wiemy, że funkcja ta jest ciągła i ma pochodne cząstkowe we wszystkich punktach nie wyłączając początku układu współrzędnych $(0, 0)$, przy czym

$$f_{,x}(0, 0) = 0, \quad f_{,y}(0, 0) = 0.$$

Zauważmy jednak, że pochodne cząstkowe (jako funkcje dwóch zmiennych) nie są ciągłe w $(0, 0)$. Założenia powyższego twierdzenia o pochodnych funkcji złożonej nie są spełnione i nie można go stosować.

Żeby się o tym przekonać, wprowadzimy nową zmienną t przyjmując $x = t$ i $y = t$ otrzymując funkcję złożoną zmiennej t . Stosując wzór z powyższego twierdzenia pochodna tej funkcji w punkcie $t = 0$ byłaby równa

$$f_{,t} = f_{,x}x_{,t} + f_{,y}y_{,t} = 0.$$

Z drugiej strony, wstawiając jawnie t do tej funkcji dostajemy

$$f(x(t), y(t)) = \frac{t}{2}$$

i pochodna tej funkcji po t jest różna od zera.

3.8 Pochodne kierunkowe. Gradient

Pochodne cząstkowe charakteryzują zmianę funkcji wraz ze zmianą jednej ze zmiennych. Na przykład, pochodna cząstkowa po x funkcji $f(x, y)$ mówi o szybkości zmian tej funkcji wraz ze zmianą x przy ustalonej wartości y . Pochodne cząstkowe ∂_x i ∂_y możemy nazwać pochodnymi kierunkowymi w kierunku osi x i y . Pojęcie to możemy uogólnić na dowolny kierunek na płaszczyźnie Oxy .

Niech funkcja $z = f(x, y)$ będzie określona w pewnym otoczeniu otwartym punktu (x_0, y_0) , a p będzie półprostą wychodzącą z tego punktu sparametryzowaną za pomocą parametru $t \geq 0$ w następujący sposób

$$x = x_0 + t\alpha, \quad y = y_0 + t\beta,$$

gdzie $\alpha^2 + \beta^2 = 1$. Jeśli przestrzeń zmiennych (x, y) potraktujemy jako przestrzeń wektorową, to liczby α i β mają interpretację składowych kierunkowego wektora $\vec{p} = (\alpha, \beta)$ o długości 1 określającego kierunek tej półprostej na płaszczyźnie Oxy .

Definicja pochodnej kierunkowej:

Pochodną kierunkową w punkcie (x_0, y_0) w kierunku $\vec{p}$ funkcji $f(x, y)$, którą będziemy oznaczali przez $\nabla_p f$, nazywamy granicę jednostronną (jeśli ona istnieje)

$$\nabla_p f|_{x_0, y_0} = \lim_{t \rightarrow 0+} \frac{f(x_0 + t\alpha, y_0 + t\beta) - f(x_0, y_0)}{t}.$$

Powyższa granica jest pochodną zwyczajną (jednostronną) funkcji złożonej $f(x_0 + t\alpha, y_0 + t\beta)$ jednej zmiennej t w punkcie $t = 0$ i redukuje się do pochodnej cząstkowej $\partial_x f$ w kierunku półosi dodatniej x dla $\alpha = 1, \beta = 0$ i do pochodnej cząstkowej $\partial_y f$ w kierunku półosi dodatniej y , gdy $\alpha = 0, \beta = 1$.

Jeśli funkcja $f(x, y)$ ma w otoczeniu (x_0, y_0) ciągle pochodne cząstkowe $\partial_x f$ i $\partial_y f$, to na mocy wzoru na pochodną funkcji złożonej pochodna kierunkowa wyraża się wzorem

$$\nabla_p f = \frac{\partial f}{\partial x} \alpha + \frac{\partial f}{\partial y} \beta = \vec{\nabla} f \cdot \vec{p}.$$

Ponieważ α i β są składowymi wektora $\vec{p} = (\alpha, \beta)$, to powyższy wzór po prawej stronie równości zinterpretowaliśmy jako iloczyn skalarny wektora $\vec{p}$ i wektora zbudowanego z pochodnych cząstkowych zwanego *gradientem* funkcji f ,

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (f_{,x}, f_{,y}).$$

Z tego powodu wygodnie jest wprowadzić pojęcie wektorowego operatora różniczkowego

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$$

zwanego *nabla* lub operatorem gradientu, który funkcji $f(x, y)$ przyporządkowuje gradient tej funkcji $(f_{,x}, f_{,y})$.

Gradient funkcji ma bardzo interesującą własność. Aby się o tym przekonać, postawmy pytanie *w jakim kierunku w danym punkcie funkcja rośnie najszybciej?* W tym celu przekształcimy wyrażenie dla pochodnej kierunkowej do postaci

$$\nabla_p f = f_{,x} \alpha + f_{,y} \beta = \sqrt{(f_{,x})^2 + (f_{,y})^2} (a\alpha + b\beta),$$

gdzie wprowadziliśmy oznaczenia

$$a = \frac{f_{,x}}{\sqrt{(f_{,x})^2 + (f_{,y})^2}}, \quad b = \frac{f_{,y}}{\sqrt{(f_{,x})^2 + (f_{,y})^2}}, \quad a^2 + b^2 = 1.$$

W danym punkcie pochodne cząstkowe są ustalone, zmieniamy jedynie kierunek wektora $\vec{p}$. Pochodna kierunkowa przyjmie maksymalną wartość wtedy, gdy iloczyn $a\alpha + b\beta$ osiągnie maksimum. Ponieważ iloczyn ten jest równy iloczynowi skalarnemu wektora jednostkowego $\vec{p} = (\alpha, \beta)$

i wektora jednostkowego (a, b) , to maksimum jest osiągane, gdy te wektory są równoległe. Tym samym wykazaliśmy, że funkcja rośnie najszybciej w kierunku wyznaczonym przez gradient tej funkcji.

Powyższą dyskusję przeprowadziliśmy dla funkcji dwóch zmiennych. Uogólnienie na przypadek funkcji n -zmiennych jest natychmiastowe. Na przykład, dla trzech zmiennych operator gradientu ma postać

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

W powyższej dyskusji użyliśmy zwrotu operatora wektorowego bez wykazania, że rzeczywiście ten operator ma własności “wektorowe”, tak jak zwykle wektory. Na przykład, przy zamianie bazy współrzędne wektora ulegają zmianie. Należałoby więc wykazać, że przy zmianie bazy składowe gradientu odpowiednio się transformują. W przyszłości to zostanie wykazane. Dla nas wystarczy przypomnieć sobie, że iloczyn skalarny dwóch wektorów nie zależy od bazy. Pochodna kierunkowa funkcji w ustalonym kierunku nie może zależeć od wyboru bazy, co sugeruje, że rzeczywiście gradient funkcji zachowuje się jak dobry wektor.

3.9 Różniczka zupełna. Różniczkowalność

Przypomnijmy najpierw przypadek funkcji jednej zmiennej $y = f(x)$.

Interesował nas przyrost wartości funkcji Δy wraz z przyrostem zmiennej x od wartości x_0 do $x_0 + h$ i możliwość przedstawienia tego przyrostu w postaci członu liniowego w h i reszty $R(x_0, h)$ znikającej szybciej niż h , tzn.

$$\Delta y = f(x_0 + h) - f(x_0) = Ah + R(x_0, h).$$

Tutaj A jest stałą, która może zależeć jedynie od x_0 , natomiast reszta R musi spełniać warunek

$$\lim_{h \rightarrow 0} \frac{R(x_0, h)}{h} = 0.$$

Aby to było możliwe, potrzeba i wystarcza, żeby istniała skończona pochodna funkcji $f(x)$ w punkcie x_0 i równość powyższa jest realizowana dla $A = f'(x_0)$. Dowód był bardzo prosty. Z powyższej równości mamy bowiem

$$\lim_{h \rightarrow 0} \frac{R(x_0, h)}{h} = \lim_{h \rightarrow 0} \left(\frac{f(x_0 + h) - f(x_0)}{h} - A \right).$$

Jeśli więc istnieje pochodna $f'(x_0)$ i wybierzemy $A = f'(x_0)$, to rzeczywiście reszta jest małą wyższego rzędu, tzn. $\lim_{h \rightarrow 0} \frac{R(x_0, h)}{h} = 0$.

Część liniową przyrostu funkcji nazywamy różniczką tej funkcji i oznaczamy przez $dy = df = f'(x_0)h$. Jeśli przyrost zmiennej x oznaczmy przez δx , to mamy tradycyjny zapis dla różniczki

$$df = f' dx = \frac{df}{dx} dx.$$

Przyrost funkcji jest więc sumą różniczki i reszty wyższego rzędu, $\Delta f = df + R$.

Geometrycznie pochodna wyznacza tangens kąta nachylenia stycznej do krzywej w punkcie x_0 i równanie tej stycznej ma postać

$$y = y_0 + f'(x_0)(x - x_0), \quad y_0 = f(x_0).$$

Przechodzimy teraz do dyskusji funkcji wielu zmiennych. Jak zwykle rozpatrzmy przypadek funkcji dwóch zmiennych $z = f(x, y)$; rozszerzenie na przypadek n zmiennych jest natychmiastowe.

Ponownie stawiamy pytanie, kiedy przyrost wartości funkcji można przedstawić jako sumę wyrazu liniowego w przyrostach zmiennych i reszty będącej małą wyższego rzędu. Tym razem trzeba będzie nałożyć na funkcję f mocniejsze wymagania.

Wybieramy punkt $P_0 = (x_0, y_0)$ i pytamy o przyrost wartości funkcji przy przejściu do punktu $P_0 + h = (x_0 + h_x, y_0 + h_y)$, gdzie $h = (h_x, h_y)$. Ponieważ teraz mamy do czynienia z przyrostami dwóch zmiennych (w ogólności n zmiennych), to wprowadzimy miarę tego przyrostu za pomocą normy zdefiniowanej po "pitagorejsku", tzn $\|h\| = \sqrt{h_x^2 + h_y^2}$.

Definicja:

Jeśli istnieje wektor $A = (A_x, A_y)$ o tej własności, że dla dowolnego wektora $P_0 + h$ leżącego w pewnym otoczeniu P_0 zachodzi

$$\Delta z = f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0) = A_x h_x + A_y h_y + R(P, h),$$

gdzie A_x i A_y to stałe, natomiast

$$\lim_{\|h\| \rightarrow 0} \frac{R(P_0, h)}{\|h\|} = 0,$$

to funkcja $f(P)$ jest różniczkowalna w P_0 , a wektor A nazywamy wektorem pochodnej lub krótko pochodną funkcji w punkcie P_0 .

Jeżeli f jest różniczkowalna w pewnym otoczeniu $\mathcal{O}$, to odwzorowanie $\mathcal{O} \ni P \rightarrow (A_x, A_y) \in \mathcal{R}^2$ nazywamy pochodną. Jeżeli dodatkowo to odwzorowanie jest ciągłe, to znaczy $A_x(P)$, $A_y(P)$ są ciągłymi funkcjami argumentu P , to mówimy, że $f(x, y)$ jest różniczkowalna w sposób ciągły na $\mathcal{O}$. Mówimy wówczas, że funkcja jest różniczkowalna w sposób mocny (w sensie Frecheta).

Twierdzenie:

Funkcja ciągła i mająca pochodne cząstkowe $f_{,x}$ i $f_{,y}$ w pewnym otoczeniu $\mathcal{O}$ punktu P_0 i ciągłe w punkcie P_0 jest różniczkowalna w P_0 i pochodna w tym punkcie wynosi $A = (f_{,x}, f_{,y})$.

Dowód:

Konieczność: Z warunku różniczkowalności mamy

$$f(P_0 + h) - f(P_0) = A_x h_x + A_y h_y + R(P_0, h),$$

gdzie $R(P, h)$ jest małą wyższego rzędu. Obliczamy:

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \lim_{h \rightarrow 0} \frac{A_x h + R(P_0, h)}{h} = A_x$$

i podobnie dla $\frac{\partial f}{\partial y}$.

Różniczkowalność f w punkcie P_0 pociąga też ciągłość f w P_0 , gdyż

$$|f(P_0 + h) - f(P_0)| = \left| \frac{\partial f}{\partial x} h_x + \frac{\partial f}{\partial y} h_y + R \right| \leq \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \|h\| + |R(P_0, h)|$$

i rzeczywiście

$$\lim_{\|h\| \rightarrow 0} |f(P_0 + h) - f(P_0)| = 0.$$

Komentarz: jak pamiętamy, istnienie pochodnych cząstkowych nie gwarantuje ciągłości funkcji. Ciągłość pochodnych cząstkowych implikuje nie tylko ciągłość funkcji, ale też jej różniczkowalność, o czym traktuje druga część dowodu.

Dostateczność:

Rozważmy $P_0 \in \mathcal{O}$ oraz $P_0 + h \in \mathcal{O}$. Przyrost funkcji wynosi

$$\begin{aligned} \Delta f(P_0) &= f(P_0 + h) - f(P_0) = f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0) \\ &= f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0 + h_y) + f(x_0, y_0 + h_y) - f(x_0, y_0) \end{aligned}$$

Ponieważ z założenia funkcja ma ciągle pochodne na obszarze otwartym zawierającym P_0 , to różnicę $f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0 + h_y)$, traktowaną jako różnicą funkcji pierwszej zmiennej przy ustalonej wartości drugiej, można z twierdzenia Lagrange'a o wartości średniej zapisać w postaci

$$f(x_0 + h_x, y_0 + h_y) - f(x_0, y_0 + h_y) = \frac{\partial f}{\partial x}(x_0 + \theta_x h_x, y_0 + h_y) h_x, \quad (0 \leq \theta_x \leq 1).$$

Analogicznie

$$f(x_0, y_0 + h_y) - f(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0 + \theta_y h_y) h_y, \quad (0 \leq \theta_y \leq 1).$$

Jeśli więc wybierzemy

$$A_x = \frac{\partial f}{\partial x}(x_0, y_0), \quad A_y = \frac{\partial f}{\partial y}(x_0, y_0),$$

to reszta

$$\frac{R(P_0, h)}{\|h\|} = \left[\frac{\partial f}{\partial x}(x_0 + \theta_x h_x, y_0 + h_y) - A_x \right] \frac{h_x}{\|h\|} + \left[\frac{\partial f}{\partial y}(x_0, y_0 + \theta_y h_y) - A_y \right] \frac{h_y}{\|h\|}$$

dąży do zera przy $\|h\| \rightarrow 0$ na mocy ciągłości pochodnych cząstkowych.

cbdo

Przykład:

pokazujący konieczność założenia ciągłości pochodnych cząstkowych. Niech

$$f(x, y) = \frac{x^2 y}{x^2 + y^2} \quad \text{dla } (x, y) \neq (0, 0), \quad f(0, 0) = (0, 0).$$

Funkcja jest ciągła w całej płaszczyźnie, ale pochodne cząstkowe

$$f_{,x} = \frac{2xy^3}{(x^2 + y^2)^2}, \quad f_{,y} = \frac{x^2(x^2 - y^2)}{(x^2 + y^2)^2}$$

są nieciągłe w $(0,0)$, gdyż obliczając bezpośrednio z definicji mamy

$$f_{,x}(0,0) = 0, \quad f_{,y}(0,0) = 0.$$

Gdyby na ślepo skorzystać z tezy twierdzenia, to mielibyśmy

$$\Delta f(0,0) = 0 + R(h).$$

Przekonajmy się, czy rzeczywiście reszta R jest małą wyższego rzędu. Obliczając jawnie przyrost funkcji mamy

$$\Delta f(0,0) = \frac{h_x^2 h_y}{h_x^2 + h_y^2} = 0 + R(h_x, h_y),$$

Biorąc na przykład $h_x = h_y$, dostajemy

$$\lim_{h_x \rightarrow 0} \frac{R(h_x, h_x)}{h_x} = \frac{1}{2},$$

a więc reszta nie znika dostatecznie szybko.

Definicja różniczki zupełnej:

Jeśli funkcja jest różniczkowalna, to część liniową w przyrostach zmiennych (oznaczonych przez dx i dy) przyrostu funkcji $f(x, y)$

$$df = f_{,x}dx + f_{,y}dy$$

nazywamy różniczką zupełną. Możemy wtedy napisać

$$\Delta f = df + R(dx, dy).$$

Interpretacja geometryczna pochodnej

W przypadku jednej zmiennej pochodna określa tangens nachylenia stycznej do krzywej danej równaniem $y = f(x)$. Dla funkcji dwóch zmiennych wykresem funkcji jest powierzchnia dana przez $z = f(x, y)$. Domyślamy się, że teraz będzie chodziło o płaszczyznę styczną do tej powierzchni. Przypomnijmy, że prosta M_0T jest styczną do krzywej $\mathcal{K}$ w punkcie M_0 , jeśli odległość MP zmiennego punktu M na krzywej $\mathcal{K}$ od prostej M_0T jest małą wyższego rzędu niż odległość MM_0 , tzn. $\lim_{MM_0 \rightarrow 0} MP/M_0P = 0$.

Analogicznie definiujemy płaszczyznę styczną: płaszczyzna $\mathcal{P}$ jest styczna w punkcie M_0 powierzchni $\mathcal{S}$, jeśli odległość MP zmiennego punktu M powierzchni $\mathcal{S}$ od tej płaszczyzny jest małą wyższego rzędu od MM_0 , tzn. $\lim_{MM_0 \rightarrow 0} MP/MM_0 = 0$.

Równanie płaszczyzny przechodzącej przez punkt $M_0(x_0, y_0, z_0)$ jest dane równaniem

$$z - z_0 = A_x(x - x_0) + A_y(y - y_0),$$

gdzie A_x, A_y są współczynnikami kierunkowymi tej płaszczyzny. Jeśli porównamy to z wyrażeniem na przyrost różniczkowalnej funkcji $z = f(x, y)$ w punkcie x_0, y_0 ,

$$\Delta z = f_{,x}dx + f_{,y}dy + R(dx, dy),$$

to dochodzimy do następującego

Twierdzenia:

Aby powierzchnia $z = f(x, y)$ miała w punkcie $M_0(x_0, y_0, z_0)$ płaszczyznę styczną potrzeba i wystarcza, żeby w $x = x_0, y = y_0$ funkcja $f(x, y)$ była różniczkowalna. Współczynniki kierunkowe A_x, A_y tej płaszczyzny są równe pochodnym cząstkowym $f_{,x}, f_{,y}$ w tym punkcie.

Dowód pominiemy.

Przykłady:

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

a) pochodne cząstkowe

$$f_{,x} = -2x, \quad f_{,y} = -2y$$

są ciągle wszędzie, funkcja jest więc różniczkowalna.

b) różniczka zupełna w punkcie x_0, y_0 wynosi

$$df = -2x_0 dx - 2y_0 dy$$

c) sprawdzimy, że rzeczywiście reszta zachowuje się poprawnie. Obliczamy jawnie przyrost funkcji

$$\Delta f = f(x_0 + dx, y_0 + dy) - f(x_0, y_0) = -(x_0 + dx)^2 - (y_0 + dy)^2 + x_0^2 + y_0^2 = df - dx^2 - dy^2.$$

Widać, że reszta jest małą wyższego rzędu niż przyrosty zmiennych dx i dy .

d) równanie płaszczyzny stycznej w punkcie x_0, y_0

$$z + x_0^2 + y_0^2 = -2x_0(x - x_0) - 2y_0(y - y_0).$$

2. Jeśli funkcja jest różniczkowalna w punkcie x_0, y_0 , to ma w tym punkcie wszystkie pochodne kierunkowe i różniczkę zupełną. Ale może być tak, że funkcja ma w x_0, y_0 wszystkie pochodne kierunkowe, a nie mieć w tym punkcie różniczki zupełnej. Przykładem takiej funkcji jest

$$f(x, y) = \frac{xy^2}{x^2 + y^4}, \quad \text{dla } x^2 + y^2 > 0, \quad f(0, 0) = 0.$$

Obliczamy pochodne kierunkowe w kierunku zadanym przez $x = \alpha t$, $y = \beta t$

$$f(\alpha t, \beta t) = \frac{\alpha \beta^2 t}{\alpha^2 + \beta^4 t^2}, \quad \frac{df}{dt}|_{t=0} = \beta^2 / \alpha.$$

Pochodne cząstkowe w $(0,0)$ są równe zero, a przyrost funkcji w tym punkcie wynosi

$$\Delta f = \frac{dx \, dy^2}{dx^2 + dy^4}.$$

Gdyby funkcja miała różniczkę zupełną, to przyrost powinien być małą wyższego rzędu niż dx, dy , a nie jest co widać wybierając np. $dx = dy^2$, bo wtedy $\Delta f = 1/2$.

Jeżeli istnieją wszystkie (dowolne) pochodne kierunkowe funkcji f w punkcie P_0 , to mówimy, że jest ona różniczkowalna w sensie słabym (Gateaux) w tym punkcie. Nie gwarantuje to różniczkowalności w sensie mocnym.

Dla funkcji wielu zmiennych zachodzi

Twierdzenie:

Wzór Lagrange'a

Jeśli funkcja $f(x, y)$ jest różniczkowalna wewnątrz obszaru D , to dla punktów $(x + h_x, y + h_y)$ i (x, y) leżących wewnątrz D i które można połączyć odcinkiem prostym leżącym całkowicie w D istnieje takie $0 < \theta < 1$, że zachodzi wzór

$$f(x + h_x, y + h_y) - f(x, y) = f_{,x}(x + \theta h_x, y + \theta h_y)h_x + f_{,y}(x + \theta h_x, y + \theta h_y)h_y,$$

który jest analogiem znanego wzoru Lagrange'a dla funkcji jednej zmiennej.

Dowód: Definiujemy funkcję złożoną $F(t) = f(x + th_x, y + th_y)$. Funkcja ta jest ciągła i różniczkowalna dla $0 < t < 1$. Jej pochodna wynosi

$$F'(t) = f_{,x}(x + th_x, y + th_y)h_x + f_{,y}(x + th_x, y + th_y)h_y.$$

Jeśli do $F(t)$ zastosujemy wzór Lagrange'a i położymy $t = 1$, to dostajemy tezę twierdzenia. cbdo

3.10 Różniczki wyższych rzędów

Niech $P_0 = (x_0, y_0)$ i $P' = (x_0 + dx, y_0 + dy)$ oraz $T = dP = (dx, dy)$. Zdefiniujemy różniczkę zupełną rzędu k jako

$$d^k f(x, y) = \left(dx \frac{\partial}{\partial x} + dy \frac{\partial}{\partial y} \right)^k f(x, y).$$

Dla $k = 1$ mamy oczywiście $df = dx \frac{\partial}{\partial x} f + dy \frac{\partial}{\partial y} f$. Korzystając ze wzoru Newtona mamy

$$\left(dx \frac{\partial}{\partial x} + dy \frac{\partial}{\partial y} \right)^n f = \sum_{k=0}^n \binom{n}{k} \frac{\partial^n f}{\partial x^k \partial y^{n-k}} dx^k dy^{n-k}.$$

3.11 Wzór Taylora dla funkcji wielu zmiennych

Przypomnijmy, że poprzednio udowodniliśmy twierdzenie, które mówi, że jeżeli funkcja f posiada n pochodnych w otoczeniu punktu $x = a$, to

$$f(x) = f(a) + \sum_{k=1}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + R_n(x, a),$$

gdzie $R_n(x, a)$ jest małą wyższego rzędu niż $(x-a)^n$ – inaczej mówiąc,

$$\lim_{x \rightarrow a} \frac{R_n(x, a)}{(x-a)^n} = 0.$$

Zanim uogólnimy to twierdzenie na przypadek wielu zmiennych, przepisujemy je za pomocą różniczek. Korzystając z

$$df = f'(a)(x-a), \quad d^2 f = f''(a)(x-a), \quad \dots \quad d^n f = f^{(n)}(a)(x-a)^n,$$

dostajemy wzór Taylora w następującej postaci

$$f(x) = f(a) + \sum_{k=1}^n \frac{d^k f(a)}{k!} + R_n(x, a).$$

Jeśli funkcja ma $n+1$ pochodną, to resztę R_n można zapisać w postaci Lagrange'a

$$R_n(x, a) = \frac{d^{n+1}(a + \theta(x-a))}{(n+1)!}, \quad (0 < \theta < 1).$$

Możemy teraz sformułować

Twierdzenie – Wzór Taylora:

Niech $f(x, y)$ będzie klasy $C^{(n+1)}$ (tzn. $(n+1)$ -krotnie różniczkowalna w sensie mocnym) w pewnym obszarze zawierającym odcinek o końcach $A(x_0, y_0)$ i $B(x_0 + h_x, y_0 + h_y)$. Na odcinku AB znajduje się taki punkt $C(x_0 + \theta h_x, y_0 + \theta h_y)$, że

$$f(x + h_x, y + h_y) = f(x_0, y_0) + \sum_{k=1}^n \frac{d^k f}{k!} + R_n,$$

gdzie

$$d^k f = (h_x \frac{\partial}{\partial x} + h_y \frac{\partial}{\partial y})^k f, \quad R_n = \frac{1}{(n+1)!} (h_x \frac{\partial}{\partial x} + h_y \frac{\partial}{\partial y})^{(n+1)} f$$

i gdzie pochodne funkcji f w wyrażeniach $d^k f$ dla $k \leq n$ są obliczone w punkcie A , a pochodne w reszcie R_n w punkcie C .

Dowód: Rozpatrzmy funkcję złożoną $F(t) = f(x_0 + th_x, y_0 + th_y)$ zmiennej $t \in [0, 1]$. Funkcja ta ma pochodne aż do rzędu $n+1$. Stosując dla niej rozwinięcie wokół $t = 0$ i podstawiając w tym rozwinięciu $t = 1$ dostajemy tezę twierdzenia. cbdo

W powyższym twierdzeniu skorzystaliśmy z założenia, że funkcja f jest $(n+1)$ -krotnie razy różniczkowalna, co pozwoliło skorzystać z postaci Lagrange’a dla reszty R_n . Jeśli nie ma $n+1$ pochodnej, to powyższe twierdzenie też zachodzi z warunkiem dla reszty (słusznym dla R_n w dowolnej postaci – nie tylko Lagrange’a)

$$\lim_{\|h\| \rightarrow 0} \frac{R_n(A, h)}{\|h\|^n} = 0.$$

Zauważmy, że wzór Taylora można równoważnie zapisać za pomocą pochodnych kierunkowych, gdyż

$$d^k f = (h_x \frac{\partial}{\partial x} + h_y \frac{\partial}{\partial y})^k f = (\vec{h} \cdot \vec{\nabla})^k f.$$

Dla funkcji wielu zmiennych obowiązuje również twierdzenie o wartości średniej (Taylora). Rozważmy funkcję f różniczkowalną na obszarze otwartym $\mathcal{O}$ i niech punkty $A(x_0, y_0)$, $B(x_0 + h_x, y_0 + h_y) \in \mathcal{O}$. Zakładamy, że $\mathcal{O}$ zawiera cały odcinek łączący punkty AB , to znaczy $C(t) = (x_0 + th_x, y_0 + th_y) \in \mathcal{O}$, $t \in [0, 1]$. Obowiązuje wówczas

Twierdzenie o wartości średniej:

Jeżeli f jest funkcją różniczkowalną na $\mathcal{O}$ zawierającym punkty $C(t)$, to istnieje $0 \leq \theta \leq 1$ takie, że

$$f(B) - f(A) = \vec{h} \cdot \vec{\nabla} f(C(\theta)).$$

Dowód: Definiujemy funkcję złożoną jednej zmiennej $g(t) = f(C(t))$. Wówczas $g'(t) = \vec{h} \cdot \vec{\nabla} f(C(t))$. Z twierdzenia o wartości średniej dla funkcji jednej zmiennej mamy $g(1) - g(0) = g'(\theta) \cdot 1$, czyli tezę twierdzenia. cbdo

3.12 Ekstrema funkcji wielu zmiennych

W teorii funkcji jednej zmiennej pojęcie pochodnej służyło do badania przebiegu zmienności funkcji, w szczególności szukania ekstremów (minimów i maksimów) funkcji. Pokazaliśmy, że funkcja może mieć ekstremum lokalne tylko w takich punktach, w których pierwsza pochodna znika. Jest to warunek konieczny, ale nie wystarczający. Warunek dostateczny istnienia lokalnego ekstremum mówi, że pierwszą nieznikającą pochodną musi być pochodna parzystego rzędu i w zależności od znaku tej pochodnej mamy minimum, gdy ta pochodna jest dodatnia, i maksimum w przeciwnym wypadku.

Przejdziemy teraz do badania ekstremów funkcji wielu zmiennych ograniczając się, jak zwykle, do przypadku funkcji dwóch zmiennych. Najpierw

Definicja ekstremum lokalnego:

Mówimy, że funkcja $f(x, y)$ ma maksimum lokalne w punkcie $P(x_0, y_0)$, gdy istnieje taka liczba δ , że

$$f(x_0, y_0) \geq f(x_0 + h_x, y_0 + h_y) \quad \text{dla } h_x^2 + h_y^2 < \delta^2.$$

Innymi słowy, w kole o środku P i promieniu δ funkcja ma wartości nie większe niż w P . Jeśli nierówność jest w przeciwną stronę, to funkcja ma w P minimum lokalne.

Twierdzenie – warunek konieczny istnienia ekstremum lokalnego:

Funkcja różniczkowalna $f(x, y)$ w pewnym obszarze może mieć ekstremum lokalne tylko w takich punktach $P(x_0, y_0)$, w których pochodne cząstkowe znikają, to znaczy

$$f_{,x}(P) = f_{,y}(P) = 0.$$

Dowód: Rzeczywiście, niech w punkcie $P = (x_0, y_0)$ będzie np. lokalne maksimum. Wówczas funkcja jednej zmiennej $f(x, y_0)$ ma w punkcie $x = x_0$ maksimum, więc jej pochodna $f_{,x}(x_0, y_0)$ znika. Podobnie musi być $f_{,y}(x_0, y_0) = 0$. cbdo

Zauważmy, że warunek konieczny istnienia ekstremum można zapisać tak

$$df(x_0, y_0) = 0,$$

bo jeśli pochodne cząstkowe w x_0, y_0 są równe zero, to różniczka zupełna znika dla wszystkich dx i dy . I na odwrót, jeśli równość $df = 0$ jest spełniona w pewnym punkcie, to wobec dowolności dx i dy każda z pochodnych cząstkowych musi zniknąć.

Podobnie jak w przypadku funkcji jednej zmiennej punkt, w którym $\frac{\partial f}{\partial x} = 0$ i $\frac{\partial f}{\partial y} = 0$ nazywamy punktem stacjonarnym. Jeżeli funkcja jest różniczkowalna w punkcie stacjonarnym, to równanie płaszczyzny stycznej ma postać $z = (x - x_0)\frac{\partial f}{\partial x} + (y - y_0)\frac{\partial f}{\partial y} + z_0$, a więc jest ona równoległa do płaszczyzny Oxy . Ale znikanie pochodnych cząstkowych (gradientu) nie gwarantuje występowania ekstremum.

Przykład:

Niech $f(x, y) = x^2 - y^2$. Wtedy w punkcie stacjonarnym $(0, 0)$ mamy $f_{,x} = f_{,y} = 0$. W tym nie ma jednak ekstremum lokalnego. Znikanie pierwszych pochodnych jest tylko warunkiem koniecznym, ale nie dostatecznym. W tym przykładzie widać, że jeśli ustalimy $y = 0$, to funkcja $f(x, 0) = x^2$ jako funkcja jednej zmiennej x ma minimum w $x = 0$. Ustalając natomiast $x = 0$, funkcja $f(0, y) = -y^2$ ma maksimum w $y = 0$. Jak widać, że punkt $(0, 0)$ jest punktem siodłowym.

Twierdzenie – warunek wystarczający dla ekstremum:

Niech f posiada wszystkie pochodne cząstkowe do rzędu drugiego, ciągłe w obszarze otwartym $\mathcal{O}$. Jeżeli punkt $P \in \mathcal{O}$ jest punktem stacjonarnym i dla dowolnego kierunku $\vec{t}$ pochodna kierunkowa w P spełnia

$$(\vec{t} \cdot \vec{\nabla})^2 f(P) > 0,$$

to f ma minimum w punkcie P .

Jeśli znak nierówności jest przeciwny, to f ma w tym punkcie maksimum. Jeśli w zależności od kierunku $\vec{t}$ powyższe wyrażenie ma różne znaki, to punkt P jest punktem siodłowym.

Nie podajemy ścisłego dowodu, gdyż twierdzenie jest dosyć intuicyjne. Intuicję naszą opieramy na naszej znajomości funkcji jednej zmiennej i różniczek funkcji wielu zmiennych. Dla ustalenia, czy funkcja ma w punkcie $P(x_0, y_0)$ ekstremum czy nie, można rozważyć różnicę $\Delta f = f(x, y) - f(x_0, y_0)$ w dostatecznie małym otoczeniu P o promieniu ρ . Jeśli $\Delta f > 0$, to mamy minimum. Rozkładając ją według wzoru Taylora do rzędu 2, wobec stacjonarności punktu P mamy

$$\Delta f = \frac{1}{2!}(h_x^2 f_{,xx} + 2h_x h_y f_{,xy} + h_y^2 f_{,yy}),$$

gdzie pochodne cząstkowe są obliczone w pewnym punkcie $(x_0 + \theta h_x, y_0 + \theta h_y)$, a $h_x = x - x_0$ i $h_y = y - y_0$. Normalizując Δf do długości wektora $\|h\| = \rho$ i przechodząc z $\rho \rightarrow 0$, wobec ciągłości drugich pochodnych, mamy

$$\Delta f / \rho \rightarrow (\vec{t} \cdot \vec{\nabla})^2 f(P),$$

gdzie $t_x = h_x / \rho$, $t_y = h_y / \rho$.

Przykład:

Niech $f(x) = x^4 + y^4$. W punkcie stacjonarnym $(0, 0)$ znikają drugie pochodne. Nasze kryterium niczego nie rozstrzyga, choć wiemy, że jest tam minimum. W takich przypadkach należy badać pochodne wyższych rzędów.

Aby to kryterium miało znaczenie praktyczne, musimy mieć prosta metodę ustalania znaku drugiej pochodnej dla wszystkich pochodnych kierunkowych. Oznaczając $\vec{t} = (t_x, t_y)$ mamy

$$A(t) = (\vec{t} \cdot \vec{\nabla})^2 f = t_x^2 f_{,xx} + 2t_x t_y f_{,xy} + t_y^2 f_{,yy},$$

gdzie pochodne są obliczone w punkcie stacjonarnym $P(x_0, y_0)$. $A(t)$ jest więc formą kwadratową wektora $t = (t_x, t_y)$ o współczynnikach danych przez wartości drugich pochodnych cząstkowych w P . Interesuje nas znak tej formy. Niech forma kwadratowa $A(t)$ będzie dodatnio określona, to znaczy jest dodatnia dla dowolnego t , $A(t) > 0$. Wówczas nie mogą zniknąć jednocześnie $f_{,xx}$ i $f_{,yy}$. Gdyby zniknęły, to $A(t) = 2t_x t_y f_{,xy}$ miałoby przeciwny znak dla $t_1 = (t_0, t_0)$ i $t_2 = (t_0, -t_0)$. Bez zmniejszenia ogólności możemy wybrać $f_{,xx} \neq 0$. Wtedy

$$A(t) = f_{,xx} \left(t_x^2 + 2t_x t_y \frac{f_{,xy}}{f_{,xx}} \right) + t_y^2 f_{,yy}.$$

Dopełniamy nawias do pełnego kwadratu

$$\begin{aligned} A(t) &= f_{,xx} \left(t_x^2 + 2t_x t_y \frac{f_{,xy}}{f_{,xx}} + t_y^2 \frac{f_{,xy}^2}{f_{,xx}^2} \right) - t_y^2 \left(\frac{f_{,xy}^2}{f_{,xx}^2} - f_{,yy} \right) = f_{,xx} \left(t_x + t_y \frac{f_{,xy}}{f_{,xx}} \right)^2 + t_y^2 \left(f_{,yy} - \frac{f_{,xy}^2}{f_{,xx}^2} \right). \\ &= \frac{1}{f_{,xx}} [(t_x f_{,xx} + t_y f_{,xy})^2 + t_y^2 (f_{,xx} f_{,yy} - f_{,xy}^2)] \end{aligned}$$

Oznacza to, że aby $A(t) > 0$, musi być $f_{,xx} > 0$ oraz $f_{,xx} f_{,yy} - f_{,xy}^2 > 0$. Gdybyśmy chcieli $A(t) < 0$, to $f_{,xx} < 0$ oraz $f_{,xx} f_{,yy} - f_{,xy}^2 > 0$.

Udowodniliśmy więc

Twierdzenie:

Jeżeli

$$f_{,xx}f_{,yy} - f_{,xy}^2 > 0,$$

to w badanym punkcie stacjonarnym (x_0, y_0) funkcja $f(x, y)$ ma ekstremum:minimum lokalne przy $f_{,xx} > 0$ maksimum lokalne przy $f_{,xx} < 0$.Jeśli zaś $f_{,xx}f_{,yy} - f_{,xy}^2 < 0$, to ekstremum nie ma, bo $A(t)$ zmienia znak w zależności od t .

W przypadku $f_{,xx}f_{,yy} - f_{,xy}^2 = 0$ do rozstrzygnięcia zagadnienia istnienia ekstremum trzeba użyć wyższych pochodnych; tym zagadnieniem nie będziemy się zajmowali.

3.12.1 Macierz drugiej pochodnej

Zauważmy, że warunek $f_{,xx}f_{,yy} - f_{,xy}^2 > 0$ ma postać żądania postawionego wobec wyznacznika macierzy utworzonego z pochodnych cząstkowych drugiego rzędu. Macierz tę definiujemy jako

$$M = \begin{bmatrix} f_{,xx} & f_{,xy} \\ f_{,xy} & f_{,yy} \end{bmatrix}.$$

Minory główne tej macierzy to $f_{,xx}$ oraz $f_{,xx}f_{,yy} - f_{,xy}^2$. Zakładamy, że pochodne cząstkowe mieszane są przemienne.

Wniosek

Warunkiem dostatecznym występowania minimum lokalnego w punkcie stacjonarnym P_0 jest dodatniość formy kwadratowej drugiej pochodnej, co jest równoważne dodatniości minorów głównych macierzy drugich pochodnych.

Ponieważ forma ujemnie określona po zamianie znaków wszystkich wyrazów przechodzi w formę dodatnio określoną, to warunek na maksimum lokalne ma postać $f_{,xx} < 0$ oraz $f_{,xx}f_{,yy} - f_{,xy}^2 > 0$.

W tej postaci warunek dostateczny łatwo przenosi się na przypadek funkcji wielu zmiennych $f(x_1, x_2, \dots, x_n)$. Funkcja ta ma minimum lokalne w punkcie stacjonarnym (w którym wszystkie pochodne cząstkowe znikają), jeśli wszystkie minory główne macierzy drugich pochodnych

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix},$$

gdzie $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$, są dodatnie. To znaczy

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \dots, \quad \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} > 0.$$

Jeśli w powyższym ciągu zamienimy znaki $>$ na $<$ w nierówności pierwszej, trzeciej, piątej itd. to dostaniemy warunek na istnienie maksimum lokalnego.

Powyższe sformułowanie dostarcza praktycznego sposobu wyznaczania ekstremum lokalnego.

Przykład:

Niech $f(x, y) = x^2 + y^2 + xy - y - 2x$. Punkt stacjonarny wyznaczamy z $f_{,x} = 2x + y - 2 = 0$ oraz $f_{,y} = 2y + x - 1 = 0$, czyli $P = (1, 0)$. Drugie pochodne $f_{,xx} = 2$, $f_{,yy} = 2$, $f_{,xy} = 1$, zatem

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Ponieważ minory główne są dodatnie, to w punkcie P funkcja $f(x, y)$ ma minimum.

3.13 Ekstremum globalne

Jeżeli funkcja $f(x, y)$ jest ciągła na domkniętym i ograniczonym zbiorze $D \subset \mathbb{R}^2$, to osiąga ona swoje wartości ekstremalne: maksimum i minimum (twierdzenie Weierstrassa). Wartości ekstremalne mogą odpowiadać punktom stacjonarnym wewnątrz D (i są to wtedy ekstrema lokalne), albo punktom leżącym na brzegu. Aby znaleźć ekstrema globalne f na D , należy:

1. Znaleźć punkty stacjonarne $f(x, y)$ i obliczyć wartości f w tych punktach
2. Wyznaczyć maksymalną i minimalną wartość $f(x, y)$ na brzegu D
3. Z otrzymanych w punktach 1) i 2) wartości wybrać najmniejszą i największą.

Punkt 2 wymaga dodatkowego komentarza praktycznego. Jeśli brzeg D jest konturem “gładkim” (w jakimś sensie różniczkowalnym), to parametryzujemy ten kontur pewnym parametrem, szukamy punktów stacjonarnych na tym konturze i obliczamy wartości funkcji w tych punktach. Jeśli brzeg ma kilka “gładkich” części połączonych ze sobą w sposób nieróżniczkowalny (tworzące “dzióbki”), to szukamy punktów stacjonarnych na każdej części “gładkiej”, obliczamy tam wartości funkcji i jeszcze dodatkowo należy obliczyć wartości funkcji we wszystkich “dzióbkach”. Dla wyjaśnienia tej procedury i użytego nazewnictwa rozpatrzmy następujący

Przykład:

Wyznaczyć ekstrema globalne funkcji $f(x, y) = 10 + 2xy - x^2$ na obszarze $0 \leq y \leq 4 - x^2$.

Jak widać, obszar D jest zawarty pomiędzy osią x a parabolą. Brzeg składa się z paraboli i odcinka (części gładkie) połączonych dzióbkami w punktach $(x, y) = (-2, 0)$ i $(2, 0)$. Postępujemy zgodnie z powyższym przepisem:

1. Szukamy punktów stacjonarnych wewnątrz obszaru D z równań

$$f_{,x} = 2y - 2x = 0, \quad f_{,y} = 2x = 0.$$

Jedynym rozwiązaniem jest $P_0 = (0, 0)$, który leży na brzegu. Wartość funkcji w tym punkcie wynosi $f(P_0) = 10$.

2. Brzeg składa się z dwóch części gładkich: paraboli, którą możemy sparametryzować w postaci $y = g(x) = 4 - x^2$, $x \in [-2, 2]$ i odcinka $y = 0$, $x \in [-2, 2]$ oraz dwóch dzióbek w punktach $(-2, 0)$ i $(2, 0)$.

Najpierw badamy parabolę $y = g(x)$:

$$h(x) = f(x, g(x)) = 10 + 2x(4 - x^2) - x^2, \quad \Rightarrow \quad \frac{dh(x)}{dx} = -6x^2 - 2x + 8 = 0$$

Stąd dostajemy dwa punkty stacjonarne $P_1 = (-4/3, 20/9)$, $P_2 = (1, 3)$, w których wartości funkcji wynoszą $f(P_1) = 62/27$ i $f(P_2) = 15$.

Teraz badamy odcinek

$$h(x) = f(x, 0) = 10 - x^2, \quad \Rightarrow \quad \frac{dh(x)}{dx} = -2x$$

Rozwiązaniem jest punkt $P_0 = (0, 0)$.

I jeszcze wartości funkcji w dzióbkach: $f(-2, 0) = 6$, $f(2, 0) = 6$.

3. Z porównania powyższych wartości funkcji w P_0 , P_1 i P_2 dostajemy więc $f_{\max} = 15$ oraz $f_{\min} = 62/27$.

3.14 Pochodna odwzorowania przestrzeni wektorowych

Do tej pory rozważaliśmy funkcje wielu zmiennych, to znaczy odwzorowanie $\mathbb{R}^n \rightarrow \mathbb{R}^1$. Teraz rozpatrzmy odwzorowania przestrzeni wektorowej w przestrzeń wektorową $\Phi: X \rightarrow Y$, $\Phi(X) = Y$, gdzie $X \ni x = (x_1 \dots x_n)$, $Y \ni y = (y_1 \dots y_m)$. Oznacza to, że wektorowi x z przestrzeni X przyporządkowany jest wektor y z przestrzeni Y (wymiarów tych przestrzeni mogą być różne. Jeśli w tych podprzestrzeniach wybierzemy jakieś bazy, to odwzorowanie to jest w pełni określone przez podanie m funkcji n zmiennych $y_i = y_i(x_1 \dots x_n) = y_i(x)$).

Definicja:

Mówimy, że Φ jest odwzorowaniem różniczkowalnym w punkcie x , gdy funkcje y_i są funkcjami różniczkowalnymi. Oznacza to, że istnieje otoczenie punktu x , w którym przyrost wartości funkcji y_i można zapisać jako

$$y_i(x + h) - y_i(x) = \sum_{j=1}^n \frac{\partial y_i}{\partial x_j} h_j + r_i(x, h),$$

przy czym $|r_i(x, h)|/\|h\| \rightarrow 0$.

W powyższej definicji pochodne cząstkowe są obliczone w punkcie x . Możemy te pochodne ustawić w macierz

$$A = (y_{i,j}) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \dots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} = \begin{pmatrix} (\nabla y_1)^T \\ \vdots \\ (\nabla y_m)^T \end{pmatrix}.$$

Jeśli reszty ustawimy w wektor $R(x, h) = (r_1(x, h) \dots r_m(x, h))$, to warunek różniczkowalności w x odwzorowania Φ możemy zapisać w następującej postaci

$$\Phi(x + h) - \Phi(x) = A h + R(x, h),$$

a warunek na znikanie reszt

$$\frac{\|R(X, h)\|}{\|h\|} = \frac{\sqrt{\sum_i r_i^2(X, h)}}{\|h\|} \leq \frac{\sum_i |r_i|}{\|h\|} \rightarrow 0.$$

Macierz A bywa nazywana macierzą gradientów, macierzą pochodnych (lub pochodną) bądź (i to będzie określenie używane przez nas) macierzą Jacobiego odwzorowania Φ .

Przykłady:

1. Pochodna odwzorowania liniowego $\mathbb{R}^n \ni x \longrightarrow y = A \cdot x \in \mathbb{R}^m$.

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Z liniowości tego odwzorowania mamy $A \cdot (x + h) - A \cdot x = A \cdot h$. Zatem $(A \cdot x)' = A$, to znaczy A jest macierzą pochodnych, a reszta $R(x, h) = 0$.

2. Biegunowy układ współrzędnych $(x, y) \in \mathcal{R}^2 \rightarrow (\rho, \phi) \in [0, \infty[\times [0, 2\pi[\subset \mathcal{R}^2$, $x = \rho \cos \phi$, $y = \rho \sin \phi$. Macierz Jacobiego

$$\begin{pmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos \phi & -\rho \sin \phi \\ \sin \phi & \rho \cos \phi \end{pmatrix}. \quad (3.1)$$

Odwzorowanie jest lokalnie wzajemnie jednoznaczne z wyjątkiem $\rho = 0$.

3. Sferyczny układ współrzędnych, $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, $z = r \cos \theta$. Macierz Jacobiego

$$\begin{pmatrix} \sin \theta \cos \phi & r \cos \theta \cos \phi & r \sin \theta \sin \phi \\ \sin \theta \sin \phi & r \cos \theta \sin \phi & r \sin \theta \cos \phi \\ \cos \theta & -\sin \theta & 0 \end{pmatrix}. \quad (3.2)$$

Wprowadzenie macierzy Jacobiego pozwala nam zrozumieć pewne zależności pomiędzy pochodnymi cząstkowymi. Rozważmy sekwencje odwzorowań

$$\Phi : X \rightarrow Y = (y_1(x_1 \dots x_n) \dots y_m(x_1 \dots x_n)), \quad \Psi : Y \rightarrow Z = (z_1(y_1 \dots y_m) \dots z_l(y_1 \dots y_m)).$$

Definiujemy złożenie odwzorowań jako $(\Psi \circ \Phi)(x) = \Psi(\Phi(x))$.

Twierdzenie:

J eżeli odwzorowanie $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ jest różniczkowalne w punkcie x oraz odwzorowanie

$\Psi : \mathbb{R}^m \rightarrow \mathbb{R}^l$ jest różniczkowalne w punkcie $y = \Phi(x)$, to złożenie odwzorowań $(\Psi \circ \Phi) : \mathbb{R}^n \rightarrow \mathbb{R}^l$ jest różniczkowalne w punkcie x , a macierz pochodnej odwzorowania złożonego jest iloczynem macierzy pochodnych

$$(\Psi \circ \Phi)'(x) = \Psi'(\Phi(x))\Phi'(x).$$

Nie podajemy dowodu: twierdzenie jest dosyć naturalne, a dowód choć dość prosty, jest żmudny. Zamiast tego zapiszemy powyższą równość w sposób jawny

$$\Psi'(y) = \begin{pmatrix} \frac{\partial z_1}{\partial y_1} & \cdots & \frac{\partial z_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial z_l}{\partial y_1} & \cdots & \frac{\partial z_l}{\partial y_m} \end{pmatrix}, \quad \Phi'(x) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

Czyli iloczyn tych dwu macierzy daje nam pochodną superpozycji odwzorowań:

$$(\Psi'(y) \cdot \Phi'(x))_{ij} = \frac{\partial z_i}{\partial x_j} = \sum_{k=1}^m \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_j}.$$

Przykłady:

1 Pochodna funkcji wzdłuż krzywej.

Niech $f : \mathbb{R}^3 \rightarrow \mathbb{R}^1$, taka że $f : (x_1, x_2, x_3) \rightarrow f(x_1, x_2, x_3)$. Poziomicą tej funkcji nazywamy zbiór $l_c(f) = \{x : f(x) = c\}$. Jest ona pewną powierzchnią zanurzoną w $\mathbb{R}^3$. Rozważmy krzywą $g : \mathbb{R}^1 \rightarrow \mathbb{R}^3$, taką że $g(t) = (x_1(t), x_2(t), x_3(t))$. Złożenie dwóch funkcji f i g jest funkcją jednej zmiennej w pewien podzbiór $\mathbb{R}^1$. Obliczymy pochodną przekształcenia $f \circ g$, która jest pochodną funkcji f wzdłuż krzywej $g(t)$.

$$(f \circ g)'(t) = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \frac{dx_2}{dt} + \frac{\partial f}{\partial x_3} \frac{dx_3}{dt} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right) \begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{pmatrix}.$$

Pochodna $f \circ g$ po t jest iloczynem skalarnym gradientu funkcji f i wektora stycznego do krzywej g . Zauważmy, że jeżeli $g(t)$ jest zawarta w poziomicy, to $\frac{df}{dt} = 0$. Wektor gradientu jest prostopadły do poziomicy w danym punkcie (dokładniej mówiąc do stycznej do poziomicy). Równanie płaszczyzny stycznej do poziomicy w punkcie (x_1^0, x_2^0, x_3^0) :

$$(x_1 - x_1^0) \frac{\partial f}{\partial x_1} + (x_2 - x_2^0) \frac{\partial f}{\partial x_2} + (x_3 - x_3^0) \frac{\partial f}{\partial x_3} = 0.$$

2. Niech $f(x, y, z) = z + x^2 + y^2$.

Poziomica zerowa dana przez $f(x, y, z) = 0$ ma równanie $z = -(x^2 + y^2)$. Jest to równanie paraboloidy obrotowej. Wybierzmy na tej paraboloidzie kontur $g(t) = (\cos t, \sin t, -1)$. Wtedy $f(g(t)) = 0$ i $\frac{df}{dt} = 0$. Obliczmy tę pochodną jako pochodną funkcji złożonej

$$\frac{df(g(t))}{dt} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix} = (2x, 2y, 1) \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} = 0.$$

3. Równanie gazu doskonałego $pv = RT$.

Rozpatrzmy przemianę gazu daną przez $p = p_0 + a \cdot t$, $v = v_0 - b \cdot t$, $a, b > 0$. Szukamy zmiany temperatury w funkcji czasu

$$\frac{dT}{dt} = \left[\left(\frac{\partial T}{\partial p} \right)_v \frac{dp}{dt} + \left(\frac{\partial T}{\partial v} \right)_p \frac{dv}{dt} \right] = \frac{v}{R}a - \frac{p}{R}b = \frac{v_0a - p_0b}{R} - \left(\frac{ab}{R} + \frac{ab}{R} \right)t = \frac{v_0a - p_0b}{R} - 2\frac{ab}{R}t.$$

Bezpośredni rachunek daje to samo

$$\frac{d}{dt} \frac{1}{R} (p_0 + at)(v_0 - bt) = \frac{1}{R} [av_0 - bp_0 - 2abt].$$

3.15 Funkcje uwikłane

Do tej pory mówiąc o funkcjach mieliśmy na myśli funkcje zdefiniowane przez podanie przepisu na obliczenie wartości funkcji w postaci

$$y = f(x), \quad z = f(x, y), \quad \text{itd.}$$

Powyższe definicje można równoważnie zapisać jako

$$y - f(x) = 0, \quad z - f(x, y) = 0, \quad \text{itd.}$$

Są to szczególne postaci ogólniejszych relacji

$$F(x, y) = 0, \quad F(x, y, z) = 0, \quad \text{itd.}$$

Postawimy teraz pytanie, kiedy relacje typu $F(x, y) = 0$, których nie można jednoznacznie rozwikłać do postaci $y = f(x)$, definiują funkcje.

W ogólności zbiór E punktów (x, y) spełniających relacje typu $F(x, y) = 0$, tzn. $E = \{(x, y) : F(x, y) = 0\}$, może być bardzo różnorodny. Np. dla $F = x^2 + y^2 + 1$ zbiór E jest pusty, a dla $F = x^2 + y^2 - 4$ jest okręgiem. W pierwszym przypadku F nie definiuje żadnej funkcji, w drugim relacji $x^2 + y^2 - 4 = 0$ nie można rozwikłać globalnie, bo wykresem żadnej funkcji $y = f(x)$ nie może być okrąg. Można jednak tę relację rozwikłać lokalnie wokół niektórych punktów. Na przykład, dla punktów spełniających tę relację i $y > 0$ można to zrobić dostając $y = \sqrt{4 - x^2}$, natomiast jeśli $y < 0$, to $y = -\sqrt{4 - x^2}$. Tej relacji nie można jednak rozwikłać do postaci $y = f(x)$ wokół dowolnie małego otoczenia punktu $(x, y) = (2, 0)$. Ale w otoczeniu tego punktu można powyższą relację rozwikłać do postaci $x = \sqrt{4 - y^2}$.

Definicja funkcji uwikłanej:

Niech zbiór $E = \{(x, y) : F(x, y) = 0\}$ zawiera punkt (x_0, y_0) . Jeśli istnieje otoczenie $\mathcal{O}$ punktu (x_0, y_0) , w którym $y = y(x)$ (lub $x = x(y)$) dla wszystkich $(x, y) \in E$, to funkcję $y(x)$ ($x(y)$) nazywamy elementem funkcji uwikłanej w otoczeniu (x_0, y_0) .

Jasne jest, że termin “funkcja uwikłana” odnosi się tylko do sposobu przedstawienia tej funkcji, a nie do jej istoty.

Interesować nas będzie zagadnienie istnienia i jednoznaczności funkcji uwikłanej. O tym mówi następujące

Twierdzenie:

Jeśli funkcja $F(x, y)$ ma pochodne cząstkowe $F_{,x}$ i $F_{,y}$ ciągłe w otoczeniu punktu (x_0, y_0) i jeżeli

$$F(x_0, y_0) = 0, \quad i \quad F_{,y}(x_0, y_0) \neq 0,$$

to:

(a) istnieje dostatecznie małe otoczenie x_0 , w którym relacji tej odpowiada dokładnie jedno rozwiązanie $y = y(x)$,

(b) funkcja $y(x)$ ta jest tam ciągła i ma w nim ciągłą pochodną wyrażoną wzorem

$$y' = -\frac{F_{,x}(x, y)}{F_{,y}(x, y)}.$$

Dowód pominiemy. Intuicyjnie to twierdzenie jest jasne: tam, gdzie styczna do krzywej $f(x, y) = 0$ nie jest pionowa, można podać lokalnie $y = y(x)$. Jeśli $F_{,y}(x_0, y_0) = 0$, ale $F_{,x}(x_0, y_0) \neq 0$, to można to rozwickłać do postaci $x = x(y)$, gdyż tam styczna nie jest pozioma. Jeżeli natomiast jednocześnie

$$F(x_0, y_0) = 0, \quad F_{,x}(x_0, y_0) = 0 \quad i \quad F_{,y}(x_0, y_0) = 0,$$

to tej funkcji nie da się rozwickłać ani względem x , ani względem y . Taki punkt nazywamy punktem osobliwym równania $f(x, y) = 0$.

Uwaga: Zwróćmy uwagę na *lokalny* charakter tych stwierdzeń na temat funkcji uwikłanej. Odnoszą się one jedynie do elementu funkcji uwikłanej w pewnym otoczeniu rozpatrywanego punktu. Niemniej jednak, i takie stwierdzenia są bardzo ważne i praktyczne. Pozwala szybko rozstrzygnąć niełatwe zagadnienie, np. czy równanie $x^2y - x^2y^2 + (1 - x^2)\sin y = 0$ ma rozwiązanie w postaci $y = y(x)$ w otoczeniu $x = 0$? Odpowiedź jest twierdząca. Punkt $(0, 0)$ spełnia powyższą relację i jej pochodna po y w tym punkcie jest dodatnia.

Zauważmy, że jeśli istnieje $y(x)$, to wzór na jej pochodną można uzyskać różniczkując po x relację $F(x, y(x)) = 0$. Dostajemy

$$F_{,x} + y'f_{,y} = 0,$$

z czego wyprowadzamy wzór na y' . Różniczkując tę tożsamość ponownie po x dostajemy

$$F_{,xx} + y'F_{,xx} + y'F_{,xy} + y'^2F_{,yy} + y''F_{,y} = 0,$$

skąd po uwzględnieniu wzoru na y' otrzymujemy wzór na drugą pochodną

$$y'' = -\frac{F_{,xx}F_{,y}^2 - 2F_{,xy}F_{,x}F_{,y} + F_{,yy}F_{,x}^2}{F_{,y}^3}.$$

Rozszerzenie tych rozważań na funkcje uwikłane wielu zmiennych jest proste. Zapytajmy, kiedy można rozwickłać lokalnie relację $F(x, y, z) = 0$ do postaci $z = z(x, y)$?

Twierdzenie:

Jeśli funkcja $F(x, y, z)$ ma pochodne cząstkowe $F_{,x}$, $F_{,y}$, $F_{,z}$ ciągłe w otoczeniu punktu (x_0, y_0, z_0) , w którym

$$f(x_0, y_0, z_0) = 0, \quad i \quad F_{,z}(x_0, y_0, z_0) \neq 0,$$

to istnieje otoczenie (x_0, y_0) , w którym jest dokładnie jedno rozwiązanie $z = z(x, y)$ ciągłe z ciągłymi pochodnymi cząstkowymi

$$\frac{\partial z}{\partial x} = -\frac{F_{,x}}{F_{,z}}, \quad \frac{\partial z}{\partial y} = -\frac{F_{,y}}{F_{,z}}.$$

3.16 Ekstrema funkcji uwikłanej

Powyższe wzory na y' i y'' elementu $y = y(x)$ funkcji uwikłanej pozwalają na wyznaczenie ekstremów lokalnych tej funkcji nawet bez możliwości jej rozwikłania! W tym celu należy rozwiązać układ równań

$$F(x, y) = 0, \quad F_{,x} = 0, \quad \frac{F'_{,xx}}{F_{,y}} \neq 0,$$

gdyż w punkcie podejrzanym o ekstremum wyrażenie na drugą pochodną upraszcza się do $y'' = -F_{,xx}/F_{,y}$.

Przykłady:

1. Obliczyć y' i y'' funkcji danej relacją

$$\ln \sqrt{x^2 + y^2} = \operatorname{arctg} \frac{y}{x}.$$

Różniczkując obie strony względem x (i traktując $y = y(x)$), dostajemy

$$\frac{2x + 2yy'}{x^2 + y^2} = \frac{xy' - y}{x^2 + y^2}, \quad \text{czyli} \quad y' = \frac{x + y}{x - y}.$$

Różniczkując drugi raz i wstawiając y' , dostajemy

$$y'' = 2 \frac{x^2 + y^2}{(x - y)^3}.$$

Podobnie można obliczyć wyższe pochodne y po x .

2. Znaleźć ekstremum funkcji uwikłanej $y(x)$ danej równaniem

$$F(x, y) = x^3 + y^3 - 3axy = 0.$$

Obliczamy pochodne cząstkowe

$$F_{,x} = 3(x^2 - ay), \quad F_{,y} = 3(y^2 - ax).$$

Rozwiązaniem układu $F = 0$ i $F_{,x} = 0$ są dwa punkty

$$(x, y) = (0, 0) \quad i \quad (x, y) = a(\sqrt[3]{2}, \sqrt[3]{4}).$$

Ponieważ w pierwszym punkcie $F_{,y} = 0$, to ten punkt jest osobliwy. W drugim punkcie $F_{,y} = 3a^2\sqrt[3]{2} \neq 0$. Obliczamy w drugim punkcie $F_{,xx} = 6x > 0$, więc $y'' < 0$, co oznacza, że jest tam maksimum.

3. Niech z będzie określone jako funkcja x i y równaniem

$$z = x + y\phi(z).$$

Pokażemy, że zachodzi

$$\frac{\partial z}{\partial y} = \phi(z) \frac{\partial z}{\partial x},$$

przy założeniu $1 - y\phi'(z) \neq 0$.

Ze wzorów na pochodne cząstkowe mamy

$$\frac{\partial z}{\partial x} = \frac{1}{1 - y\phi'(z)}, \quad \frac{\partial z}{\partial y} = \frac{\phi(z)}{1 - y\phi'(z)},$$

z czego już wynika powyższa równość.

4. Znaleźć ekstremum funkcji uwikłanej $y(x)$ danej równaniem

$$F(x, y) = x^4 + y^4 - 4xy = 0.$$

Obliczamy pochodne cząstkowe

$$F_{,x} = 4x^3 - 4y, \quad F_{,y} = 4y^3 - 4x.$$

Rozwiązaniem układu $F = 0$ i $F_{,x} = 0$ są trzy punkty: $P_1 = (0, 0)$, $P_1 = (\sqrt[8]{3}, \sqrt[3]{27})$ i $P_2 = (-\sqrt[8]{3}, -\sqrt[3]{27})$. W punkcie P_0 znika $F_{,y}$, a więc jest to punkt osobliwy. Druga pochodna $F_{,xx} = 12x^2$ w obu punktach P_1 i P_2 jest dodatnia. Z kolei $F_{,y} = 4y^3 - 4x$ jest dodatnia w P_1 i ujemna w P_2 . Stąd wnioskujemy, że w P_1 jest lokalne maksimum, a w P_2 lokalne minimum funkcji uwikłanej.

3.17 Ekstrema związane

Stawiamy następujący problem. Chcemy znaleźć maksimum funkcji $f(x, y)$, ale tylko dla punktów spełniających pewne równanie $g(x, y) = 0$ zwane równaniem więzów.

Definicja ekstremum związanego:

Mówimy, że funkcja $f(x, y)$ ma w punkcie (x_0, y_0) maksimum związane (warunkowe) przy warunku $g(x, y) = 0$, jeśli

$$f(x_0, y_0) \geq f(x, y)$$

we wszystkich punktach (x, y) pewnego otoczenia $(x - x_0)^2 + (y - y_0)^2 < r^2$, które spełniają ten warunek. O funkcjach f i g zakładamy, że są określone i ciągłe w otoczeniu (x_0, y_0) .

Analogicznie definiujemy minimum warunkowe.

Założmy, że funkcje f i g mają pochodne cząstkowe pierwszego i drugiego rzędu ciągłe i że $g_{,y}(x_0, y_0) \neq 0$. Wtedy w okolicy (x_0, y_0) warunek $g(x, y) = 0$ można rozwikłać do postaci $y = y(x)$, więc zagadnienie szukania ekstremum związanego sprowadza się do problemu znalezienia ekstremum funkcji jednej zmiennej $F(x) = f(x, y(x))$. Warunki na istnienie ekstremum są postaci

$$F'(x) = f_{,x} + f_{,y}y' = 0, \quad F''(x) = f_{,xx} + 2f_{,xy}y' + f_{,yy}(y')^2 + f_{,y}y'' \neq 0.$$

Pochodne y' i y'' funkcji uwikłanej realizującej więzy $g(x, y) = 0$ są dane przez

$$y' = -\frac{g_{,x}}{g_{,y}}, \quad y'' = -\frac{g_{,xx}g_{,y}^2 - 2g_{,xy}g_{,x}g_{,y} + g_{,yy}g_{,x}^2}{g_{,y}^3}.$$

Uwzględniając je, dostajemy

$$F'(x) = \frac{f_{,x}g_{,y} - f_{,y}g_{,x}}{g_{,y}}$$

i podobnie dla F'' . Dostajemy więc następujące

Twierdzenie:

Funkcja $f(x, y)$ może mieć ekstremum przy warunku $g(x, y) = 0$ tylko w takich punktach, w których $g(x, y) = 0$ oraz

$$f_{,x}g_{,y} - f_{,y}g_{,x} = 0$$

i ma maksimum (minimum), gdy $F''(x) < 0$ ($F''(x) > 0$).

Zauważmy, że równanie $f_{,x}g_{,y} - f_{,y}g_{,x} = 0$ można otrzymać jako wynik usunięcia parametru λ z układu równań:

$$f_{,x} + \lambda g_{,x} = 0, \quad f_{,y} + \lambda g_{,y} = 0$$

i lewe strony tych równań są pochodnymi cząstkowymi funkcji $\Phi(x, y) = f + \lambda g$. Wobec tego powyższe twierdzenie można sformułować w następujący sposób:

Tworzymy funkcję $\Phi(x, y) = f + \lambda g$, gdzie λ jest stałą i szukamy ekstremum Φ traktując x i y jako niezależne. Z warunków koniecznych na istnienie ekstremum Φ i równania więzów $g(x, y) = 0$ wyznaczamy następnie x i y (i λ).

Metodę tę nazywa się *metodą mnożników nieoznaczonych Lagrange'a*. Łatwo jest ją uogólnić na przypadek szukania ekstremum $f(x_1, \dots, x_n)$ funkcji n zmiennych przy zadanych r ($r < n$) równaniach więzów. Tworzymy funkcję Lagrange'a Φ :

$$\Phi = f(x_1, \dots, x_n) + \sum_{i=1}^r \lambda_i g_i(x_1, \dots, x_n),$$

gdzie λ_i są dowolnymi czynnikami stałymi i szukamy jej ekstremów, tak jakby $x_1, \dots, x_n$ były niezależne. Z warunków na ekstrema Φ i z równań więzów $g_i = 0$ znajdujemy szukane $x_1^0, \dots, x_n^0$ (oraz λ_i^0).

Powyższe twierdzenie określa warunki konieczne na istnienie ekstremum. Warunek dostateczny możemy uzyskać badając, jak dla zwykłego ekstremum, znak przyrostu funkcji

$$\Delta = f(x_1, \dots, x_n) - f(x_1^0, \dots, x_n^0),$$

z tym, że argumenty muszą spełniać równania więzów. Stąd warunek ten można zastąpić przez przyrost funkcji Φ ,

$$\Delta = \Phi(x_1, \dots, x_n) - \Phi(x_1^0, \dots, x_n^0),$$

w której czynniki $\lambda_i = \lambda_i^0$. Z wcześniejszych rozważań dla funkcji wielu zmiennych wiemy już, że warunkiem dostatecznym na minimum (maksimum) jest dodatnia (ujemna) określoność formy kwadratowej różniczek $dx_1, \dots, dx_n$, której współczynniki są dane przez pochodne cząstkowe $a_{ij} = \Phi_{,x_i x_j}$ obliczone w punkcie $x_1^0, \dots, x_n^0$, przy czym różniczki muszą spełniać warunek wynikający z równania więzów $dg = g_x dx + g_y dy = 0$.

Przykłady:

1. Znaleźć ekstrema funkcji $f(x, y) = xy$, na okręgu $x^2 + y^2 = 1$.

Funkcja Lagrange'a $F(x, y) = xy + \lambda(x^2 + y^2 - 1)$. Warunki konieczne na istnienie ekstremum Φ + równanie więzów mają postać:

$$\frac{\partial F}{\partial x} = 0 = y + 2\lambda x, \quad \frac{\partial F}{\partial y} = 0 = x + 2\lambda y, \quad x^2 + y^2 = 1.$$

Rozwiązując dostajemy dwa rozwiązania:

$$\begin{aligned} \lambda = -\frac{1}{2}, \quad x = y = \pm \frac{1}{\sqrt{2}}, \\ \lambda = \frac{1}{2}, \quad x = -y = \pm \frac{1}{\sqrt{2}} \end{aligned}$$

W punktach $(-1/\sqrt{2}, -1/\sqrt{2})$ i $(1/\sqrt{2}, 1/\sqrt{2})$, $f(x, y) = 1/2$.

W punktach $(-1/\sqrt{2}, 1/\sqrt{2})$ i $(1/\sqrt{2}, -1/\sqrt{2})$, $f(x, y) = -1/2$.

Mamy tu zatem maksima i minima, bo na kole (zbiór domknięty i ograniczony) funkcja osiąga swoje kresy górny i dolny.

2. Znaleźć ekstremum funkcji $f(x, y, z, t) = x + y + z$, $x, y, z \geq 0$, przy warunku $g(x, y, z) = xyz - a^3 = 0$.

Wprowadzamy funkcję Lagrange'a $\Phi = f + \lambda g$. Z warunków na istnienie ekstremum Φ :

$$\Phi_{,x} = 1 + \lambda yz = 0, \dots \quad \Phi_{,z} = 1 + \lambda xy = 0$$

mamy

$$yz = xz = xy \quad \Rightarrow \quad x = y = z = a, \quad \lambda = -\frac{1}{a^2}.$$

Żeby sprawdzić, czy jest w tym punkcie ekstremum, obliczamy drugą różniczkę

$$\Phi = x + y + z - \frac{xyz}{a^2}$$

w tym punkcie:

$$d^2\Phi = -\frac{2}{a}(dxdy + dxdz + dydz).$$

Z równania więzów różniczki zmiennych spełniają warunek

$$dx + dy + dz = 0.$$

Wyznaczając stąd np. dz i wstawiając do $d^2\Phi$, dostajemy

$$-\frac{2}{a}(dxdy - (dx + dy)^2) = \frac{1}{a}((dx)^2 + (dy)^2 + dxdy).$$

Ponieważ ta forma jest dodatnio określona, to w znalezionym punkcie jest minimum warunkowe.

3.18 Zastosowania

W tej części będziemy rozpatrywać przestrzeń wektorową $\mathbb{R}^3$

Jeśli każdemu wektorowi $(x, y, z) \in D \subset \mathbb{R}^3$ przyporządkowana jest pewna liczba $\phi = \phi(x, y, z)$, to mówimy, że w D jest określone pole skalarne, a funkcję ϕ skalarą. Jeśli każdemu wektorowi (x, y, z) jest przyporządkowany wektor $\vec{w} = (w_x, w_y, w_z)$, gdzie $w_x = w_x(x, y, z)$, $w_y = w_y(x, y, z)$ i $w_z = w_z(x, y, z)$, to mówimy, że jest określone pole wektorowe. Wielkości fizyczne, jak siła, pęd, są wektorami, natomiast masa lub temperatura, są skalarami.

3.18.1 Gradient i potencjał

Jak wiemy, gdy w obszarze D jest pole skalarne, które jest funkcją różniczkowalną, to możemy określić pole wektorowe będące gradientem tego pola

$$\vec{w} = (\text{grad})\phi = (\phi_x, \phi_y, \phi_z).$$

Jeśli pole wektorowe $\vec{w}$ wyrazić można jako gradient pewnego pola skalarne ϕ , to mówimy, że pole wektorowe $\vec{w}$ jest potencjalne, a pole skalarne ϕ nazywamy potencjałem pola wektorowego.

Przykład:

Potencjał grawitacyjny punktu materialnego o masie m umieszczonego w punkcie (x_0, y_0, z_0) w punkcie o współrzędnych (x, y, z) jest dany wzorem

$$\phi = k \frac{m}{r},$$

gdzie $r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$. Potencjał ma symetrię sferyczną względem punktu materialnego, gdyż powierzchnie stałej wartości potencjału (powierzchnie ekwipotencjalne) są sferami o środku w tym punkcie materialnym.

Natężenie pola grawitacyjnego jest polem wektorowym równym gradientowi potencjału

$$\vec{g} = -\text{grad}\phi.$$

Obliczmy składową natężenia pola w kierunku osi x .

$$g_x = -\frac{\partial \phi}{\partial x} = k \frac{m}{r^2} \frac{x - x_0}{r} = k \frac{m(x - x_0)}{r^3}.$$

Podobnie obliczamy pozostałe składowe. Widzimy, że pole grawitacyjne ma postać

$$\vec{g} = km \frac{\vec{r}}{r^3}, \quad \vec{r} = (x - x_0, y - y_0, z - z_0),$$

to znaczy jest centralne, skierowane wzdłuż wektora od masy m do punktu obserwacji i jego wartość maleje z kwadratem odległości.

Nie każde pole wektorowe ma potencjał. Warunkiem koniecznym na istnienie potencjału pola wektorowego $\vec{w}$ jest

$$w_{x,y} - w_{y,x} = 0, \quad w_{y,z} - w_{z,y} = 0, \quad w_{x,z} - w_{z,x} = 0,$$

co wynika z wyrażenia na gradient i równości drugich pochodnych mieszanych potencjału.

3.18.2 Dywergencja i rotacja

Mając określone pole wektorowe $\vec{w}$, którego składowe są określone przez funkcje różniczkowalne możemy zdefiniować dywergencję (rozbieżność) tego pola przez

$$\operatorname{div} \vec{w} = \frac{\partial w_x}{\partial x} + \frac{\partial w_y}{\partial y} + \frac{\partial w_z}{\partial z} \equiv w_{x,x} + w_{y,y} + w_{z,z}.$$

Korzystając z definicji operatora gradientu (nabla),

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right),$$

możemy dywergencję pola $\vec{w}$ uważać za wynik mnożenia skalarnego operatora gradientu przez pole $\vec{w}$, tzn.

$$\operatorname{div} \vec{w} = \vec{\nabla} \cdot \vec{w}.$$

Bardzo wygodna w użyciu jest notacja wskaźnikowa, którą już częściowo poznaliśmy. Zamiast składowych x, y, z możemy użyć wskaźników 1, 2, 3 do numerowania współrzędnych. i -ita składowa gradientu pola skalarnego ϕ wynosi

$$(\operatorname{grad} \phi)_i = \phi_{,i}$$

a dywergencja pola wektorowego $\vec{w}$ jest równa

$$\operatorname{div} \vec{w} = w_{1,1} + w_{2,2} + w_{3,3} = \sum_{i=1}^3 w_{i,i} \equiv w_{i,i},$$

gdzie w ostatniej równości przyjmujemy konwencję, że po powtarzającym się wskaźniku należy rozumieć sumowanie po całym zakresie jego zmienności. Ta konwencja pozwala uprościć notację, co znacznie ułatwia zrozumienie wyrażeń. Ale też wymaga pewnej ostrożności w nazywaniu wskaźników, po których ma być sumowanie.

Ponieważ, gradient pola skalarnego jest polem wektorowym, to możemy obliczyć teraz dywergencję gradientu pola skalarnego

$$\operatorname{div} \operatorname{grad} \phi = \vec{\nabla} \cdot (\vec{\nabla} \phi) = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \equiv \Delta \phi$$

gdzie przez Δ oznaczyliśmy bardzo ważny operator różniczkowy drugiego rzędu

$$\Delta = \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

zwany laplasjanem. Operator ten można stosować do funkcji dwukrotnie różniczkowalnych. Pojawia się on w wielu równaniach fizyki. Powyższy wynik mówi, że rozbieżność (dywergencja) pola potencjalnego jest równa laplasjanowi potencjału. W notacji wskaźnikowej laplasjan pola skalarnego to

$$\Delta \phi = \phi_{,ii}.$$

Ponieważ operator nabra traktujemy jak “wektor”, to możemy “obliczyć” iloczyn wektorowy tego operatora z pewnym polem wektorowym. Wynik tego działania nazywamy rotacją (krążeniem) pola wektorowego, której jawna postać jest następująca

$$\operatorname{rot} \vec{w} = \vec{\nabla} \times \vec{w} = \left(\frac{\partial w_z}{\partial y} - \frac{\partial w_y}{\partial z}, \frac{\partial w_x}{\partial z} - \frac{\partial w_z}{\partial x}, \frac{\partial w_y}{\partial x} - \frac{\partial w_x}{\partial y} \right).$$

Wynik ten łatwo zapamiętać korzystając z mnemotechnicznego sposobu obliczania wyniku iloczynu wektorowego dwóch wektorów. i -ta składowa tego iloczynu to

$$(\vec{A} \times \vec{B})_i = \epsilon_{ijk} A_j B_k$$

gdzie ϵ_{ijk} jest obiektem (tensorem) całkowicie antysymetrycznym zdefiniowanym w następujący sposób: jeśli wskaźniki ijk są parzystą permutacją 123, to wartość jest $+1$, jeśli jest to nieparzysta permutacja, to wartość jest -1 , a gdy jakiś wskaźnik się powtarza, to 0. Korzystając z tej formalnej analogii operatora różniczkowego nabra $\vec{\nabla}$ ze zwykłym wektorem, możemy napisać dla i -tej składowej rotacji

$$(\operatorname{rot} \vec{w})_i = \epsilon_{ijk} \nabla_j w_k = \epsilon_{ijk} w_{k,j}$$

W tym wyrażeniu jest suma po wskaźnikach j, k , ale na skutek własności ϵ_{ijk} z tej sumy zostają niezerowe tylko dwa wyrazy, gdy j, k są różne od siebie i różne od i .

Zauważmy, że z formalnego zapisu operatora nabra jako wektora, łatwo widać, że

$$\operatorname{rot}(\operatorname{grad} \phi) = \vec{\nabla} \times \vec{\nabla} \phi = 0, \quad \operatorname{div}(\operatorname{rot} \vec{w}) = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{w}) = 0,$$

co jest też łatwo sprawdzić bezpośrednim rachunkiem. Wzory te wyrażają ważne własności pól: pole potencjalne jest bezwirowe – jego rotacja wynosi zero, pole wirowe jest bezźródłowe – jego dywergencja wynosi zero.

3.18.3 Funkcje jednorodne

Funkcję $f(x, y, z, \dots)$ zmiennych $x, y, z, \dots$ nazywamy jednorodną stopnia a , jeśli przy przeskalowaniu wszystkich zmiennych o czynnik λ zachodzi

$$f(\lambda x, \lambda y, \lambda z, \dots) = \lambda^a f(x, y, z, \dots).$$

Dla funkcji jednorodnych zachodzi wzór Eulera

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} + z \frac{\partial f}{\partial z} + \dots = a f(x, y, z, \dots),$$

co łatwo widać różniczkując powyższą definicję po λ i wstawiając $\lambda = 1$.

Jeśli zachodzi wzór Eulera, to funkcja musi być jednorodna. Oznaczając $g(\lambda) = f(\lambda x, \lambda y, \lambda z, \dots) - \lambda^a f(x, y, z, \dots)$, dostajemy równanie

$$g'(\lambda) = \frac{a}{\lambda} g(\lambda).$$

Weźmy $\gamma \lambda = \lambda^{-a} g(\lambda)$. Różniczkując dostajemy $\gamma'(\lambda) = 0$, co oznacza, że γ jest funkcją stałą. Ponieważ $\gamma(1) = g(1) = 0$, to $\gamma(\lambda) = 0$, a zatem pokazaliśmy, że f jest funkcją jednorodną.