

Wykład 4.

Matematyka 2, semestr letni 2010/2011

Po nauczaniu się solidnej porcji algebry liniowej możemy w końcu przejść do analizy funkcji wielu zmiennych rzeczywistych. Będziemy pracować także z odwzorowaniami z \mathbb{R}^n do \mathbb{R}^m , tzn odwzorowaniami „składającymi się” z m funkcji n zmiennych:

$$F : \mathbb{R}^n \ni (x^1, x^2, \dots, x^n) \longmapsto (F^1(x^1, \dots, x^n), F^2(x^1, \dots, x^n), \dots, F^m(x^1, \dots, x^n)) \in \mathbb{R}^m$$

Zanim zajmniemy się różniczkowaniem musimy omówić problemy związane z ciągłością i pojęciem granicy odwzorowania w punkcie. Przypomnijmy sobie definicję ciągłości funkcji jednej zmiennej w punkcie. Niech I będzie odcinkiem otwartym w \mathbb{R} (bez końców). Mówimy, że funkcja $f : I \rightarrow \mathbb{R}$ jest ciągła w punkcie $x_0 \in I$ jeśli wartości funkcji w punktach bliskich x_0 są bliskie $f(x_0)$. Precyzyjnie warunek ciągłości zapisujemy

$$(1) \quad \forall \varepsilon > 0 \quad \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Powyższą definicję (zwaną *definicją ciągłości Cauchy’ego*) można uogólnić na przypadek odwzorowań z \mathbb{R}^n do \mathbb{R}^m posługując się pojęciem odległości. Załóżmy, że potrafimy mierzyć odległość między punktami w \mathbb{R}^n i w \mathbb{R}^m . Niech $x = (x^1, x^2, \dots, x^n)$, $y = (y^1, y^2, \dots, y^n)$ będą elementami \mathbb{R}^n , wtedy odległość między nimi oznaczmy $d(x, y)$. W przestrzeni wartości, czyli w \mathbb{R}^m użyjemy symbolu ρ , tzn jeśli $a = (a^1, a^2, \dots, a^m)$, $b = (b^1, b^2, \dots, b^m)$ to odległość między punktami a i b oznaczmy $\rho(a, b)$. Teraz możemy napisać warunek ciągłości w punkcie $x_0 \in \mathbb{R}^n$: Mówimy, że odwzorowanie

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

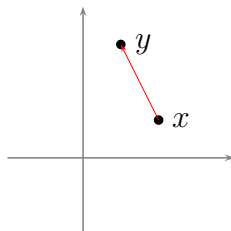
jest ciągłe w punkcie x_0 , jeśli spełniony jest warunek

$$(2) \quad \forall \varepsilon > 0 \quad \exists \delta > 0 : d(x, x_0) < \delta \Rightarrow \rho(F(x), F(x_0)) < \varepsilon.$$

Warunek (2) jest bardzo podobny do (1) - wartość bezwzględną różnicy liczb należy zamienić na odległość między punktami odpowiednio w \mathbb{R}^n i \mathbb{R}^m . Wygląda więc na to, że ciągłość odwzorowania zależy od sposobu mierzenia odległości w dziedzinie i zbiorze wartości. W ogólności rzeczywiście tak jest. Oto przykład:

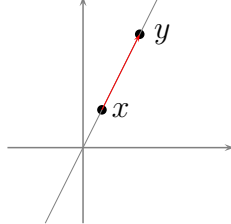
Przykład 1. Rozważmy dwa sposoby mierzenia odległości między punktami na płaszczyźnie (\mathbb{R}^2). Pierwszy sposób - euklidesowy:

$$x = (x^1, x^2), \quad y = (y^1, y^2), \quad d_2(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2},$$



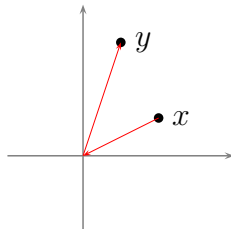
a drugi kolejowy (stacja węzłowa i tory): jeśli punkty x i y leżą na jednej prostej przechodzącej przez punkt $(0, 0)$ wtedy odległość mierzymy jak zwykle:

$$d_k(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2},$$



w przeciwnym przypadku jedziemy najpierw do stacji węzłowej:

$$d_k(x, y) = \sqrt{(x^1)^2 + (x^2)^2} + \sqrt{(y^1)^2 + (y^2)^2}.$$



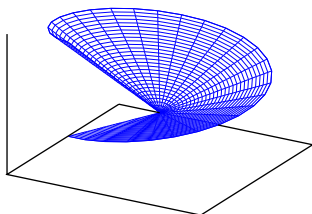
Rozważmy teraz funkcję, która przypisuje punktowi $x = (x^1, x^2)$ wartość

$$F(x^1, x^2) = r(x)\varphi(x),$$

gdzie $r(x)$ jest odległością punktu x od $(0, 0)$ w sensie euklidesowym a $\varphi(x) \in [0, 2\pi[$ jest kątem jak we współrzędnych biegunowych. Rozważmy punkt $x_0 = (1, 0)$ Funkcja ta jest nieciągła w x_0 względem metryki euklidesowej. Istotnie, Odległość euklidesowa od punktu x_0 do punktu $x = (1, -a)$ dla $a > 0$ jest równa a . Różnica w wartości zaś jest bliska 2π i dla a dążącego do 0 rośnie do 2π a nie zmniejsza się. W metryce kolejowej punkt $(1, -a)$ nie jest dobrym punktem do konstruowania kontrprzykładu na ciągłość, gdyż odległość kolejowa

$$d_k(x_0, x) = 1 + \sqrt{1 + a^2} > 2.$$

Naprawdę blisko punktu x_0 leżą jedynie punkty położone na osi $0X$. Dla tych punktów wartość funkcji jest stała i równa zero. Funkcja F jest więc ciągła względem odległości kolejowej. Oto wykres funkcji F :



Korzystając z pojęcia odległości możemy również mówić o ciągach zbieżnych w \mathbb{R}^n . Niech $(x_k)_{k=1}^\infty$ będzie ciągiem punktów w \mathbb{R}^n (tzn. każde x_k ma n współrzędnych: $(x_k^1, x_k^2, \dots, x_k^n)$). Mówimy, że *ciąg* (x_k) *jest zbieżny do* $x_0 = (x_0^1, x_0^2, \dots, x_0^n)$ jeśli spełniony jest warunek

$$(3) \quad \forall \varepsilon > 0 \quad \exists N \in \mathbb{N} : \quad \forall k > N d(x_k, x_0) \leq \varepsilon.$$

Pojęcie ciągłości w \mathbb{R}^n ma też swoją ciągową wersję, podobną do definicji Heinego:

Fakt 1. *Odwzorowanie $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ jest ciągłe w $x_0 \in \mathbb{R}^n$ wtedy i tylko wtedy, gdy dla każdego ciągu (x_k) zbieżnego do x_0 , ciąg $F(x_k)$ jest zbieżny do $F(x_0)$.*

Nie będziemy dowodzić tego faktu, gdyż dowód nie różni się od dowodu dla jednego wymiaru. Ciągowa wersja definicji ciągłości nadaje się dobrze do wykazywania, że jakaś funkcja jest nieciągła. Wystarczy wtedy podać jeden ciąg (x_k) stanowiący kontrprzykład, tzn taki, który jest zbieżny do x_0 , ale ciąg wartości $F(x_k)$ nie jest zbieżny do $F(x_0)$. Definicja Cauchy'ego nadaje się lepiej do dowodzenia ciągłości.

W definicji ciągłości Cauchy'ego badamy wartości funkcji F na punktach \mathbb{R}^n znajdujących się w odległości mniejszej niż δ od x_0 . Zbiór tych punktów nazywać będziemy kulą otwartą. Dokładniej, *kulą otwartą* w \mathbb{R}^n o środku w x_0 i promieniu δ względem odległości d nazywamy zbiór

$$K_d(x_0, \delta) = \{ x \in \mathbb{R}^n : d(x, x_0) < \delta \}.$$

Warunek ciągłości można przeformułować używając kul:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : x \in K_d(x_0, \delta) \Rightarrow F(x) \in K_\rho(F(x_0), \varepsilon).$$

W pojęciu ciągłości ważny jest więc kształt kul względem danej odległości. Zdarza się, że różne sposoby mierzenia odległości prowadzą do takiego samego zbioru funkcji ciągłych. Zanim omówimy to szczegółowo sformułujmy precyzyjnie pojęcie *sposobu mierzenia odległości*:

Sposób mierzenia odległości wyrażamy matematycznie za pomocą funkcji zwanej metryką. Mówimy, że $d : X \times X \rightarrow \mathbb{R}$ jest *metryką* na zbiorze X gdy spełnione są warunki

- (1) $\forall x, y \in X \quad d(x, y) \geq 0$
- (2) $\forall x, y \in X \quad d(x, y) = d(y, x)$
- (3) $d(x, y) = 0 \Leftrightarrow x = y$
- (4) $\forall x, y, z \in X \quad d(x, z) \leq d(x, y) + d(y, z)$

Warunek (2) oznacza, że metryka jest symetryczna, warunek (3) wyraża niezdegenerowanie, warunek (4) nosi nazwę nierówności trójkąta. Oprócz całego mnóstwa dziwnych metryk na \mathbb{R}^n istnieją trzy bardzo zwyczajne, których będziemy używać: Pierwsza metryka to *metryka euklidesowa*, zapewne najbardziej dla państwa naturalna, ale i najtrudniejsza w użyciu ze względu na nieporęczny pierwiastek:

$$d_2(x, y) = \sqrt{(x^1 - y^1)^2 + \dots + (x^n - y^n)^2}.$$

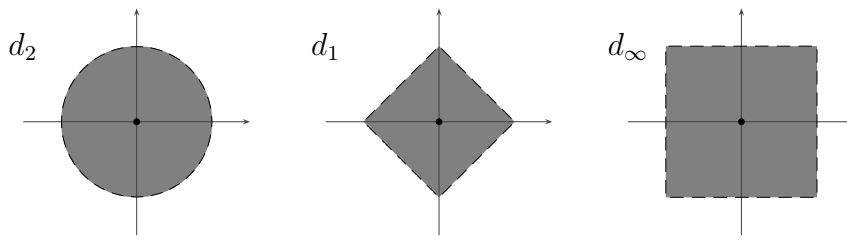
Druga metryka nazywana jest *metryką miejską*, gdyż odległość mierzona jest wzdłuż prostokątnych ulic:

$$d_1(x, y) = |x^1 - y^1| + \dots + |x^n - y^n|.$$

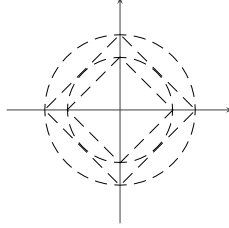
Ostatnia metryka, to tak zwana *metryka maksimum*:

$$d_\infty(x, y) = \max\{|x^1 - y^1|, \dots, |x^n - y^n|\}.$$

Wspomnieliśmy wcześniej, że dla pojęcia ciągłości istotny jest kształt kuli względem metryki. Żeby wyrobić sobie jakieś wyobrażenie o trzech używanych przez nas metrykach narysujmy kule o środku w $(0, 0)$ i promieniu 1 w \mathbb{R}^2 względem każdej z nich:

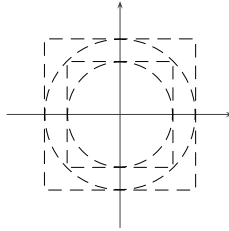


Zauważmy także, że dobierając odpowiednio promień możemy zawrzeć jedną kulę w drugiej: np dla d_2 i d_1 :



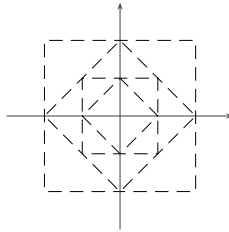
$$K_1((0,0), \frac{\sqrt{2}}{2}) \subset K_2((0,0), \frac{\sqrt{2}}{2}) \subset K_1((0,0), 1) \subset K_2((0,0), 1)$$

Podobnie dla pary d_2, d_∞ :



$$K_2((0,0), \frac{\sqrt{2}}{2}) \subset K_\infty((0,0), \frac{\sqrt{2}}{2}) \subset K_2((0,0), 1) \subset K_\infty((0,0), 1)$$

To samo dla pary d_1, d_∞ :



$$K_1((0,0), \frac{1}{2}) \subset K_\infty((0,0), \frac{1}{2}) \subset K_1((0,0), 1) \subset K_\infty((0,0), 1)$$

Jeśli więc funkcja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ jest ciągła względem d_2 , tzn dla każdego ε znajdziemy δ taką, że warunek ciągłości (2) zachodzi dla metryki d_2 , to będzie zachodził także dla $\delta' = \frac{\sqrt{2}}{2}\delta$ dla metryki d_∞ (bo kula względem d_∞ i promieniu δ' zawiera się w kuli o promieniu δ względem metryki d_2). Warunek ten będzie zachodził także dla d_1 jeśli weźmiemy promień $\delta'' = \delta$. Podobnie, jeśli zapiszemy warunek dla metryki d_1 , to możemy dobrać takie promienie względem d_2 i d_∞ , żeby odpowiednie kule się zawierały w kuli względem d_1 i warunek zachodził. To samo, jeżeli zaczniemy od d_∞ . Widać więc, że metrykom d_1, d_2, d_∞ odpowiadają takie same zbiory funkcji

ciągłych. Dopuszczenie wielu wymiarów w przestrzeni wartości nie zmienia sytuacji: także te trzy metryki mają te same zbiory odwzorowań ciągłych. Będziemy więc we wszelkich dowodach ciągłości i zbieżności używać tych trzech metryk wymiennie - w zależności od wygody rachunkowej. Podobnie jest ze zbieżnością ciągów w \mathbb{R}^n . Wszystkie trzy metryki definiują takie same zbiory ciągów zbieżnych. Niech (x_k) będzie ciągiem zbieżnym do x_0 względem którejkolwiek z trzech metryk. Wtedy jest zbieżny także względem metryki maksimum. Oznacza to, że

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n > N \quad \max\{|x_k^1 - x_0^1|, \dots, |x_k^n - x_0^n|\} < \varepsilon$$

Skoro największa z różnic jest mniejsza niż ε to także każda z nich jest mniejsza. To prowadzi do wniosku, że każda ze współrzędnych ciągu (x_k) zbiega oddzielnie do odpowiedniej współrzędnej punktu granicznego x_0 . Odwrotnie, jeśli każda ze współrzędnych ciągu (x_k) zbiega do odpowiedniej współrzędnej punktu (x_0) , to znaczy, że dla każdego indeksu i możemy znaleźć N_i takie, że dla $n > N_i$ zachodzi nierówność $|x_k^i - x_0^i| < \varepsilon$. Największe z N_i zapewnia spełnienie warunku zbieżności ciągu x_k do x_0 . Zbieżność w \mathbb{R}^n oznacza więc zbieżność po współrzędnych.

Oto dwa przykłady niecodziennych sytuacji, które możemy napotkać badając ciągłość funkcji zależnych od wielu zmiennych:

Przykład 2. Rozważmy funkcję $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ w otoczeniu punktu $(0, 0)$

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

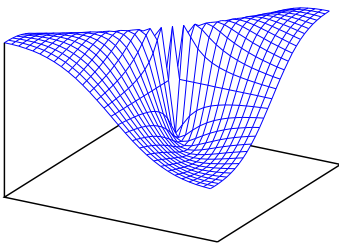
Na osiach współrzędnych, czyli na zbiorach $\{(0, y)\}$ i $\{(x, 0)\}$ funkcja f przyjmuje wartość 0. Gdy jednak zmierzamy do zera wzdłuż innej prostej, np. $y = ax$ okazuje się, że f jest nieciągła w $(0, 0)$. Niech $\alpha_k = (\frac{1}{n}, \frac{a}{n})$. Jest oczywiste, że

$$\lim_{k \rightarrow \infty} \alpha_k = (0, 0)$$

Ciąg wartości jest ciągiem stałym

$$f(\alpha_k) = f\left(\frac{1}{n}, \frac{a}{n}\right) = \frac{\frac{a}{n^2}}{\frac{1}{n^2} + \frac{a^2}{n^2}} = \frac{a}{1 + a^2}$$

z granicą $\frac{a}{1+a^2} \neq 0$.



Przykład 3. Rozważmy funkcję $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ w otoczeniu punktu $(0, 0)$

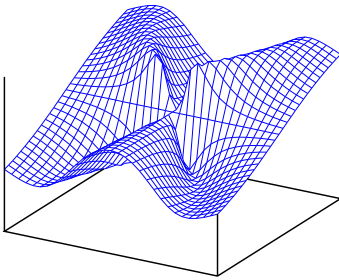
$$\begin{cases} g(x, y) = \frac{xy^2}{x^2 + y^4} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Na osiach współrzędnych, czyli na zbiorach $\{(0, y)\}$ i $\{(x, 0)\}$ funkcja g przyjmuje wartość 0. Na dowolnej prostej przechodzącej przez zero ($y = ax$) otrzymujemy

$$g(x, ax) = \frac{a^2 x^3}{x^2 + a^4 x^4} = \frac{a^2}{1 + a^4 x^2} \xrightarrow{(x \rightarrow 0)} 0$$

Gdy jednak zmierzamy do zera wzdłuż innej krzywej, np $x = ay^2$ okazuje się, że g jest nieciągła w $(0, 0)$.

$$g(ay^2, y) = \frac{ay^4}{a^2 y^4 + y^4} = \frac{a}{1 + a^2} \neq 0 \quad \text{dla } a \neq 0.$$



Różniczkowanie funkcji i odwzorowań zależnych od wielu zmiennych rzeczywistych.

Oznaczmy przez F pewne odwzorowanie z \mathbb{R}^n do \mathbb{R}^m . Oznacza to, że F składa się z m funkcji n zmiennych:

$$F : \mathbb{R}^n \ni (x^1, \dots, x^n) \longmapsto (F^1(x^1, \dots, x^n), \dots, F^m(x^1, \dots, x^n)) \in \mathbb{R}^m.$$

Mówimy, że F jest *różniczkowalne w punkcie* $x_0 = (x_0^1, \dots, x_0^n)$ jeśli istnieje odwzorowanie liniowe $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ takie, że

$$F(x_0 + h) = F(x_0) + A(h) + r(x_0, h),$$

gdzie reszta $r(x_0, h)$ spełnia warunek

$$\lim_{h \rightarrow 0} \frac{\|r(x_0, h)\|}{\|h\|} = 0$$

W powyższym wzorze $\|h\|$ oznacza długość wektora h . Odwzorowanie A nazywamy pochodną F w punkcie x_0 . Oznaczamy je także $F'(x_0)$.

Powyższa definicja nie jest niestety konstruktywna, tzn. nie precyzuje postaci odwzorowania A . Spróbujmy ją odgadnąć. Ustalmy pewne odwzorowanie F i załóżmy, że jest ono różniczkowalne. $A = F'(x_0)$ jest odwzorowaniem liniowym z \mathbb{R}^n do \mathbb{R}^m , czyli macierzą mającą m -wierszy i n -kolumn. Wyrazy macierzowe znajdziemy (jak zawsze dla odwzorowań liniowych) obliczając wartość odwzorowania na wektorach bazowych. Użyć należy oczywiście bazy standardowej.

Niech $e = (e_1, \dots, e_n)$ oznacza bazę standardową w \mathbb{R}^n a $f = (f_1, \dots, f_m)$ bazę standardową w \mathbb{R}^m . Wyraz macierzowy A^i_j jest i -tą współrzędną wektora $A(e_j)$ w bazie f , tzn:

$$A(e_j) = A^1_j f_1 + \dots + A^m_j f_m.$$

Znajdźmy $A(e_j)$. Z definicji wiemy, że dla przyrostu $h = te_j$, $t \in \mathbb{R}$, mamy

$$F(x_0 + te_j) = F(x_0) + A(te_j) + r(x_0, te_j) = F(x_0) + tA(e_j) + r(x_0, te_j).$$

Wyznaczamy $A(e_j)$:

$$tA(e_j) = F(x_0 + te_j) - F(x_0) - r(x_0, te_j), \quad \text{tzn.} \quad A(e_j) = \frac{F(x_0 + te_j) - F(x_0)}{t} - \frac{r(x_0, te_j)}{t}.$$

W granicy, dla $t \rightarrow 0$ wyrażenie z resztą znika (co wynika z definicji różniczkowalności) i otrzymujemy wzór:

$$A(e_j) = \lim_{t \rightarrow 0} \frac{F(x_0 + te_j) - F(x_0)}{t}$$

bardzo przypominający definicję pochodnej dla funkcji jednej zmiennej. Pamiętać jednak należy, że w liczniku jest różnica dwóch punktów z przestrzeni \mathbb{R}^m , a więc także element \mathbb{R}^m . Jeśli skorzystamy z bazy standardowej f otrzymamy

$$A(e_j) = \left(\lim_{t \rightarrow 0} \frac{F^1(x_0 + te_j) - F^1(x_0)}{t} \right) f_1 + \dots + \left(\lim_{t \rightarrow 0} \frac{F^m(x_0 + te_j) - F^m(x_0)}{t} \right) f_m.$$

Wyraz macierzowy A^i_j ma więc postać

$$A^i_j = \lim_{t \rightarrow 0} \frac{F^i(x_0 + te_j) - F^i(x_0)}{t}.$$

Wypiszmy jeszcze explicite argumenty funkcji F^i :

$$A^i_j = \lim_{t \rightarrow 0} \frac{F^i(x_0^1, x_0^2, \dots, x_0^{j-1}, x_0^j + t, x_0^{j+1}, \dots, x_0^n) - F^i(x_0^1, x_0^2, \dots, x_0^{j-1}, x_0^j, x_0^{j+1}, \dots, x_0^n)}{t}.$$

Powyższe wyrażenie nazywamy *pochodną cząstkową funkcji F^i w punkcie x_0 po współrzędnej x^j* i oznaczamy

$$\frac{\partial F^i}{\partial x^j}(x_0).$$

Praktycznie pochodną cząstkową obliczamy różniczkując funkcję F^i względem x^j , zgodnie z zasadami obowiązującymi dla funkcji jednej zmiennej, traktując pozostałe zmienne jako ustalone. Okazało się, że jeśli F jest różniczkowalne, to jego pochodna w punkcie x_0 ma postać:

$$F'(x_0) = \begin{bmatrix} \frac{\partial F^1}{\partial x^1}(x_0) & \frac{\partial F^1}{\partial x^2}(x_0) & \dots & \frac{\partial F^1}{\partial x^n}(x_0) \\ \frac{\partial F^2}{\partial x^1}(x_0) & \frac{\partial F^2}{\partial x^2}(x_0) & \dots & \frac{\partial F^2}{\partial x^n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F^m}{\partial x^1}(x_0) & \frac{\partial F^m}{\partial x^2}(x_0) & \dots & \frac{\partial F^m}{\partial x^n}(x_0) \end{bmatrix}$$

Przykład 4. Jako przykład rozważmy odwzorowanie $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ($m = 2$, $n = 2$), dane wzorem

$$\mathbb{R}^2 \ni (r, \varphi) \longmapsto (x(r, \varphi), y(r, \varphi)) = (r \cos \varphi, r \sin \varphi) \in \mathbb{R}^2,$$

które obcięte do odpowiedniego podzbioru w dziedzinie opisuje współrzędne biegunowe na płaszczyźnie. Sprawdźmy, czy odwzorowanie to jest różniczkowalne w punkcie (r_0, φ_0) . Naturalnym kandydatem na pochodną jest macierz złożona z pochodnych cząstkowych funkcji składowych:

$$A = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix}.$$

Oznaczając przyrosty względem r i φ przez h_r i h_φ otrzymujemy

$$\Phi(r_0 + h_r, \varphi_0 + h_\varphi) = \Phi(r_0, \varphi_0) + A \begin{bmatrix} h_r \\ h_\varphi \end{bmatrix} + r((r_0, \varphi_0), (h_r, h_\varphi))$$

Badanie różniczkowalności polega na analizowaniu postaci reszty. Wyznaczamy resztę z powyższego równania

$$r((r_0, \varphi_0), (h_r, h_\varphi)) = \Phi(r_0 + h_r, \varphi_0 + h_\varphi) - \Phi(r_0, \varphi_0) - A \begin{bmatrix} h_r \\ h_\varphi \end{bmatrix} =$$

Wstawiamy wzory określające Φ i A :

$$\begin{aligned} &= ((r_0 + h_r) \cos(\varphi_0 + h_\varphi), (r_0 + h_r) \sin(\varphi_0 + h_\varphi)) - \\ &\quad (r_0 \cos \varphi_0, r_0 \sin \varphi_0) - \begin{bmatrix} \cos \varphi_0 & -r_0 \sin \varphi_0 \\ \sin \varphi_0 & r_0 \cos \varphi_0 \end{bmatrix} \begin{bmatrix} h_r \\ h_\varphi \end{bmatrix} = \end{aligned}$$

i wykonujemy wskazane działania:

$$\begin{aligned} &= ((r_0 + h_r) \cos(\varphi_0 + h_\varphi) - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi, \\ &\quad (r_0 + h_r) \sin(\varphi_0 + h_\varphi) - r_0 \sin \varphi_0 - \sin \varphi_0 h_r - r_0 \cos \varphi_0 h_\varphi) \end{aligned}$$

Każdą ze współrzędnych, niebieską i czerwoną, omówimy oddzielnie. Zaczynamy od niebieskiej, stosujemy wzór

$$\cos(\varphi_0 + h_\varphi) = \cos \varphi_0 \cos h_\varphi - \sin \varphi_0 \sin h_\varphi$$

$$\begin{aligned} (r_0 + h_r) \cos(\varphi_0 + h_\varphi) - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi &= \\ (r_0 + h_r)(\cos \varphi_0 \cos h_\varphi - \sin \varphi_0 \sin h_\varphi) - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi &= \\ r_0 \cos \varphi_0 \cos h_\varphi - r_0 \sin \varphi_0 \sin h_\varphi + h_r \cos \varphi_0 \cos h_\varphi - & \\ h_r \sin \varphi_0 \sin h_\varphi - r_0 \cos \varphi_0 - \cos \varphi_0 h_r + r_0 \sin \varphi_0 h_\varphi &= \end{aligned}$$

W ostatnich dwóch linijkach wzoru grupujemy wyrazy: pierwszy z piątym, drugi z ostatnim i trzeci z przedostatnim:

$$= r_0 \cos \varphi_0 (\cos h_\varphi - 1) + r_0 \sin \varphi_0 (h_\varphi - \sin h_\varphi) + h_r \cos \varphi_0 (\cos h_\varphi - 1) - h_r \sin \varphi_0 \sin h_\varphi =$$

Wypisujemy przybliżone wartości pomijając wyrazy z potęgami przyrostów h_r i h_φ wyższymi niż trzecia (korzystamy z odpowiednich rozwinięć w szereg):

$$\approx r_0 \cos \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + r_0 \sin \varphi_0 \left(\frac{h_\varphi^3}{6} \right) + h_r \cos \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + \sin \varphi_0 h_r \left(h_\varphi - \frac{h_\varphi^3}{6} \right).$$

Okazuje się, że w wyrażeniu na pierwszą współrzędną nie ma składników rzędu niższego niż dwa. To samo robimy z czerwoną współrzędną. Rachunki pomijamy (są bardzo podobne) i otrzymujemy

$$(r_0 + h_r) \sin(\varphi_0 + h_\varphi) - r_0 \sin \varphi_0 - \sin \varphi_0 h_r - r_0 \cos \varphi_0 h_\varphi \approx \\ r_0 \sin \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + r_0 \cos \varphi_0 \left(-\frac{h_\varphi^3}{6} \right) + \sin \varphi_0 h_r \left(-\frac{h_\varphi^2}{2} \right) + \cos \varphi_0 h_r \left(h_\varphi - \frac{h_\varphi^3}{6} \right).$$

Warunek na resztę jest postaci:

$$\lim_{h \rightarrow 0} \frac{\|r((r_0, \varphi_0), (h_r, h_\varphi))\|}{\|h\|} = 0.$$

Wygodnie zamiast metryki euklidesowej użyć metryki maksimum. Wówczas

$$\|h\| = \max\{|h_r|, |h_\varphi|\},$$

metryki maksimum używamy także dla reszty. Jeśli wykazemy, że każda ze współrzędnych reszty podzielona przez $\|h\|$ znika, gdy h dąży do 0, będziemy wiedzieć, że większa z tych współrzędnych będąca długością reszty znika. Zaczniemy od niebieskiej współrzędnej:

$$\frac{r_0 \cos \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + r_0 \sin \varphi_0 \left(\frac{h_\varphi^3}{6} \right) + h_r \cos \varphi_0 \left(-\frac{h_\varphi^2}{2} \right) + \sin \varphi_0 h_r \left(h_\varphi - \frac{h_\varphi^3}{6} \right)}{\|h\|}$$

Najniższa potęga przyrostu w liczniku jest dwa (także wyrażenie $h_r h_\varphi$ traktujemy jak kwadratowe, bo oba czynniki dążą do zera). Mianownik jest rzędu jeden (jest równy $|h_r|$ lub $|h_\varphi|$). Oznacza to, że cały ułamek jest rzędu jeden w przyroście. Gdy przyrost dąży do zera, cały ułamek także. Identycznie rzecz się ma z czerwoną współrzędną. W ten sposób pokazaliśmy, że warunek na resztę w definicji różniczkowalności jest spełniony. Odwzorowanie Φ jest różniczkowalne w (r_0, φ_0) i jego pochodna jest równa

$$\Phi'(r_0, \varphi_0) = \begin{bmatrix} \cos \varphi_0 & -r_0 \sin \varphi_0 \\ \sin \varphi_0 & r_0 \cos \varphi_0 \end{bmatrix}$$



Powyższy przykład i wcześniejsze rozumowanie pokazują, że różniczkowanie funkcji i odwzorowań zależnych od wielu zmiennych polega na obliczaniu odpowiednich pochodnych cząstkowych i zapisywaniu ich w macierzach. Badanie różniczkowalności w sposób przedstawiony w przykładzie (4) jest nużące – na szczęście mamy do dyspozycji twierdzenie:

Twierdzenie 1. *Jeśli dla odwzorowania $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ istnieją wszystkie pochodne cząstkowe $\frac{\partial F^i}{\partial x^j}$ w punkcie (x_0) i są ciągłe w x_0 jako funkcje określone na \mathbb{R}^n , to odwzorowanie F jest różniczkowalne w x_0 i jego pochodna traktowana jako odwzorowanie*

$$\mathbb{R}^n \ni x \longmapsto F'(x) \in L(\mathbb{R}^n, \mathbb{R}^m).$$

jest ciągła w x_0 . Mówimy w takim przypadku, że F jest różniczkowalna w x_0 w sposób ciągły.

Przestrzeni odwzorowań liniowych, która jest oczywiście przestrzenią wektorową można liczyć długość wektorów, jeśli jest określona długość w dziedzinie i obrazie. Jeśli $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ to

$$(4) \quad \|A\| = \sup_{\|v\|_n=1} \|A(v)\|_m,$$

gdzie $\|\cdot\|_n$ jest długością wektora w dziedzinie a $\|\cdot\|_m$ długością wektora w przestrzeni wartości odwzorowania A . W powyższym twierdzeniu ciągłość odwzorowania o wartościach w $L(\mathbb{R}^n, \mathbb{R}^m)$ określana jest względem tej właśnie długości. Warto także wiedzieć, że ponieważ \mathbb{R}^n i \mathbb{R}^m są przestrzeniami skończenie-wymiarowymi, to także $L(\mathbb{R}^n, \mathbb{R}^m)$ jest przestrzenią skończenie-wymiarową (wymiaru nm), ponadto może być ona w sposób naturalny utożsamiona z \mathbb{R}^{nm} . Oznacza to, że mamy w niej trzy metryki d_1, d_2, d_∞ względem których można definiować ciągłość. Okazuje się, że zbiór odwzorowań ciągłych względem krórejkoľwiek z metryk i zbiór odwzorowań ciągłych względem długości zdefiniowanej w (4) są identyczne.

Dla funkcji jednej zmiennej obowiązują prawa ułatwiające różniczkowanie skomplikowanych funkcji: reguła Leibniza różniczkowania iloczynu, wzór na pochodną funkcji złożonej, wzór na pochodną ilorazu itd. W świecie wielu zmiennych także są podobne prawa:

Twierdzenie 2 (Formalne prawa różniczkowania). *Dla funkcji i odwzorowań określonych na \mathbb{R}^n obowiązują następujące reguły:*

- (1) *Jeśli odwzorowania $F, G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ są różniczkowalne w x_0 to także*

$$\alpha F + \beta G, \quad (\alpha F + \beta G)(x) = \alpha F(x) + \beta G(x)$$

jest różniczkowalne i

$$(\alpha F + \beta G)'(x_0) = \alpha F'(x_0) + \beta G'(x_0).$$

- (2) *Jeśli odwzorowanie $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ jest różniczkowalne w x_0 oraz funkcja $f : \mathbb{R} \rightarrow \mathbb{R}$ jest różniczkowalna w x_0 to odwzorowanie*

$$fF : \mathbb{R}^n \ni x \mapsto f(x)F(x) \in \mathbb{R}^m$$

także jest różniczkowalne w x_0 i pochodna ma postać

$$(fF)'(x_0) = f'(x_0)F(x_0) + f(x)F'(x_0),$$

tzn.

$$(fF)'(x_0)h = [f'(x_0)h]F(x_0) + f(x)[F'(x_0)h].$$

- (3) *Jeśli $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ i $G : \mathbb{R}^m \rightarrow \mathbb{R}^k$ oraz F jest różniczkowalne w x_0 a G jest różniczkowalne w $F(x_0)$ to złożenie*

$$G \circ F : \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad G \circ F(x) = G(F(x))$$

także jest odwzorowaniem różniczkowalnym w x_0 . Jego pochodna wyraża się wzorem

$$(G \circ F)'(x_0) = G'(F(x_0)) \cdot F'(x_0),$$

gdzie kropka po prawej stronie równania oznacza mnożenie macierzy.

Powyższego twierdzenia nie będziemy dowodzić. Zajmiemy się jedynie analizą przykładów. Punkt (1) wyraża zasadę liniowości pochodnej i nie wymaga komentowania. Punkt (2) jest wielozmiennową wersją reguły Leibniza. Zwróćmy uwagę na znaczenie wzoru na pochodną:

$$(fF)'(x_0) = f'(x_0)F(x_0) + f(x)F'(x_0) \quad \text{tzn.} \quad (fF)'(x_0)h = [f'(x_0)h]F(x_0) + f(x)[F'(x_0)h].$$

Pochodna $f'(x_0)$ jest odwzorowaniem liniowym z \mathbb{R}^n do \mathbb{R} . Działając na przyrost daje liczbę przez którą można mnożyć wartość odwzorowania F . Pochodna $F'(x_0)$ jest odwzorowaniem liniowym z \mathbb{R}^n do \mathbb{R}^m . Działając na przyrost h daje element \mathbb{R}^m , który następnie może być pomnożony (każda jego współrzędna) przez wartość $f(x_0)$.

Przykład 5. Niech Φ będzie odwzorowaniem związanym z biegunowym układem współrzędnych. Pomnóżmy Φ przez funkcję $f(r, \varphi) = r\varphi$ i obliczmy pochodną. Zrobimy to na dwa sposoby: wprost i korzystając ze wzoru: Najpierw wprost:

$$f(r, \varphi)\Phi(r, \varphi) = r\varphi(r \cos \varphi, r \sin \varphi) = (r^2\varphi \cos \varphi, r^2\varphi \sin \varphi)$$

Pochodna ma zatem postać

$$\begin{bmatrix} 2r\varphi \cos \varphi & r^2(\cos \varphi - \varphi \sin \varphi) \\ 2r\varphi \sin \varphi & r^2(\sin \varphi + \varphi \cos \varphi) \end{bmatrix}.$$

Używając wzoru z punktu (2) twierdzenia musimy zauważyć, że pierwszy człon tzn $f'(x_0)F(x_0)$ należy zapisać w odwrotnej kolejności jeśli chce się korzystać z rachunku macierzowego. Najpierw bowiem działamy $f'(x_0)$ na przyrost a dopiero potem mnożymy przez $F(x_0)$. W naszym przykładzie mamy

$$f(r, \varphi) = r\varphi, \quad f'(r, \varphi) = [\varphi \ r].$$

pochodną mnożymy z lewej strony przez $\Phi(r, \varphi)$ zapisane jako wektor w kolumnie:

$$\begin{bmatrix} r \cos \varphi \\ r \sin \varphi \end{bmatrix} [\varphi \ r] = \begin{bmatrix} \varphi r \cos \varphi & r^2 \cos \varphi \\ \varphi r \sin \varphi & r^2 \sin \varphi \end{bmatrix}$$

Dodajemy drugi składnik sumy, tzn $f(x_0)F'(x_0)$:

$$\begin{bmatrix} \varphi r \cos \varphi & r^2 \cos \varphi \\ \varphi r \sin \varphi & r^2 \sin \varphi \end{bmatrix} + r\varphi \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix} = \begin{bmatrix} 2\varphi r \cos \varphi & r^2 \cos \varphi - r^2\varphi \sin \varphi \\ 2\varphi r \sin \varphi & r^2 \sin \varphi + r^2\varphi \cos \varphi \end{bmatrix}$$

Obie metody doprowadziły do tego samego wyniku. ♣

punkt (3) twierdzenia o formalnych prawach różniczkowania mówi o różniczkowaniu złożenia.

Przykład 6. Rozważmy funkcję

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = ax^2 + by^2, \quad a, b > 0$$

Jej pochodną w biegunowym układzie współrzędnych możemy obliczyć na dwa sposoby: wprost i korzystając ze wzoru na pochodną złożenia. Funkcja f w biegunowym układzie współrzędnych jest złożeniem:

$$f(\Phi(r, \varphi)) = f(r \cos \varphi, r \sin \varphi) = ar^2 \cos^2 \varphi + br^2 \sin^2 \varphi = r^2(a \cos^2 \varphi + b \sin^2 \varphi).$$

Otrzymaną funkcję różniczkujemy:

$$\begin{aligned} [2r(a \cos^2 \varphi + b \sin^2 \varphi), \quad r^2(-2a \cos \varphi \sin \varphi + 2b \cos \varphi \sin \varphi)] = \\ [2r(a \cos^2 \varphi + b \sin^2 \varphi), \quad r^2 \sin(2\varphi)(b - a)]. \end{aligned}$$

Używając wzoru na pochodną złożenia obliczmy

$$\Phi'(r, \varphi) = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix}, \quad f'(x, y) = [2ax, \quad 2by], \quad f'(\Phi(r, \varphi)) = [2ar \cos \varphi, \quad 2br \sin \varphi],$$

i pomnożymy:

$$\begin{aligned} (f \circ \Phi)'(r, \varphi) &= f'(\Phi(r, \varphi))\Phi'(r, \varphi) = [2ar \cos \varphi, \quad 2br \sin \varphi] \cdot \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix} = \\ &[2ar \cos^2 \varphi + 2br \sin^2 \varphi, \quad -2ar^2 \cos \varphi \sin \varphi + 2br^2 \cos \varphi \sin \varphi] = \\ &[2r(a \cos^2 \varphi + b \sin^2 \varphi), \quad r^2 \sin(2\varphi)(b - a)] \end{aligned}$$

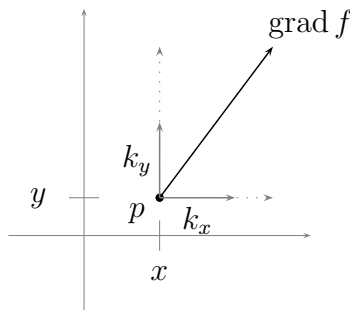
Obie metody doprowadziły do tego samego wyniku. ♣

Przykład 7 (Pojęcie gradientu funkcji). Z całą pewnością byli już państwo zmuszeni na ćwiczeniach z fizyki używać gradientu funkcji określonej na \mathbb{R}^2 , lub \mathbb{R}^3 , zapisanego w biegunowym (w \mathbb{R}^2), walcowym i sferycznym (w \mathbb{R}^3) układzie współrzędnych. Pamiętam z moich czasów studenckich, że rachunki tego rodzaju sprawiały mi kłopoty. Później okazało się, że źródłem tych kłopotów był brak zrozumienia struktur geometrycznych, których używa się do definicji gradientu. Być może chodziło o to, że tak naprawdę pojęcie to zaliczane jest do dziedziny matematyki nazywanej *geometrią różniczkową*. My obracamy się raczej w ramach klasycznej analizy. Opisywanie układów fizycznych wymaga jednak używania wielu narzędzi matematycznych już na początkowym etapie nauki - nie możemy czekać aż odbędzie państwo kurs geometrii różniczkowej.

Zastanówmy się więc najpierw co to jest gradient funkcji. Dla ustalenia uwagi będziemy pracować na \mathbb{R}^2 . Przestrzeń tę potraktujemy jak przestrzeń afiniczną na której określimy funkcję $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Gradient $\text{grad } f$ w punkcie $p = (x, y)$ jest wektorem zaczepionym w p o współrzędnych

$$\text{grad } f(p) = \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{bmatrix}$$

Przestrzeń wektorowa wektorów zaczepionych w punkcie p będzie oznaczana V . Jest to oczywiście także \mathbb{R}^2 . Wprowadzamy jednak abstrakcyjne oznaczenie, gdyż w przestrzeni tej będziemy używać różnych baz. Oto stosowny rysunek:



$$\text{grad } f = \frac{\partial f}{\partial x}(p)k_x + \frac{\partial f}{\partial y}(p)k_y$$

Składowe tego wektora względem bazy $k = (k_x, k_y)$ związanej z kartezjańskim układem współrzędnych są jednocześnie składowymi pochodnej funkcji f . Czasami myśli się więc o gradiencie jako o pochodnej. Jest to jednak błąd, który prowadzi do trudności w rachunkach. Gradient jest bowiem wektorem (elementem przestrzeni wektorowej) a pochodna jest odwzorowaniem liniowym na przestrzeni wektorowej. W zapisie widzimy także różnicę: gradient (wektor) jest „pionowy” a pochodna (odwzorowanie liniowe) jest „poziome”.

Przestrzeń wektorowa \mathbb{R}^2 (i każda \mathbb{R}^n) wyposażona jest w kanoniczny iloczyn skalarny. Zadaje on izomorfizm między przestrzenią wektorową $V = \mathbb{R}^2$ wektorów zaczepionych w p a przestrzenią funkcji liniowych na V , do której należy pochodna funkcji f . Funkcja liniowa odpowiadająca

wektorowi $v \in V$ to funkcja

$$V \ni w \longmapsto (v | w) \in \mathbb{R}.$$

Można ją oznaczyć $(v | \cdot)$, gdzie kropka oznacza miejsce „czekające na argument”. Odwzorowanie między V a $L(V, \mathbb{R})$ jest izomorfizmem, zatem jest odwracalne: można mówić o funkcjach liniowych odpowiadających wektorom ale także o wektorach odpowiadających funkcjom liniowym. Gradient jest właśnie takim „wektorem odpowiadającym pochodnej”. Abstrakcyjny wzór wyglądałby tak:

$$f'(p) = (\text{grad } f(p) | \cdot).$$

Jeśli formę dwuliniową będącą iloczynem skalarnym oznaczymy dodatkowo przez G (potrzebujemy jakąś literę do zapisywania macierzy) to w bazie k mamy:

$$[f'(p)]_k = ([\text{grad } f(p)]^k)^T [G]_k,$$

a w innej, dowolnej bazie v :

$$[f'(p)]_v = ([\text{grad } f(p)]^v)^T [G]_v.$$

Skupmy się na równaniu dla bazy v . Mamy wyznaczyć $[\text{grad } f(p)]^v$. Transponujemy więc obie strony równania:

$$([f'(p)]_v)^T = ([G]_v)^T [\text{grad } f(p)]^v = [G]_v [\text{grad } f(p)]^v$$

Skorzystaliśmy także z faktu, że $[G]_v$ jest macierzą symetryczną, więc transpozycja nie wpływa na postać macierzy. Macierz $[G]_v$ jest odwracalna, więc

$$(5) \quad [\text{grad } f(p)]^v = ([G]_v)^{-1} ([f'(p)]_v)^T.$$

Powyższy wzór może posłużyć do wyznaczenia postaci gradientu w dowolnym wybranym przez nas układzie współrzędnych.

Zajmijmy się teraz praktyczną stroną zapisywania gradientu we współrzędnych biegunowych. Istnieją co najmniej dwie metody z których możemy skorzystać. Pierwsza z nich, którą nazywać będę *metodą zamiany zmiennych* jest następująca: Nowe zmienne, biegunowe, związane są z kartezjańskimi odwzorowaniem

$$\Phi : U \longrightarrow \mathbb{R}^2, \quad \mathbb{R}^2 \supset U = \{(r, \varphi) : r > 0, \varphi \in]0, 2\pi[\}.$$

$$\Phi(r, \varphi) = (r \cos \varphi, r \sin \varphi).$$

Różniczkując $f \circ \Phi$ jak funkcję złożoną dostajemy równania opisujące $\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \varphi}$. Przyjęło się oznaczać wyjściową funkcję i tę złożoną z Φ tym samym symbolem, dla podkreślenia, że chodzi o ten sam obiekt, tylko zapisany w różnych układach współrzędnych.

$$\begin{aligned} \left[\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \varphi} \right] &= [(f \circ \Phi)'(r, \varphi)] = \\ &= \left[\frac{\partial f}{\partial x}(\Phi(r, \varphi)), \frac{\partial f}{\partial y}(\Phi(r, \varphi)) \right] \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix} = \\ &= \left[\cos \varphi \frac{\partial f}{\partial x} + \sin \varphi \frac{\partial f}{\partial y}, -r \sin \varphi \frac{\partial f}{\partial x} + r \cos \varphi \frac{\partial f}{\partial y} \right]. \end{aligned}$$

Żeby wyznaczyć pochodne po x i y w zależności od pochodnych po r i φ należy rozwiązać układ równań:

$$\begin{cases} \frac{\partial f}{\partial r} = \cos \varphi \frac{\partial f}{\partial x} + \sin \varphi \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial \varphi} = -r \sin \varphi \frac{\partial f}{\partial x} + r \cos \varphi \frac{\partial f}{\partial y} \end{cases}$$

Otrzymujemy:

$$\begin{cases} \frac{\partial f}{\partial x} = \cos \varphi \frac{\partial f}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial f}{\partial \varphi} \\ \frac{\partial f}{\partial y} = \sin \varphi \frac{\partial f}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial f}{\partial \varphi} \end{cases}$$

Wstawiając to wszystko do wzoru na gradient otrzymujemy

$$\text{grad } f = \frac{\partial f}{\partial x} k_x + \frac{\partial f}{\partial y} k_y = \left(\cos \varphi \frac{\partial f}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial f}{\partial \varphi} \right) k_x + \left(\sin \varphi \frac{\partial f}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial f}{\partial \varphi} \right) k_y$$

Dalej jeszcze musimy zamienić bazę na bazę ortonormalną związaną z biegunowym układem współrzędnych:

$$(6) \quad e_r = \cos \varphi k_x + \sin \varphi k_y$$

$$(7) \quad e_\varphi = -\sin \varphi k_x + \cos \varphi k_y$$

O tym, skąd wzięły się powyższe zależności będzie później. Wyznaczamy teraz k_x i k_y :

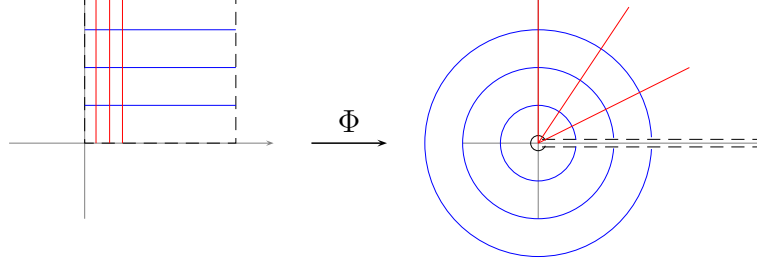
$$\begin{aligned} k_x &= \cos \varphi e_r - \sin \varphi e_\varphi \\ k_y &= \sin \varphi e_r + \cos \varphi e_\varphi \end{aligned}$$

Wstawiamy wyniki do wzoru na gradient i wykonujemy działania:

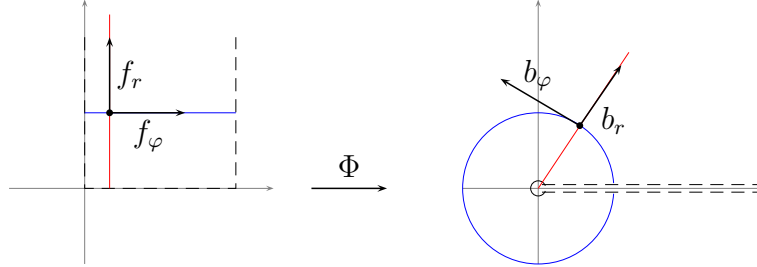
$$\begin{aligned} \text{grad } f &= \left(\cos \varphi \frac{\partial f}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial f}{\partial \varphi} \right) k_x + \left(\sin \varphi \frac{\partial f}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial f}{\partial \varphi} \right) k_y = \\ &= \left(\cos \varphi \frac{\partial f}{\partial r} - \frac{1}{r} \sin \varphi \frac{\partial f}{\partial \varphi} \right) (\cos \varphi e_r - \sin \varphi e_\varphi) + \left(\sin \varphi \frac{\partial f}{\partial r} + \frac{1}{r} \cos \varphi \frac{\partial f}{\partial \varphi} \right) (\sin \varphi e_r + \cos \varphi e_\varphi) \\ &= \frac{\partial f}{\partial r} e_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} e_\varphi. \end{aligned}$$

Druga metoda polega na *skorzystaniu z definicji gradientu*, czyli użyciu wzoru (4) W przestrzeni A wprowadzamy współrzędne na dwa różne sposoby: kartezjańskie zadające w przestrzeni V bazę $k = (k_x, k_y)$ i biegunowe, z którymi związana jest baza $b = (b_r, b_\varphi)$. Przestrzeń V wyposażona jest w iloczyn skalarny, który w bazie k związanej ze współrzędnymi kartezjańskimi ma macierz równą macierzy jednostkowej.

Przyjrzyjmy się obrazkom ilustrującym sytuację: Biegunowy układ współrzędnych związany jest z odwzorowaniem Φ obciętym do zbioru $\{(r, \varphi) : r > 0, \varphi \in]0, 2\pi[\}$:



Odwzorowanie $\Phi'(r, \varphi)$ jest izomorfizmem między przestrzeniami wektorów zaczepionych w (r, φ) i w $(x, y) = \Phi(r, \varphi)$.



Baza b otrzymana jest z działania pochodnej $\Phi'(r, \varphi)$ na wektory bazy standardowej $f = (f_r, f_\varphi)$ w przestrzeni wektorów zaczepionych w (r, φ) w dziedzinie odwzorowania Φ . Baza b składa się z wektorów prostopadłych do siebie, ale wektor b_φ nie jest długości 1. Sprawdźmy:

$$(8) \quad b_\varphi = \Phi'(r, \varphi)f_\varphi = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -r \sin \varphi \\ r \cos \varphi \end{bmatrix} = -r \sin \varphi e_x + r \cos \varphi e_y$$

Długość wektora b_φ liczymy korzystając z iloczynu skalarnego:

$$\|b_\varphi\|^2 = (b_\varphi | b_\varphi) = (-r \sin \varphi e_x + r \cos \varphi e_y | -r \sin \varphi e_x + r \cos \varphi e_y) = r^2$$

Wektor b_φ jest długości r . Wektor b_r jest długości 1 co można sprawdzić zapisując go w bazie e :

$$(9) \quad b_r = \Phi'(r_0, \varphi_0)f_r = \begin{bmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} = \cos \varphi e_x + \sin \varphi e_y$$

$$(b_r | b_r) = (\cos \varphi e_x + \sin \varphi e_y | \cos \varphi e_x + \sin \varphi e_y) = 1$$

Równania (8) i (9) wraz z informacją, że $\|b_\varphi\| = r$ stanowią uzasadnienie dla użytych wcześniej wzorów (6), (7).

Pochodna funkcji f zapisana w bazie k ma postać

$$[f'(p)]_k = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$$

zaś zapisana w bazie b związanej z biegunowym układem współrzędnych ma postać:

$$[f'(p)]_b = \left[\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \varphi} \right].$$

Wiedząc, że

$$[G]_b = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}, \quad \text{więc} \quad [G]_b^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{bmatrix}$$

wyznaczamy

$$[\text{grad } f]^b = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial r} \\ \frac{\partial f}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial r} \\ \frac{1}{r^2} \frac{\partial f}{\partial \varphi} \end{bmatrix},$$

zatem

$$\text{grad } f = \frac{\partial f}{\partial r} b_r + \frac{1}{r^2} \frac{\partial f}{\partial \varphi} b_\varphi.$$

Zazwyczaj chcemy, aby baza, z którą pracujemy była nie tylko ortogonalna ale i ortonormalna. Zamiast bazy b musimy więc wziąć bazę $e' = (e_r, e_\varphi)$, gdzie $e_r = b_r$, $e_\varphi = \frac{1}{r} b_\varphi$:

$$\text{grad } f = \frac{\partial f}{\partial r} b_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} \left(\frac{1}{r} b_\varphi \right) = \frac{\partial f}{\partial r} e_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} e_\varphi.$$

Metoda korzystania z definicji wymagała od nas trochę rozważań abstrakcyjnych. Jednak teraz możemy bez dalszych rachunków zapisać gradient w jakichkolwiek współrzędnych. Potrzebujemy do tego jedynie macierzy iloczynu skalarnego zapisanej w bazie związanej z tym układem współrzędnych. Jeśli bazę tę oznaczmy v gradient uzyskujemy według wzoru (4):

$$[\text{grad } f(p)]^v = ([G]_v)^{-1}([f'(p)]_v)^T,$$

gdzie wektor $([f'(p)]_v)^T$ jest transpozycją pochodnej, składającej się z pochodnych cząstkowych funkcji f względem nowych współrzędnych. ♣

Wykład 5.

Matematyka 2, semestr letni 2010/2011

Rachunek różniczkowy funkcji jednej zmiennej służy, między innymi, do badania przebiegu zmienności funkcji. Potrafimy znajdować punkty krytyczne, określać ich rodzaj, badać kształt wykresu (wypukły, wklęsły), znajdować asymptoty itp. Korzystając z umiejętności liczenia granic i różniczkowania potrafimy dość dokładnie naszkicować wykres funkcji. Funkcje wielu zmiennych będziemy badać jedynie w niewielkim zakresie. Zajmiemy się poszukiwaniem i określaniem typu punktów krytycznych.

Definicja 1. Mówimy, że $x_0 \in \mathbb{R}^n$ jest *maksimum lokalnym* (*minimum lokalnym*) funkcji $f : \mathbb{R}^n \rightarrow \mathbb{R}$, jeśli istnieje otoczenie \mathcal{O} punktu x_0 takie, że dla wszystkich $x \in \mathcal{O}$ zachodzi $f(x) \leq f(x_0)$ ($f(x) \geq f(x_0)$).

Doprecyzowania wymaga pojęcie otoczenia. Zazwyczaj otoczeniem punktu x_0 nazywa się dowolny zbiór otwarty zawierający x_0 . pojęcie zbioru otwartego pojawia się w analizie bardzo często, dlatego podamy definicję takiego zbioru.

Definicja 2. Mówimy, że zbiór \mathcal{O} jest *otwarty względem metryki d* jeśli każdy punkt $x \in \mathcal{O}$ jest środkiem pewnej kuli otwartej $K_d(x, \varepsilon)$ zawartej w \mathcal{O} .

Kulę otwartą zdefiniowaliśmy omawiając ciągłość. Mówiąc językiem nieprecyzyjnym zbiory otwarte są to takie zbiory które są „grube” i „nie zawierają brzegów”. Mówiąc „grube”, mam na myśli, że np. prosta w \mathbb{R}^2 nie jest otwarta względem metryki euklidesowej. Podobnie płaszczyzna czy powierzchnia sfery w \mathbb{R}^3 . Ta sama rodzina zbiorów otwartych może odpowiadać różnym metrykom. Metryki równoważne w takim sensie jak opisany przy okazji ciągłości mają takie same rodziny zbiorów otwartych. Mówiąc „zbiór otwarty w \mathbb{R}^n ” będziemy mieć na myśli zbiór otwarty względem którejkolwiek z metryk d_1, d_2, d_∞ .

Wracamy do ekstremów: Załóżmy, że punkt x_0 jest maksimum lokalnym funkcji f i funkcja f jest w tym punkcie różniczkowalna. Rozpatrzmy funkcję

$$\varphi_i : t \mapsto f(x_0 + te_i)$$

Jest ona określona na pewnym odcinku I zawierającym $t = 0$ i ze względu na własności f (punkt x_0 jest maksimum) na pewnym odcinku $I' \subset I$ spełnia warunek $\varphi_i(t) \leq \varphi_i(0)$. Oznacza to, że funkcja φ_i ma maksimum w $t = 0$. Funkcja φ_i jest też różniczkowalna, zatem zgodnie z zasadami obowiązującymi dla funkcji jednej zmiennej $\varphi_i'(0) = 0$. Pochodna φ_i w $t = 0$ jest równa pochodnej cząstkowej funkcji f po zmiennej x_i w punkcie x_0 :

$$\varphi_i'(0) = \lim_{t \rightarrow 0} \frac{\varphi_i(t) - \varphi_i(0)}{t} = \lim_{t \rightarrow 0} \frac{f(x_0 + te_i) - f(x_0)}{t} = \frac{\partial f}{\partial x^i}(x_0).$$

Okazuje się więc, że $\frac{\partial f}{\partial x^i}(x_0) = 0$. Jest to prawda dla każdego indeksu i . Podobnie byłoby w minimum. Otrzymaliśmy następujące twierdzenie:

Twierdzenie 1. Jeśli $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ma w punkcie x_0 ekstremum lokalne i jest w tym punkcie różniczkowalna to wszystkie pochodne cząstkowe tej funkcji w tym punkcie są równe zero, a co za tym idzie także pochodna $f'(x_0)$ jest równa zero.

Powyższe twierdzenie formułuje warunek konieczny istnienia ekstremum. Że nie jest on wystarczający pokazują następujące przykłady.

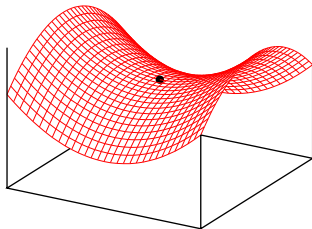
Przykład 1 (Siodło). Najprostszy przykład to funkcja $f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$ zadana wzorem

$$f_1(x, y) = x^2 - y^2.$$

Pochodna tej funkcji

$$f'_0(x, y) = [2x \quad -2y]$$

przyjmuje wartość 0 w punkcie $(0, 0)$. Jednak w tym punkcie funkcja nie ma ekstremum, co pokazuje wykres:



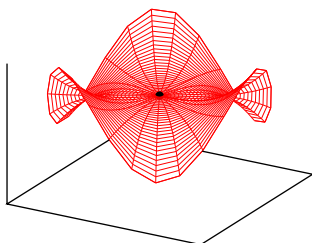
Przykład 2 (Małpie siodło). Nieco ciekawszy kształt ma wykres funkcji

$$f_2(x, y) = z = x(x^2 - 3y^2).$$

Jej pochodna

$$f'_2(0, 0) = [3x^2 - 3y^2 \quad 6yx]$$

także przyjmuje wartość 0 w punkcie $(0, 0)$:



Przykład 3 (Ciekawa funkcja). Przyjrzyjmy się funkcji

$$f_3(x, y) = (x - y^2)(x - 3y^2).$$

W punkcie $(0, 0)$ jest oczywiście punkt krytyczny:

$$f'_3 = [2x - 4y^2 \quad -4y(2x + 3y^2)].$$

Sprawdźmy jak funkcja zachowuje się wzdłuż prostej

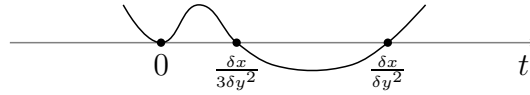
$$t \longmapsto \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$$

$$f(t\delta x, t\delta y) = (t\delta x - t^2\delta y^2)(t\delta x - 3t^2\delta y^2)$$

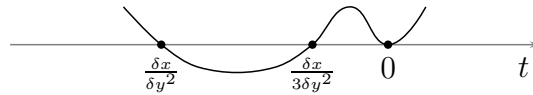
Założmy, że $\delta y \neq 0$ (co wyklucza na razie z rozważań prostą $y = 0$) wtedy zależność od t można opisać wzorem:

$$f(t\delta x, t\delta y) = 3t^2(\delta y)^4 \left(\frac{\delta x}{\delta y^2} - t \right) \left(\frac{\delta x}{3\delta y^2} - t \right).$$

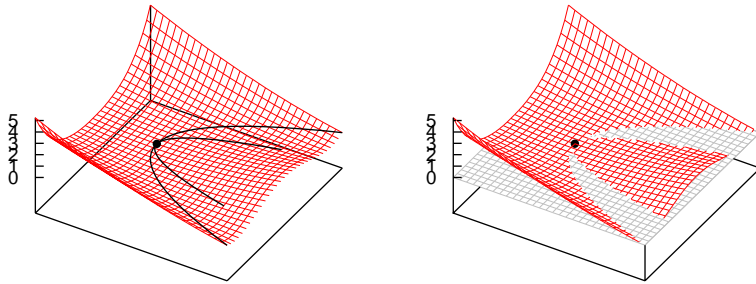
Widać więc, że wzdłuż każdej prostej funkcja ma w zerze minimum. Dla dodatnich δx sytuacja wygląda tak:



Dla ujemnych tak:



Gdy $\delta x = 0$ mamy $f(0, t\delta y) = 3t^4(\delta y)^4$ i oczywiście w dla $t = 0$ jest minimum. Dla $\delta y = 0$ otrzymujemy $f(t\delta x, 0) = t^2\delta t^2$ i także dla $t = 0$ jest minimum. A jednak w każdym dowolnie małym otoczeniu punktu $(0, 0)$ są punkty dla których wartość f jest ujemna. Ze wzoru definiującego funkcję wynika, że na krzywych $x = y^2$ i $x = 3y^2$ funkcja przyjmuje wartość zero. Pomiędzy nimi zaś wartości ujemne. Przyjrzyjmy się wykresom: na pierwszym z nich zaznaczono krzywe na których funkcja przyjmuje wartość zero. Na drugim na szaro narysowana jest płaszczyzna $z = 0$:



Potrzebujemy zatem więcej informacji na temat zachowania funkcji w punkcie krytycznym niż tylko wartość jej pochodnej. W rachunku różniczkowym funkcji jednej zmiennej dodatkowych kryteriów dostarczały wyższe pochodne. Tak samo jest tutaj. Zdefiniujmy zatem drugą pochodną funkcji odwzorowania $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Załóżmy, że F jest różniczkowalne w pewnym otoczeniu \mathcal{O} punktu x_0 . Pochodna w każdym punkcie otoczenia \mathcal{O} jest odwzorowaniem liniowym z \mathbb{R}^n do \mathbb{R}^m . Wyznaczając pochodne w każdym punkcie otoczenia \mathcal{O} otrzymujemy odwzorowanie

$$F' : \mathbb{R}^n \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m), \quad x \longmapsto F'(x).$$

Zbiór wartości powyższego odwzorowania jest przestrzenią wektorową izomorficzną z \mathbb{R}^{nm} (kolejnym wyrazom macierzowym macierzy odwzorowania liniowego przyporządkowujemy współrzędne w \mathbb{R}^{nm}). Możemy także „mierzyć długość odwzorowania liniowego” posługując się następującą definicją

$$A \in L(\mathbb{R}^n, \mathbb{R}^m), \quad \|A\| = \sup_{\|x\|=1} \|Ax\|.$$

Dla porządku zauważmy, że powyższa definicja zależy od tego jak mierzymy odległość w dziedzinie i w zbiorze wartości. Okazuje się jednak, że w sposób mierzenia odległości w przestrzeni odwzorowań liniowych w ogóle nie musimy wnikać, gdyż zbieżność i ciągłość definiowana przez tę odległość jest taka sama jak ta definiowana przez d_1 , d_2 i d_∞ . Traktujemy poprostu zbiór wartości jak \mathbb{R}^{mn} z odpowiednio dużym wykładnikiem i pracujemy na nim jak zawsze. W takim razie możemy także różniczkować odwzorowanie

$$F' : \mathbb{R}^n \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{mn}$$

w punkcie x_0 zgodnie z obowiązującymi zasadami, otrzymując drugą pochodną odwzorowania F .

Definicja 3. Jeśli odwzorowanie $F' : \mathbb{R}^n \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{mn}$ jest różniczkowalne w punkcie x_0 , to jego pochodną oznaczamy $F''(x_0)$ i nazywamy drugą pochodną odwzorowania F w x_0 .

Na pierwszy rzut oka druga pochodna jest bardzo skomplikowanym obiektem: odwzorowanie liniowe o wartościach w przestrzeni odwzorowań liniowych. Jeśli drugą pochodną w punkcie x_0 obliczymy na przyroście h otrzymamy odwzorowanie liniowe, które znowu można obliczyć na przyroście k otrzymując element z \mathbb{R}^m . Sytuację znacznie upraszcza obserwacja, że zależność od h i k jest *de facto* dwuliniowa. Druga pochodna jest więc ostatecznie odwzorowaniem dwuliniowym

$$F''(x_0) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

Obowiązuje ponadto twierdzenie:

Twierdzenie 2. Jeśli F jest dwukrotnie różniczkowalna w punkcie x_0 to druga pochodna $F''(x_0)$ jest odwzorowaniem dwuliniowym symetrycznym.

Dlatego właśnie włożyliśmy wcześniej sporo wysiłku w to, żeby poznać własności tego rodzaju odwzorowań. Ograniczymy się teraz do funkcji wielu zmiennych, tzn założymy, że wartości są w \mathbb{R} (tzn $m = 1$). Druga pochodna jest więc odwzorowaniem dwuliniowym o wartościach w \mathbb{R} , czyli formą dwuliniową. Jak wygląda macierz tej formy w bazie kanoniczej? Ponieważ druga pochodna jest pochodną pochodnej spodziewamy się, że w użyciu będą drugie pochodne cząstkowe. Istotnie, okazuje się, że dla funkcji

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

dwukrotnie różniczkowalnej w x_0 druga pochodna w x_0 ma postać:

$$F''(x_0) = \begin{bmatrix} \frac{\partial^2 f}{(\partial x^1)^2}(x_0) & \frac{\partial^2 f}{\partial x^2 \partial x^1}(x_0) & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^1}(x_0) \\ \frac{\partial^2 f}{\partial x^1 \partial x^2}(x_0) & \frac{\partial^2 f}{(\partial x^2)^2}(x_0) & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^2}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1}(x_0) & \frac{\partial^2 f}{\partial x^n \partial x^1}(x_0) & \cdots & \frac{\partial^2 f}{(\partial x^n)^2}(x_0) \end{bmatrix}$$

Z twierdzenia (2) wynika, że macierz ta powinna być symetryczna, to znaczy drugie pochodne cząstkowe mieszane nie powinny zależeć od kolejności różniczkowania. Istotnie wygląda, że tak jest. Sprawdźmy to dla funkcji z przykładów (1), (2), (3). Twierdzenia (2) nie będziemy dowodzić w całej ogólności. Pożyteczne będzie jednak przyjrzenie się najprostszej sytuacji funkcji dwóch zmiennych

$$\mathbb{R}^2 \ni (x, y) \longmapsto f(x, y) \in \mathbb{R}.$$

Założymy dodatkowo, że funkcja nie tylko jest dwukrotnie różniczkowalna w punkcie (x_0, y_0) ale także jej drugie pochodne cząstkowe istnieją także w otoczeniu tego punktu i są ciągłe.

Twierdzenie 3. *Założmy, że funkcja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ma w otoczeniu \mathcal{O} punktu (x_0, y_0) pochodne cząstkowe rzędu pierwszego i drugiego. Założmy ponadto, że pochodne rzędu drugiego są ciągłe na \mathcal{O} . Wówczas mieszane pochodne rzędu drugiego są równe.*

Dowód. Do dowodu potrzebne nam jest twierdzenie Lagrange’a dla funkcji jednej zmiennej. Twierdzenie to mówi, że jeśli funkcja $\varphi : [a, b] \rightarrow \mathbb{R}$ jest ciągła na $[a, b]$ i różniczkowalna na $]a, b[$ to istnieje punkt $c \in]a, b[$ taki, że

$$\varphi'(c)(b - a) = (\varphi(b) - \varphi(a)).$$

Geometrycznie oznacza to, że w odcinku $]a, b[$ jest punkt w którym styczna do wykresu jest równoległa do prostej przechodzącej przez punkty $(a, f(a))$ i $(b, f(b))$. Oznaczmy przez $W(t, s)$ następującą wielkość:

$$W(t, s) = \frac{1}{st} [f(x_0 + t, y_0 + s) - f(x_0 + t, y_0) - f(x_0, y_0 + s) + f(x_0, y_0)].$$

Pochodną mieszaną

$$\frac{\partial f}{\partial x \partial y}(x_0, y_0)$$

liczymy przechodząc do granicy

$$\lim_{t \rightarrow 0} \lim_{s \rightarrow 0} W(t, s)$$

a pochodną mieszaną w drugiej kolejności licząc

$$\lim_{s \rightarrow 0} \lim_{t \rightarrow 0} W(t, s).$$

Pytanie o równość pochodnych mieszanych sprowadza się do pytania o możliwość zamiany kolejności przechodzenia do granicy. Oznaczmy

$$\varphi(x) = \frac{f(x, y_0 + s) - f(x, y_0)}{s}.$$

Funkcja φ jest różniczkowalna i jej pochodna po x zapisuje się wzorem

$$\varphi'(x) = \frac{\frac{\partial f}{\partial x}(x, y_0 + s) - \frac{\partial f}{\partial x}(x, y_0)}{s}.$$

Korzystając z φ możemy zapisać $W(s, t)$ jako

$$W(s, t) = \frac{1}{t} [\varphi(x_0 + t) - \varphi(x_0)].$$

Dla funkcji φ możemy skorzystać z twierdzenia Lagrange’a:

$$W(s, t) = \frac{1}{t} \varphi'(x_0 + \theta_1 t)(x_0 + t - x_0) = \varphi'(x_0 + \theta_1 t).$$

Punkt c znajdujący się pomiędzy x_0 a $x_0 + t$ zapisaliśmy w postaci $x_0 + \theta_1 t$, gdzie θ_1 jest jakąś liczbą pomiędzy 0 i 1. Podstawiając wartość pochodnej φ' mamy

$$W(s, t) = \varphi'(x_0 + \theta_1 t) = \frac{\frac{\partial f}{\partial x}(x_0 + \theta_1 t, y_0 + s) - \frac{\partial f}{\partial x}(x_0 + \theta_1 t, y_0)}{s}$$

Korzystamy teraz z faktu, że pochodne cząstkowe f także są różniczkowalne. Można więc użyć dla nich twierdzenia Lagrange'a zastępując wyrażenie w liczniku przez wartość pochodnej po y w pewnym punkcie mnożoną przez s :

$$W(s, t) = \frac{\partial^2 f}{\partial y \partial x}(x_0 + \theta_1 t, y_0 + \theta_2 s),$$

gdzie θ_2 też jest jakąś liczbą pomiędzy 0 a 1.

Tak samo możemy podsąpić wprowadzając pomocniczą funkcję ψ

$$\psi(y) = \frac{f(x_0 + t, y) - f(x_0, y)}{t}.$$

I wykonując podobne czynności zamieniając rolami x i y . Dostajemy wtedy

$$W(s, t) = \frac{\partial^2 f}{\partial x \partial y}(x_0 + \theta_3 t, y_0 + \theta_4 s),$$

dla pewnych wartości θ_3, θ_4 pomiędzy 0 a 1. Ostatecznie

$$W(s, t) = \frac{\partial^2 f}{\partial y \partial x}(x_0 + \theta_1 t, y_0 + \theta_2 s) = \frac{\partial^2 f}{\partial x \partial y}(x_0 + \theta_3 t, y_0 + \theta_4 s).$$

Z założenia wiadomi, że drugie pochodne cząstkowe są ciągłe, zatem wyrażenie W też jest ciągłe i kolejność przechodzenia do granicy z t i s nie ma znaczenia. Granica w $s = 0$ i $t = 0$ jest równa drugim pochodnym cząstkowym, które w związku z tym są równe. \square

Twierdzenie Taylora. Znajomość pochodnej funkcji $f : \mathbb{R}^n \rightarrow \mathbb{R}$ w punkcie x_0 umożliwia nam przybliżanie zachowania tej funkcji w otoczeniu x_0 za pomocą funkcji afinicznej (liniowa + stała). Czasami jednak (na przykład przy badaniu rodzaju punktu krytycznego) potrzebujemy subtelniejszego przybliżenia. W rachunku różniczkowym jednej zmiennej mieliśmy do dyspozycji twierdzenie Taylora o przybliżaniu funkcji różniczkowalnej k -razy wielomianem stopnia k . Dla funkcji wielu zmiennych mamy podobne twierdzenie. Ograniczymy się tutaj do funkcji różniczkowalnych dwa razy, choć można oczywiście sformułować twierdzenie bardziej ogólne:

Twierdzenie 4. Niech $f : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją różniczkowalną dwa razy w punkcie x_0 . Wówczas

$$f(x_0 + h) = f(x_0) + f'(x_0)(h) + \frac{1}{2!}f''(x_0)(h, h) + r_2(x_0, h)$$

i reszta $r_2(x_0, h)$ spełnia warunek

$$\lim_{h \rightarrow 0} \frac{|r_2(x_0, h)|}{\|h\|^2} = 0.$$

Przykład 4. Wróćmy do pierwszej funkcji dwóch zmiennych, którą rozważaliśmy na naszym wykładzie w kontekście przybliżania funkcją afiniczną (przy definiowaniu pochodnej):

$$f(x, y) = x^2 y.$$

Zapiszmy dla niej wyrażenie jak w twierdzeniu (3) i znajdziemy resztę w otoczeniu punktu $(1, 2)$.
Obliczamy pierwsze pochodne cząstkowe

$$\frac{\partial f}{\partial x} = 2xy, \quad \frac{\partial f}{\partial y} = x^2$$

i ich wartości w punkcie $(1, 2)$:

$$\frac{\partial f}{\partial x}(1, 2) = 4, \quad \frac{\partial f}{\partial y}(1, 2) = 1.$$

Macierz pochodnej w punkcie $(1, 2)$ ma więc postać:

$$f'(1, 2) = [4 \ 1].$$

Obliczamy drugie pochodne cząstkowe

$$\frac{\partial^2 f}{\partial x^2} = 2y, \quad \frac{\partial^2 f}{\partial y^2} = 0, \quad \frac{\partial^2 f}{\partial x \partial y} = 2x$$

i ich wartości w punkcie $(1, 2)$:

$$\frac{\partial^2 f}{\partial x^2}(1, 2) = 4, \quad \frac{\partial^2 f}{\partial y^2}(1, 2) = 0, \quad \frac{\partial^2 f}{\partial x \partial y}(1, 2) = 2.$$

Macierz drugiej pochodnej w punkcie $(1, 2)$ ma więc postać:

$$f''(1, 2) = \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix}$$

Zapisujemy wzór Taylora:

$$f(1 + \delta x, 2 + \delta y) = 2 + [4 \ 1] \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} + \frac{1}{2} [\delta x \ \delta y] \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} + r((1, 2), (\delta x, \delta y)).$$

Po wykonaniu wskazanych działań mamy:

$$(1) \quad f(1 + \delta x, 2 + \delta y) = 2 + 4\delta x + \delta y + \frac{1}{2} (4\delta x^2 + 4\delta x\delta y) + r_2((1, 2), (\delta x, \delta y)).$$

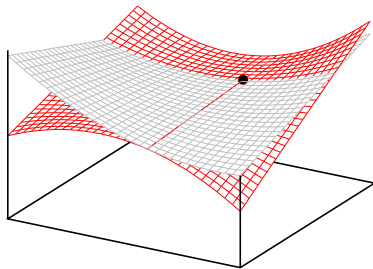
Funkcja f jest dosyć prosta, dlatego możemy *explicite* zapisać wzór na resztę:

$$(2) \quad f(1 + \delta x, 2 + \delta y) = (1 + \delta x)^2(2 + \delta y) = (1 + 2\delta x + \delta x^2)(2 + \delta y) = \\ 2 + 4\delta x + 2\delta x^2 + \delta y + 2\delta x\delta y + \delta x^2\delta y$$

Wyrazy niebieskie we wzorach (1) i (2) pokrywają się, zatem reszta jest postaci

$$r_2((1, 2), (\delta x, \delta y)) = \delta x^2\delta y.$$

Ponieważ reszta jest trzeciego rzędu w przyroście, to rzeczywiście podzielona przez kwadrat długości (drugiego rzędu) znika przy $(\delta x, \delta y)$ dążącym do 0. Rysunek przedstawia wykres funkcji f (na czerwono) i jej przybliżenie drugiego rzędu (na szaro).



Znajomość zachowania funkcji z dokładnością do drugiej pochodnej umożliwia w wielu przypadkach określenie typu punktu krytycznego. Jeśli bowiem x_0 jest punktem krytycznym funkcji f to wzór Taylora przyjmuje postać:

$$f(x_0 + h) = f(x_0) + f''(x_0)(h, h) + r_2(x_0, h).$$

Biorąc wystarczająco małe otoczenie punktu x_0 możemy zagwarantować, że o zachowaniu f decyduje wyłącznie druga pochodna. Jeśli $f''(x_0)$ jest **dodatnio określona** to w tym otoczeniu wartości $f(x_0 + h)$ będą większe niż $f(x_0)$, zatem w x_0 jest **minimum**. Jeśli w $f''(x_0)$ jest **ujemnie określona**, to w tym otoczeniu wartości $f(x_0 + h)$ będą mniejsze niż $f(x_0)$, czyli w x_0 jest **maksimum**. We wzorze Taylora używana jest forma kwadratowa odpowiadająca drugiej pochodnej (obliczamy wartość drugiej pochodnej dwa razy na tym samym przyroście). Znak tej formy możemy określić patrząc na sygnaturę: sygnatura $(n, 0)$ oznacza formę dodatnio określoną, sygnatura $(0, n)$ ujemnie. Sygnatury (p, q) dla $p \neq 0$ i $q \neq 0$ odpowiadają formom przyjmującym dodatnie i ujemne wartości. W sytuacjach kiedy druga pochodna jest nieokreślona (ale niezdegenerowana, tzn $p + q = n$) mamy inny rodzaj punktu krytycznego. W sytuacji kiedy druga pochodna jest zdegenerowana kryterium drugiego rzędu (wykorzystujące jedynie drugą pochodną) nie rozstrzyga o rodzaju punktu krytycznego. Możemy wtedy szukać innych metod. Jedną z nich może być wykorzystanie wyższych pochodnych. Jest to często skomplikowane rachunkowo.

Przykład 5. Poszukajmy punktów krytycznych funkcji

$$g(x, y) = \sin x \sin y \sin(x + y), \quad \text{dla } 0 < x < \pi, 0 < y < \pi.$$

Zaczynamy od wyznaczenia pierwszych pochodnych cząstkowych funkcji g :

$$\begin{aligned} \frac{\partial g}{\partial x} &= \cos x \sin y \sin(x + y) + \sin x \sin y \cos(x + y) = \\ &= \sin y [\cos x \sin(x + y) + \sin x \cos(x + y)] = \sin y \sin(2x + y). \end{aligned}$$

Funkcja g jest symetryczna ze względu na zamianę x i y , zatem

$$\frac{\partial g}{\partial x} = \sin x \sin(x + 2y).$$

Z układu równań

$$\sin y \sin(2x + y) = 0, \quad \sin x \sin(x + 2y) = 0$$

(po uwzględnieniu ograniczonej dziedziny) wynika, że

$$2x + y = k\pi, \quad x + 2y = l\pi \quad \text{dla pewnych } k, l \in \mathbb{Z}$$

Powyższy układ równań zapisany w postaci macierzowej przyjmuje postać:

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k\pi \\ l\pi \end{bmatrix}.$$

Macierz tego układu jest niezdegenerowana, zatem posługując się macierzą odwrotną można napisać rozwiązanie:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} k\pi \\ l\pi \end{bmatrix} = \frac{1}{3} \begin{bmatrix} (2k-l)\pi \\ (-k+2l)\pi \end{bmatrix}.$$

W dziedzinie funkcji $]0, \pi[\times]0, \pi[$ są dwa punkty krytyczne, które otrzymujemy biorąc $k = 1, l = 1$ i $k = 2, l = 2$. Są to $p_1 = (\frac{\pi}{3}, \frac{\pi}{3})$ i $p_2 = (\frac{2\pi}{3}, \frac{2\pi}{3})$. Zbadamy charakter pierwszego z nich. Wyznaczamy drugą pochodną funkcji g w punkcie p_1 :

$$\frac{\partial^2 g}{\partial x^2} = 2 \sin y \cos(2x + y)|_{(\frac{\pi}{3}, \frac{\pi}{3})} = 2 \sin \frac{\pi}{3} \cos \left(\frac{2\pi}{3} + \frac{\pi}{3} \right) = 2 \frac{\sqrt{3}}{2} (-1) = -\sqrt{3}.$$

Ze względu na symetrię stwierdzamy, że

$$\frac{\partial^2 g}{\partial y^2} \left(\frac{\pi}{3}, \frac{\pi}{3} \right) = -\sqrt{3}.$$

Obliczamy pochodne mieszane:

$$\frac{\partial^2 g}{\partial x \partial y} = \cos y \sin(2x + y) + \sin y \cos(2x + y) = \sin(2x + 2y)|_{(\frac{\pi}{3}, \frac{\pi}{3})} = \sin \left(\frac{4\pi}{3} \right) = -\frac{\sqrt{3}}{2}$$

Macierz drugiej pochodnej funkcji g w punkcie p_1 to

$$g''(p_1) = \begin{bmatrix} -\sqrt{3} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \sqrt{3} \end{bmatrix}$$

Szukamy sygnatury metodą wyznacznikową:

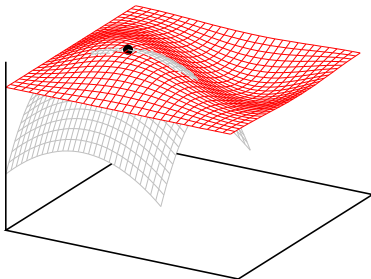
$$D_1 = -\sqrt{3}, \quad D_2 = 3 - \frac{3}{4} = \frac{9}{4}.$$

Odpowiednie liczby wyznaczające sygnaturę to:

$$D_1 = -\sqrt{3} < 0, \quad \frac{D_2}{D_1} = -\frac{9}{4\sqrt{3}} < 0.$$

Forma kwadratowa zadawana przez drugą pochodną g w punkcie p_1 ma sygnaturę $(0, 2)$, jest więc ujemnie określona. Stwierdzamy zatem, że funkcja g ma w punkcie p_1 maksimum. W ramach treningu czytelnik może zbadać rodzaj ekstremum dla funkcji g w punkcie $(\frac{2\pi}{3}, \frac{2\pi}{3})$.

Wykres przedstawia przybliżenie funkcji g (na czerwono) przez odpowiednią funkcję kwadratową (na szaro) w okolicach punktu krytycznego.



Przykład 6. Korzystając z procedury poszukiwania punktów krytycznych funkcji wielu zmiennych wyprowadzimy wzór na współczynniki równania regresji liniowej. Załóżmy, że na płaszczyźnie mamy n punktów (x_i, y_i) . Szukamy równania prostej $y = f(x) = ax + b$ takiej, aby suma kwadratów różnic $y_i - f(x_i)$ była możliwie mała. Innymi słowy szukamy minimum funkcji

$$F(a, b) = \sum_{i=1}^n (y_i - f(x_i))^2 = \sum_{i=1}^n (y_i - (ax_i + b))^2.$$

Obliczamy pierwsze pochodne cząstkowe funkcji F :

$$\begin{aligned} \frac{\partial F}{\partial a} &= \sum_{i=1}^n 2(y_i - (ax_i + b))(-x_i) = 2 \sum_{i=1}^n (-y_i x_i + a(x_i)^2 + bx_i) = \\ &= 2 \left[\sum_{i=1}^n -y_i x_i + a \sum_{i=1}^n (x_i)^2 + b \sum_{i=1}^n x_i \right] \end{aligned}$$

$$\frac{\partial F}{\partial b} = \sum_{i=1}^n 2(y_i - (ax_i + b))(-1) = 2 \sum_{i=1}^n (-y_i + ax_i + b) = 2 \left[\sum_{i=1}^n -y_i + a \sum_{i=1}^n x_i + bn \right]$$

Punkt krytyczny spełnia warunki

$$\frac{\partial F}{\partial a} = 0, \quad \frac{\partial F}{\partial b} = 0,$$

czyli

$$\sum_{i=1}^n -y_i x_i + a \sum_{i=1}^n (x_i)^2 + b \sum_{i=1}^n x_i = 0, \quad \sum_{i=1}^n -y_i + a \sum_{i=1}^n x_i + bn = 0$$

Dzieląc obie strony powyższych równań przez n , i oznaczając średnią arytmetyczną współrzędnych x_i symbolem $E(x)$, średnią arytmetyczną współrzędnych y_i symbolem $E(y)$ i podobnie $E(xy)$, $E(x^2)$ otrzymujemy:

$$aE(x^2) + bE(x) = E(xy), \quad aE(x) + b = E(y).$$

Z powyższych równań wyznaczamy a i b :

$$b = E(y) - aE(x), \quad aE(x^2) + (E(y) - aE(x))E(x) = E(xy)$$

$$a(E(x^2) - [E(x)]^2) = E(xy) - E(x)E(y).$$

Różnica $E(x^2) - [E(x)]^2$ oznaczana jest $D^2(x)$ i nazywana *wariancją* x , natomiast $E(xy) - E(x)E(y)$ oznaczana $C(x, y)$ to *kowariancja* x i y . Ostatecznie

$$a = \frac{C(x, y)}{D^2(x)}, \quad b = E(y) - \frac{C(x, y)}{D^2(x)}E(x).$$

Otrzymaliśmy współrzędne punktu krytycznego funkcji F . Z samej postaci funkcji nietrudno odgadnąć, że punkt ten jest w istocie minimum. Dla porządku sprawdzimy jednak drugą pochodną:

$$\frac{\partial^2 F}{\partial a^2} = 2 \sum_{i=1}^n (x_i)^2, \quad \frac{\partial F}{\partial b} = 2n, \quad \frac{\partial^2 F}{\partial a \partial b} = 2 \sum_{i=1}^n x_i$$

Macierz drugiej pochodnej ma postać

$$\begin{bmatrix} 2 \sum_{i=1}^n (x_i)^2 & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2n \end{bmatrix} = 2n \begin{bmatrix} E(x^2) & E(x) \\ E(x) & 1 \end{bmatrix}.$$

Współczynnik $2n$ jest dodatni, nie wpływa więc na znak formy kwadratowej zadanej powyższą macierzą. Wyznacznik pierwszego minora to $D_1 = E(x^2)$ - jest on dodatni jako średnia dodatnich liczb. Wyznacznik drugiego minora (całej macierzy) to $D_2 = E(x^2) - [E(x)]^2 = D^2(x)$. Wariancja jest także zawsze dodatnia. Sygnatura określana jest przez znaki D_1 i $\frac{D_2}{D_1}$. Obie liczby są dodatnie, zatem forma jest dodatnio określona. Znaleziony przez nas punkt krytyczny jest rzeczywiście minimum funkcji F . ♣

Wykład 6.

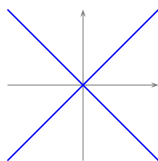
Matematyka 2, semestr letni 2010/2011

Brak fragmentu dotyczącego twierdzenia o odwzorowaniu odwrotnym

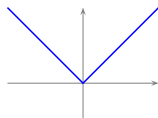
Niech $f : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją różniczkowalną na pewnym obszarze $\mathcal{O} \subset \mathbb{R}^2$. Przyjrzyjmy się zbiorowi

$$f^{-1}(0) = \{x \in \mathbb{R}^n : f(x) = 0\},$$

czyli zerowej poziomicy tej funkcji. Jeśli pochodna $f'(x)$ jest różna od zera w każdym punkcie poziomicy (to znaczy w każdym punkcie tej poziomicy nie znika przynajmniej jedna pochodna cząstkowa funkcji f), odpowiednie twierdzenia zapewniają, że poziomica ta jest powierzchnią $n - 1$ wymiarową w \mathbb{R}^n . Powierzchnię zdefiniujemy precyzyjniej w następnym wykładzie. Teraz wystarczą nam przykłady: W \mathbb{R}^2 powierzchniami jednowymiarowymi są: okrąg (tzn. czyli poziomica zerowa funkcji $F(x, y) = x^2 + y^2 - 1$), wszelkie krzywe stożkowe: elipsy, hiperbole, parabole. Krzywe stożkowe są poziomice funkcji kwadratowych na \mathbb{R}^2 . Oczywiście nie wszystkie poziomice funkcji kwadratowych są powierzchniami. Na przykład poziomica zerowa funkcji $g(x, y) = x^2 - y^2$ to dwie przecinające się proste:



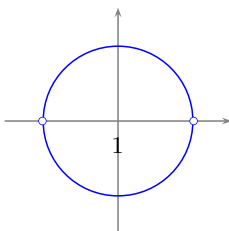
W okolicach punktu $(0, 0)$ zbiór ten nie jest powierzchnią. Wykres funkcji $x \mapsto |x|$ nie jest powierzchnią w otoczeniu $(0, 0)$ ze względu na „dzióbek”.



W \mathbb{R}^3 powierzchniami wymiaru 2 są wykresy różniczkowalnych funkcji z \mathbb{R}^2 do \mathbb{R} , kształty takie jak sfera, elipsoida, torus...

Zastanawiać się teraz będziemy nad odpowiedzią na następujące pytanie: kiedy poziomica zerowa funkcji f definiuje jedną ze zmiennych jako funkcje pozostałych. Przyjrzyjmy się funkcji f , której poziomica zerowa jest okręgiem:

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



Patrząc na wykres stwierdzamy, że poziomica zerowa definiuje dwie funkcje $x \rightarrow y_i(x)$, $i = 1, 2$. Wykresem pierwszej z nich jest górny półokrąg, wykresem drugiej dolny. Punktami kłopotliwymi są punkty $(-1, 0)$, $(0, 1)$. Mówimy w takim przypadku, że funkcje $x \rightarrow y_i(x)$ *zadane są przez równanie*

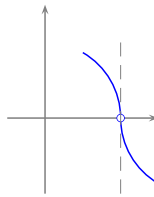
$$(1) \quad f(x, y) = 0$$

w sposób uwikłany. W tym przypadku postać funkcji f jest na tyle prosta, że równanie (1) można rozvikłać:

$$y_1(x) = \sqrt{x^2 - 1}, \quad y_2(x) = -\sqrt{x^2 - 1}.$$

Bywają jednak sytuacje w których rozwiązanie równania ze względu na jedną ze zmiennych jest trudne, albo wręcz niemożliwe. Chcielibyśmy jednak móc uzyskać jakieś informacje na temat funkcji zadawanych przez równanie. Interesować będą nas na przykład ekstrema.

Punkty $(-1, 0)$ i $(0, 1)$ nazwałam kłopotliwymi, ponieważ w ich otoczeniu nie można zdefiniować zmiennej y jako funkcji zmiennej x . Kłopot polega na tym, że każdej wartości x odpowiadałyby dwie wartości y . Innego rodzaju kłopot pojawia się w sytuacji kiedy wykres wygląda następująco:



W tej sytuacji nie ma problemu z dwuwartościowością i funkcja $x \rightarrow y(x)$ jest określona, ale styczna do wykresu w punkcie $(1, 0)$ jest pionowa, co oznacza, że funkcja ta jest nieróżniczkowalna. Wszystkie te zastrzeżenia podsumowuje następujące twierdzenie

Twierdzenie 1 (O funkcji uwikłanej, TFU). *Niech $f : \mathbb{R}^{n-1} \times \mathbb{R} \rightarrow \mathbb{R}$ będzie funkcją klasy \mathcal{C}^1 (różniczkowalną w sposób ciągły) na pewnym obszarze $\mathcal{O} \subset \mathbb{R}^{n-1} \times \mathbb{R}$. Niech także punkt (x_0, y_0) należy do \mathcal{O} . Jeśli*

$$\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$$

to istnieje otwarty zbiór $\mathcal{U} \subset \mathbb{R}^{n-1}$ zawierający x_0 i funkcja $g : \mathcal{U} \rightarrow \mathbb{R}$ różniczkowalna i taka, że zbiór

$$\{(x, y) \in \mathcal{O}, x \in \mathcal{U} : F(x, y) = 0\}$$

jest wykresem funkcji g .

Założmy teraz, że funkcja f zadaje w sposób uwikłany y jako funkcję różniczkowalną zmiennych (x^1, \dots, x^n) . Jak znaleźć punkty krytyczne $y(x)$? Warunek konieczny mówi, że w punkcie krytycznym pochodna $y'(x)$ jest macierzą zerową. Musimy więc wyznaczyć pochodne cząstkowe

$$\frac{\partial y}{\partial x^i}.$$

Użyjemy notacji

$$y_{x^i} = \frac{\partial y}{\partial x^i}, \quad F_{x^i} = \frac{\partial F}{\partial x^i}.$$

Różniczkujemy równanie

$$F(x^1, x^2, \dots, x^n, y(x^1, x^2, \dots, x^n)) = 0$$

po zmiennej x^i :

$$F_{x^i}(x^1, x^2, \dots, x^n, y(x^1, x^2, \dots, x^n)) + F_y(x^1, x^2, \dots, x^n, y(x^1, x^2, \dots, x^n))y_{x^i}(x^1, x^2, \dots, x^n) = 0.$$

Dla przejrzystości zapisu opuścimy argumenty:

$$(2) \quad F_{x^i} + F_y y_{x^i} = 0.$$

Wynika z tego, że

$$y_{x^i} = -\frac{F_{x^i}}{F_y}.$$

Pamiętać należy, że obie pochodne cząstkowe F_{x^i} i F_y są funkcjami zmiennych x i y . Aby otrzymać prawdziwą wartość pochodnej $y_x(x)$ trzeba wstawić wartość y odpowiadającą x , czyli taką, że $F(x, y) = 0$. Warunkiem istnienia funkcji uwikłanej $x \mapsto y(x)$ jest aby F_y nie zniknęła - można więc dzielić przez tę wartość. Punkt x_0 jest punktem krytycznym jeśli $y_{x^i}(x_0) = 0$. Oznacza to, że spełniony jest układ równań:

$$F_{x^i}(x_0, y_0) = 0 \text{ dla } i \in \{1, 2, \dots, n\}, \quad F(x_0, y_0) = 0.$$

Może się okazać, że znalezione przez nas punkty krytyczne należą do różnych funkcji. Przekonałiśmy się, że jedno równanie może określać wiele funkcji uwikłanych. Określenie rodzaju punktu krytycznego wymaga znalezienia drugiej pochodnej. Zrózniczkujmy jeszcze raz równanie (2), tym razem po x^j za każdym razem pamiętając, że pochodne cząstkowe funkcji F zależą od x^j także poprzez funkcję $y(x)$:

$$F_{x^i x^j} + F_{x^i y} y_{x^j} + (F_{y x^j} + F_{yy} y_{x^j}) y_{x^i} + F_y y_{x^i x^j} = 0$$

W punkcie krytycznym wszystkie pierwsze pochodne cząstkowe funkcji F i funkcji y znikają, zatem równanie przyjmuje postać:

$$F_{x^i x^j}(x_0, y_0) + F_y(x_0, y_0) y_{x^i x^j}(x_0, y_0) = 0$$

Możemy więc wyznaczyć drugą pochodną cząstkową $y_{x^i x^j}(x_0, y_0)$:

$$(3) \quad y_{x^i x^j}(x_0) = -\frac{F_{x^i x^j}(x_0, y_0)}{F_y(x_0, y_0)}.$$

Pamiętajmy, że wzór (3) obowiązuje *tylko* w punkcie krytycznym. Układając pochodne cząstkowe w macierz otrzymujemy drugą pochodną y w punkcie krytycznym:

$$y''(x_0) = \begin{bmatrix} F_{x^1 x^1}(x_0, y_0) & F_{x^1 x^2}(x_0, y_0) & \cdots & F_{x^1 x^n}(x_0, y_0) \\ F_{x^2 x^1}(x_0, y_0) & F_{x^2 x^2}(x_0, y_0) & \cdots & F_{x^2 x^n}(x_0, y_0) \\ \vdots & \vdots & \ddots & \vdots \\ F_{x^n x^1}(x_0, y_0) & F_{x^n x^2}(x_0, y_0) & \cdots & F_{x^n x^n}(x_0, y_0) \end{bmatrix}.$$

Określoność tej formy badamy jak zwykle.

Przykład 1. Niech

$$F(x, y) = x^2 + (y - \sqrt[4]{x^2})^2 - 1$$

Dziwne wyrażenie $\sqrt[4]{x^2}$ zastąpić można także $\sqrt{|x|}$. Zastanówmy się jak wygląda poziomica zero-wa tej funkcji. Spróbujemy dowiedzieć się o niej możliwie dużo stosując metody analityczne (bez patrzenia na wykres). Przyjemność narysowania wykresu przy pomocy komputera zostawimy

sobie na koniec. Używać będziemy wszelkich dostępnych metod, ze szczególnym uwzględnieniem badania funkcji zadanych w sposób uwikłany.

Poziomica zerowa funkcji F zadana jest równaniem

$$(4) \quad x^2 + \left(y - \sqrt[4]{x^2}\right)^2 = 1.$$

Wynika z tego, że żaden ze składników po lewej stronie równania nie przekracza jedynki:

$$x^2 \leq 1 \quad \text{czyli} \quad -1 \leq x \leq 1$$

oraz

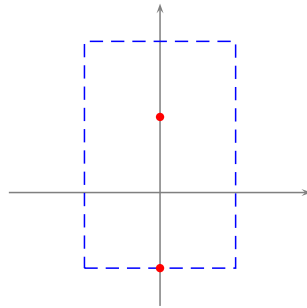
$$(y - \sqrt{|x|})^2 \leq 1 \quad \text{czyli} \quad -1 \leq y - \sqrt{|x|} \leq 1, \quad -1 + \sqrt{|x|} \leq y \leq 1 + \sqrt{|x|}.$$

Ostatecznie więc cały obrazek musi mieścić się w prostokącie $[-1, 1] \times [-1, 2]$. Równanie jest symetryczne ze względu na zamianę x na $-x$, zatem poziomica będzie symetryczna względem prostej $x = 0$.

Ponieważ funkcja $x \mapsto \sqrt{|x|}$ jest nieróżniczkowalna w $x = 0$, także poziomica będzie krzywą nieróżniczkowalną w punktach o współrzędnej $x = 0$. Czy są takie punkty na krzywej? Sprawdźmy

$$\text{dla } x = 0 \text{ mamy } F(0, y) = y^2 - 1 = 0, \quad \text{czyli } y = 1, \quad y = -1$$

Do krzywej należą więc punkty $(0, 1)$ i $(0, -1)$ i w tych punktach krzywa jest nieróżniczkowalna. Dotychczasowe informacje zaznaczmy na rysunku



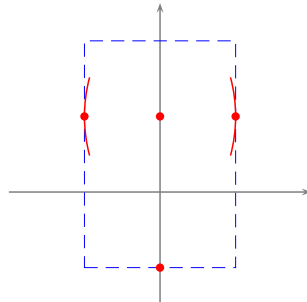
Sprawdźmy teraz czy równanie (4) określa y jako funkcję uwikłaną zmiennej x . Policzmy F_y :

$$F_y(x, y) = 2(y - \sqrt{|x|})$$

Pochodna F_y przyjmuje wartość zero w punktach w których $y = \sqrt{|x|}$. Szukamy takich punktów na krzywej:

$$F(x, \sqrt{|x|}) = x^2 + (\sqrt{|x|} - \sqrt{|x|})^2 - 1 = x^2 - 1 = 0 \quad \text{czyli} \quad x = 1, \quad x = -1.$$

Wynika z tego, że w otoczeniu punktów $(1, 1)$ i $(-1, 1)$ (o których wiemy, że należą do krzywej) równanie nie definiuje y jako funkcji od x . Oznacza to zazwyczaj następujący kształt wykresu:



Sprawdźmy, czy ewentualne funkcje $x \mapsto y(x)$ zadawane przez równanie mają punkty krytyczne. Znajdźmy pochodną F_x (dla ułatwienia policzymy dla dodatniego x):

$$F_x = 2x + 2(y - \sqrt{x}) \left(-\frac{1}{\sqrt{x}} \right) = 2x - \frac{y - \sqrt{x}}{\sqrt{x}}$$

Pochodna przyjmuje wartość 0 dla

$$y = \sqrt{x}(2x + 1)$$

Po wstawieniu do równania (4) otrzymujemy warunek na x postaci

$$4x^3 + x^2 - 1 = 0.$$

Niestety rozwiązania tego równania nie są wymierne. Można jednak oszacować mniej więcej wartość interesującego nas miejsca zerowego. Zauważmy, że wielomian

$$w(x) = 4x^3 + x^2 - 1$$

ma ekstrema w punktach spełniających równanie $w'(x) = 0$, tzn

$$0 = 12x^2 + 2x = 2x(6x + 1) \implies x = 0, \quad x = -\frac{1}{6}$$

W punkcie $x = -\frac{1}{6}$ wartość w jest ujemna, podobnie w $x = 0$. Wiadomo także, że $\lim_{x \rightarrow \infty} w = +\infty$ i dla $x > 0$ funkcja w jest ściśle rosnąca. Oznacza to, że między zerem a nieskończonością jest dokładnie jedno miejsce zerowe. Oznaczmy to miejsce zerowe x_0 . Sprawdzamy także, że $w(\frac{1}{2}) = -\frac{1}{4} < 0$ natomiast $w(\frac{3}{4}) = \frac{3}{4} > 0$. Wiemy więc, że $\frac{1}{2} < x_0 < \frac{3}{4}$. Punkt krytyczny jest więc w punkcie $(x_0, y_0 = \sqrt{x_0}(2x_0 + 1))$. Z symetrii funkcji F wynika, że podobny punkt krytyczny jest dla $-x_0$, czyli w punkcie $(-x_0, y_0)$. Sprawdźmy typ punktu krytycznego:

$$F_{xx} = \frac{\partial}{\partial x} \left(2x - \frac{y - \sqrt{x}}{\sqrt{x}} \right) = \frac{\partial}{\partial x} \left(2x - \frac{y}{\sqrt{x}} + 1 \right) = 2 + \frac{y}{x\sqrt{x}}.$$

W punkcie $(x_0, \sqrt{x_0}(2x_0 + 1))$

$$F_{xx}(x_0, \sqrt{x_0}(2x_0 + 1)) = 2 + \frac{2}{\sqrt{x_0}} + \frac{1}{x_0} > 0.$$

Wartość F_y w punkcie krytycznym to

$$F_y(x_0, \sqrt{x_0}(2x_0 + 1)) = 2(\sqrt{x_0}(2x_0 + 1) - \sqrt{x_0}) = 2x_0\sqrt{x_0} > 0$$

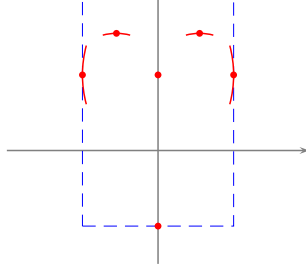
Wnioskujemy więc, że

$$y_{xx}(x_0) = -\frac{F_{xx}(x_0, y_0)}{F_y(x_0, y_0)} < 0,$$

zatem mamy do czynienia z maksimum. Podobnie jest dla $(-x_0, y_0)$. Wartość y_0 jest z całą pewnością dodatnia. Jeśli skorzystamy z informacji, że $x_0 > \frac{1}{2}$ to otrzymamy, że

$$y_0 > \sqrt{\frac{1}{2}} \left(2 \cdot \frac{1}{2} + 1 \right) = \sqrt{2}$$

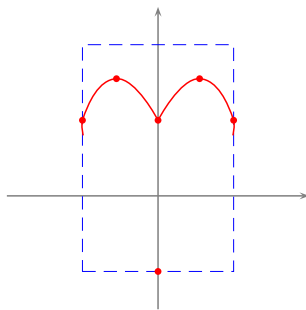
Zaznaczmy prawdopodobny przebieg krzywej na wykresie:



Policzmy teraz pochodną funkcji $x \mapsto y(x)$ zadawanej przez równanie (4), tam gdzie można to zrobić. Ograniczamy się do $x > 0$ (dla $x < 0$ wynik dostaniemy korzystając z symetrii).

$$y_x = -\frac{F_x}{F_y} = -\frac{2x\sqrt{x} - y + \sqrt{x}}{2\sqrt{x}(y - \sqrt{x})} = \frac{y - \sqrt{x}(2x + 1)}{2\sqrt{x}(y - \sqrt{x})}.$$

W „górnej” części wykresu, tzn. dla $x > 0$, ($y > \sqrt{x} > 0$), mianownik przyjmuje wartości dodatnie, i o znaku pochodnej decyduje licznik. Licznik zmienia znak w punkcie (x_0, y_0) i dla $x < x_0$ jest dodatni a dla $x > x_0$ ujemny. Funkcja $x \mapsto y(x)$ rośnie więc od wartości $y = 1$ w $x = 0$ do maksimum y_0 w x_0 a następnie maleje do wartości 1 w $x = 1$. Pochodna y_x dąży do $+\infty$ dla x zbliżających się do 0 od prawej strony. Dla $x < 0$ wykres jest symetryczny. Oznacza to, że w $x = 0$ funkcja ma minimum o kształcie „dzióbka”:



Dolną część wykresu zrekonstruujemy przyglądając się funkcjom $y \mapsto x(y)$ zadanych przez nasze równanie. Badając F_x szukamy kłopotliwych punktów:

$$F_x = 2x - \frac{y - \sqrt{x}}{\sqrt{x}}.$$

Wiemy już, że $F_x = 0$ dla (x_0, y_0) i dla $(-x_0, y_0)$. Widać także, że niedobre są punkty dla których $x = 0$, w tych punktach nie możemy różniczkować. Szukamy teraz ekstremów funkcji $y \mapsto x(y)$ (nie wiedząc oczywiście na razie ile tych funkcji jest). Pochodna F_y ma postać

$$F_y(x, y) = 2(y - \sqrt{|x|}).$$

Punkty krytyczne zatem to $(1, 1)$ i $(-1, 1)$, czyli te, które wcześniej znaleźliśmy jako niedobre do wyznaczania funkcji $x \mapsto y(x)$. Zbadamy rodzaj ekstremum w punkcie $(1, 1)$:

$$F_{yy}(x, y) = 2, \quad F_x = 2x - \frac{y - \sqrt{x}}{\sqrt{x}} \Big|_{(1,1)} = 2,$$

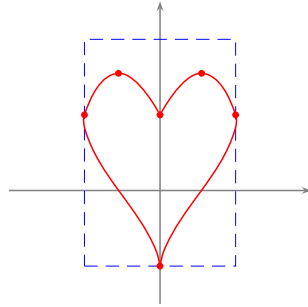
zatem

$$x_{yy}(1, 1) = -\frac{2}{2} = -1 < 0.$$

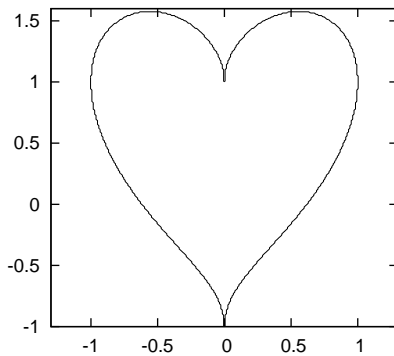
Punkt $(1, 1)$ jest więc maksimum funkcji. Z symetrii wykresu wnioskujemy, że w $(-1, 1)$ otrzymamy minimum odpowiedniej funkcji. Nasze przewidywania dotyczące kształtu wykresu okazały się więc słuszne. Zbadajmy jeszcze pochodną x_y dla $y < 1$. Pochodna ta opisana jest wzorem:

$$x_y = -\frac{F_y}{F_x} = -\frac{2\sqrt{x}(y - \sqrt{x})}{\sqrt{x}(2x + 1) - y}.$$

Znak powyższego wyrażenia łatwo określić gdy $y < 0$ i $x > 0$. Mianownik jest wtedy dodatni, licznik ujemny. Minus przed ułamkiem powoduje, że cała pochodna jest ujemna. Funkcja $y \mapsto x(y)$ jest więc rosnąca przynajmniej dla $y < 0$. Dla x dążącego do 0 i y dążącego do -1 pochodna dąży do zera. Krzywa zmierza więc do punktu $(0, -1)$ „w sposób styczny” do osi pionowej. Spróbujmy uzupełnić wykres:



A ciekawe jak to wygląda naprawdę?



Przykład 2. Znaleźć i zbadać punkty krytyczne funkcji

$$(x, y) \mapsto z(x, y)$$

określonych równaniem

$$(x + z)(y + z)\left(1 - \frac{z}{xy}\right) = 8.$$

Przykład ten był rozwiązany na wykładzie, być może wkrótce notatki zostaną uzupełnione

Wykład 7.

Matematyka 2, semestr letni 2010/2011

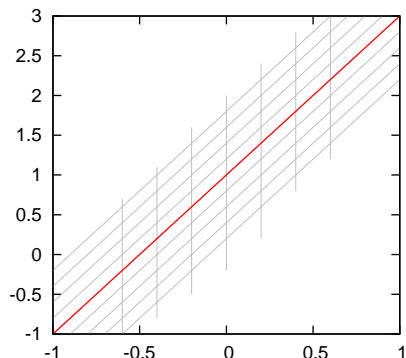
Powierzchnią wymiaru k zanurzoną w \mathbb{R}^n ($k < n$) nazywamy podzbiór $S \subset \mathbb{R}^n$ mający następującą własność: dla każdego punktu $p \in S$ istnieje otoczenie \mathcal{O} punktu p w \mathbb{R}^n oraz układ współrzędnych Φ w \mathcal{O} taki, że zbiór $\mathcal{O} \cap S$ opisany jest warunkiem znikania $n - k$ ostatnich współrzędnych związanych z układem współrzędnych Φ . Mówiąc *układ współrzędnych* mam na myśli odwzorowanie

$$\Phi : \mathcal{V} \longrightarrow \mathcal{O}.$$

spełniające założenia twierdzenia o lokalnej odwracalności na całym \mathcal{V} . Oznacza to, że odwzorowanie Φ jest bijekcją \mathcal{V} i \mathcal{O} , jest różniczkowalne i jego odwrotność też jest różniczkowalna. Przyjrzyjmy się konkretnym przykładom powierzchni. Zaczynamy od powierzchni jednowymiarowych w \mathbb{R}^2 :

Przykład 1. Najprostszym przykładem powierzchni jednowymiarowej w \mathbb{R}^2 jest prosta:

$$L = \{(x, y) : y = 2x + 1\}.$$



Żeby pokazać, że jest to powierzchnia musimy wprowadzić w otoczeniu każdego punktu układ współrzędnych taki, żeby prosta L zadana była warunkiem znikania drugiej współrzędnej. Ze względu na szczególnie nieskomplikowaną powierzchnię układ współrzędnych może być globalny, tzn zdefiniowany na całym \mathbb{R}^2 a nie tylko w otoczeniu jednego punktu. Istnieje wiele odpowiednich układów współrzędnych. Siatka współrzędnych związana z jednym z nich zaznaczona jest na rysunku. Nowe współrzędne punktu $p = (x, y)$ oznaczmy (ξ, η) . Współrzędna ξ jest identyczna z x . Współrzędną η punktu p obliczymy znajdując punkt przecięcia prostej równoległej do L i przechodzącej przez p z osią pionową. Wartości przesuniemy tak, aby 0 odpowiadało właśnie prostej L . Takie określenie układu współrzędnych prowadzi do wzorów:

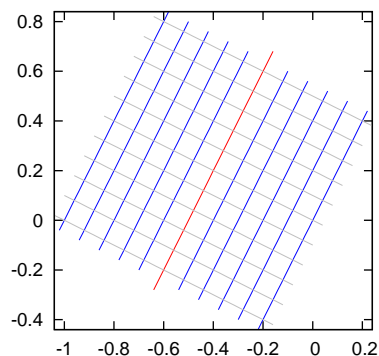
$$\begin{cases} \xi = x \\ \eta = y - 2x - 1 \end{cases}$$

Odwzorowanie odwrotne (które zazwyczaj kojarzymy z nazwą „układ współrzędnych”) ma postać

$$\Phi : \mathbb{R}^2 \ni (\xi, \eta) \longmapsto (x(\xi, \eta), y(\xi, \eta)) \in \mathbb{R}^2$$

$$\begin{cases} x = \xi \\ y = \eta + 2\xi + 1 \end{cases}$$

Możliwy jest także inny układ współrzędnych. Jeśli zażądamy, aby druga współrzędna zmieniła się wzdłuż prostych prostopadłych do L otrzymamy



Odpowiednie odwzorowanie $\Psi : (s, t) \mapsto (x(s, t), y(s, t))$ zapisuje się wzorami:

$$\begin{cases} x = \frac{1}{5}s - \frac{2}{5}t + \frac{2}{5} \\ y = \frac{2}{5}s + \frac{1}{5}t + \frac{1}{5} \end{cases}$$

i odwrotne

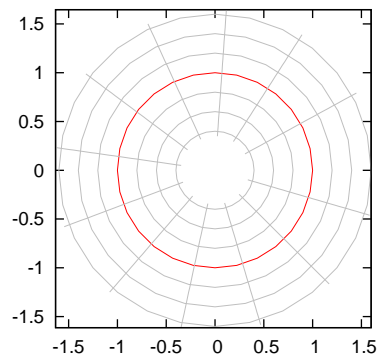
$$\begin{cases} s = 2y + x \\ t = y - 2x - 1 \end{cases}$$



Przykład 2. Drugi standardowy przykład to okrąg:

$$S = \{(x, y) : x^2 + y^2 = 1\}.$$

W tym przypadku najłatwiej użyć biegunowego układu współrzędnych:



Wzory są nam znane od dawna. Żeby zachować warunek „druga współrzędna równa zero” musimy zmienić kolejność współrzędnych i przesunąć wartości r :

$$\Phi_1(\varphi, r) = (x(\varphi, r), y(\varphi, r)), \quad \varphi \in]0, 2\pi[, \quad r \in]-1, 1[$$

$$\begin{cases} x = (r + 1) \cos \varphi \\ y = (r + 1) \sin \varphi \end{cases}$$

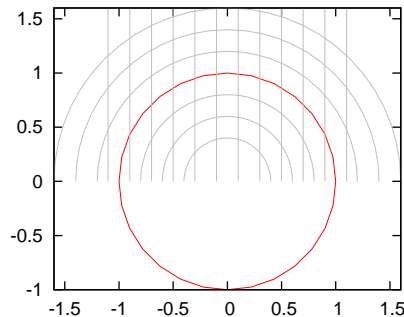
Tym razem odwrotnego odwzorowania nie napiszemy tak łatwo. Ponadto nie wystarczy jeden układ współrzędnych. Ten przedstawiony powyżej jest dobry dla każdego punktu oprócz punktu $(1, 0)$ (we współrzędnych kartezjańskich). W otoczeniu $(1, 0)$ możemy wziąć odwzorowanie zadane tymi samymi wzorami, ale określone na innej dziedzinie:

$$\Phi_2(\varphi, r) = (x(\varphi, r), y(\varphi, r)), \quad \varphi \in]-\pi, \pi[, \quad r \in]-1, 1[.$$

Dla okręgu także można wybierać inne układy współrzędnych. Na przykład taki:

$$\begin{cases} \xi = x \\ \rho = \sqrt{x^2 + y^2} - 1 \end{cases}$$

określony w otoczeniu części okręgu położonego w górnej półpłaszczyźnie:

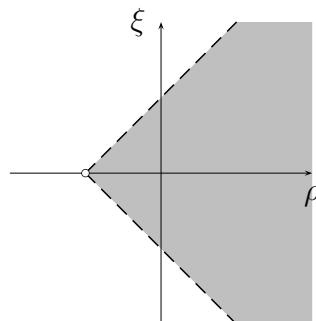


Odwzorowanie odwrotne

$$\begin{cases} x = \xi \\ y = \sqrt{(\rho + 1)^2 - \xi^2} \end{cases}$$

określone jest w obszarze

$$\mathcal{V} = \{(\xi, \rho) : \rho > -1, -\rho - 1 < \xi < \rho + 1\}$$



Do opisanego całego okręgu potrzebujemy czterech takich układów współrzędnych. ♣

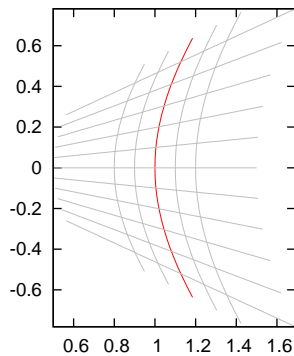
Przykład 3. Kolejnym przykładem niech będzie krzywa zadana równaniem

$$x^2 - y^2 = 1$$

dla $x > 0$. Krzywa ta to hiperbola. Odpowiedni układ współrzędnych znaleźć można korzystając z funkcji hiperbolicznych:

$$\begin{cases} x = (r + 1) \cosh(u) \\ y = (r + 1) \sinh(u) \end{cases}$$

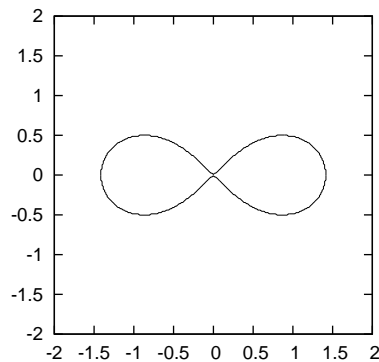
Współrzędna r należy do półprostej $] - 1, \infty[$, zaś współrzędna u może przyjmować dowolne wartości rzeczywiste. Obrazem $] - 1, \infty[\times \mathbb{R}$ w tym odwzorowaniu jest obszar pod prostą $y = x$, nad prostą $y = -x$ i dla $x > 0$.



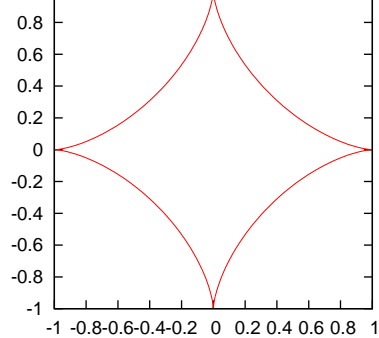
Przykład 4. Powierzchnią nie jest tzw. *lemniskata Bernoulliego* zadana równaniem

$$(x^2 + y^2)^2 = 2(x^2 - y^2)$$

Problemy są w otoczeniu punktu $(0, 0)$. Samoprzecięcia nie są dozwolone.



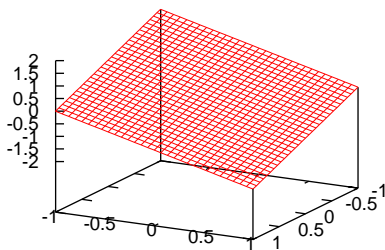
Przykład 5. Powierzchnią nie jest też. *hipocykloida* Problemy są w otoczeniu punktów $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$. Nie są dozwolone także dzióbki:



Przejdźmy teraz do przykładów w \mathbb{R}^3 :

Przykład 6. Powierzchnią dwuwymiarową (najprostszą) jest płaszczyzna. Weźmy płaszczyznę o równaniu

$$x + y + z = 0$$



Układ współrzędnych konstruujemy korzystając z algebry liniowej. Nasza płaszczyzna jest jednocześnie podprzestrzenią wektorową. Bazą tej podprzestrzeni są na przykład wektory

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

Do bazy w \mathbb{R}^3 uzupełnimy je dodając wektor prostopadły

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Układ współrzędnych (α, β, γ) związany jest z współrzędnymi kartezjańskimi równaniem:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

czyli

$$\begin{cases} x = \alpha + \gamma \\ y = \beta + \gamma \\ z = -\alpha - \beta + \gamma \end{cases}$$



Zastanówmy się teraz jak można opisywać powierzchnię. We wszystkich powyższych przykładach powierzchnia opisana była równaniem, czyli przedstawiona jako poziomica zerowa jakiejś funkcji. O tym jakie warunki powinna spełniać funkcja, żeby jej poziomica zerowa była powierzchnią mówi twierdzenie o maksymalnym rzędzie:

Twierdzenie 1 (O maksymalnym rzędzie). *Niech $F : \mathbb{R}^n \supset \mathcal{O} \longrightarrow \mathbb{R}^m$, $m < n$ będzie odwzorowaniem różniczkowalnym. Niech także $S = F^{-1}(0, \dots, 0)$ będzie zawarte w \mathcal{O} . Wówczas jeśli w każdym punkcie x pochodna $F'(x)$ ma maksymalny rząd (czyli równy m) to S jest powierzchnią w \mathbb{R}^n wymiaru $k = n - m$.*

Powierzchnię można także zadawać poprzez parametryzację, tzn obraz odwzorowania

$$\Psi : \mathbb{R}^k \supset \mathcal{V} \longrightarrow \mathbb{R}^n.$$

Okrąg o promieniu 1 można sparаметryzować następująco

$$\Psi :]0, 2\pi[\ni \varphi \longmapsto (\cos \varphi, \sin \varphi) \in \mathbb{R}^2.$$

Obrazem tej parametryzacji jest okrąg bez jednego punktu. Zazwyczaj nie da się sparаметryzować całej powierzchni za pomocą jednego odwzorowania. Parametryzacja musi także spełniać pewne warunki. Musi to być różniczkowalna iniekcja, której pochodna w każdym punkcie ma maksymalny rząd. Oto stosowne twierdzenie:

Twierdzenie 2. *Niech $\Psi : \mathbb{R}^k \supset \mathcal{V} \longrightarrow \mathbb{R}^n$, $k < n$ będzie odwzorowaniem różniczkowalnym. Wówczas jeśli $p \in \mathcal{V}$ i $\Psi'(p)$ jest rzędu k to istnieje otoczenie \mathcal{O} punktu p takie, że $\Psi(\mathcal{O})$ jest powierzchnią k -wymiarową w \mathbb{R}^n .*

Twierdzenie dotyczące parametryzacji jest jedynie lokalne, dlatego oprócz rzędu odwzorowania należy sprawdzać także iniekctywność.

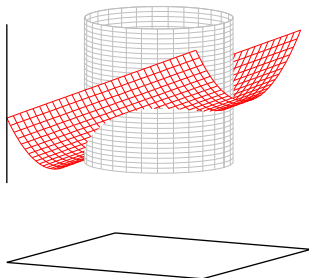
W dalszym ciągu zajmować się będziemy raczej powierzchniami zadawanymi jako poziomice zerowe. Interesować nas będą obcięcia funkcji do powierzchni. Niech S będzie okręgiem o promieniu 1 w \mathbb{R}^2 . Niech także

$$f(x, y) = x + y^2$$

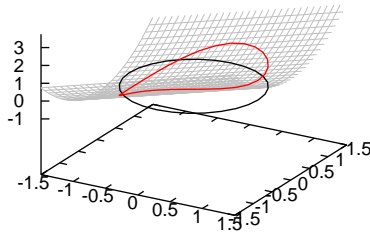
będzie funkcją na \mathbb{R}^2 . Funkcja ta jest różniczkowalna na całym \mathbb{R}^2 . Jej pochodna w punkcie (x, y) jest macierzą jednowierszową

$$f'(x, y) = [1 \quad 2y].$$

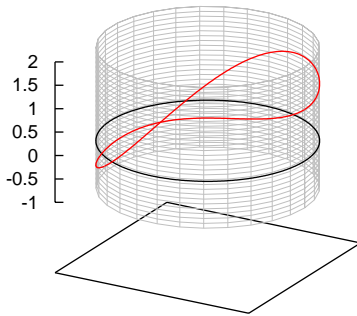
Pierwszy wyraz tej macierzy jest niezależny od punktu i równy 1, zatem funkcja ta nie ma punktów krytycznych w \mathbb{R}^2 . Rozpatrzmy teraz obcięcie tej funkcji do okręgu S . Żeby zrozumieć o co chodzi przyjrzyjmy się rysunkom:



Na czerwono zaznaczono wykres funkcji f . Szary walec narysowany jest nad okręgiem S . Przecięcie tych dwóch powierzchni ilustruje obcięcie f do S . Zamiast walca rysujemy teraz (na czerwono) krzywą będącą przecięciem walca i wykresu funkcji oraz (na czarno) okrąg S na płaszczyźnie $z = 0$:



Być może lepiej będzie widać, jeśli zamiast wykresu f zaznaczymy na rysunku walec:



Funkcja f obcięta do S przyjmuje wartości jedynie z pewnego odcinka. Z wykresów widać, że ma ona punkty krytyczne. Naszym zadaniem będzie znalezienie tych punktów. Jeden sposób poszukiwania punktów krytycznych jest łatwo wymyśleć. Jeśli sparametryzujemy okrąg S (na przykład kątem biegunowym) i złożymy funkcję f z parametryzacją otrzymamy funkcję jednej zmiennej rzeczywistej, określoną na odcinku, którą można badać zwykłymi metodami. Pamiętać musimy jednak, że parametryzacja nie obejmuje całego okręgu. W tym szczególnym przypadku można sobie łatwo poradzić z tym problemem. W ogólności jednak znalezienie parametryzacji dla powierzchni bywa kłopotliwe. Poszukamy zatem innej metody znalezienia punktów krytycznych: takiej, która wykorzystywałaby raczej funkcję, za pomocą której powierzchnia zadawana jest jako poziomica zerowa. W przypadku okręgu będzie to funkcja

$$G(x, y) = x^2 + y^2 + 1.$$

Punkt krytyczny jest, jak zwykle, punktem w którym pochodna znika, jednak teraz interesuje nas jedynie wartość pochodnej na wektorach (przyrostach) stycznych do powierzchni. Na wektorach w innych kierunkach pochodna może sobie być jakakolwiek. Skoro powierzchnia S jest poziomica zerową funkcji G , to oczywiście pochodna funkcji G znika na wektorach stycznych do powierzchni, ponieważ wzdłuż powierzchni funkcja się nie zmienia: pozostaje równa zero. Wektory styczne do powierzchni można wręcz zdefiniować w taki sposób: Wektor h zaczepiony w $(x_0, y_0) \in S$ jest *styczny* do S jeśli $G'(x_0, y_0)(h) = 0$. Jeśli więc pochodna funkcji f byłaby w

punkcie (x, y) proporcjonalna do $G'(x, y)$, to wiadomo byłoby, że pochodna ta znika na wektorach stycznych w tym punkcie. **Punkt $(x, y) \in S$ jest więc punktem krytycznym funkcji f na powierzchni S jeśli istnieje liczba λ taka, że**

$$f'(x, y) = \lambda G'(x, y).$$

Liczba λ nazywana jest *mnożnikiem Lagrange'a*. W powyższym rozumowaniu myśleliśmy cały czas o konkretnej funkcji f obciętej do konkretnej powierzchni S . Wnioski do których doszliśmy mają jednak charakter ogólny. Sformułujmy je w postaci twierdzenia:

Twierdzenie 3 (Lagrange). *Niech $f : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją różniczkowalną a $S = G^{-1}(0)$ powierzchnią wymiaru $n - 1$. Jeśli $p \in S$ jest punktem krytycznym funkcji f na powierzchni S to istnieje taka liczba rzeczywista λ , że*

$$f'(p) = \lambda G'(p).$$

Znajdźmy punkty krytyczne w naszym przykładzie. Mamy

$$f(x, y) = x + y^2, \quad G(x, y) = x^2 + y^2 - 1,$$

zatem

$$f'(x, y) = [1 \quad 2y], \quad G'(x, y) = [2x \quad 2y].$$

Punkt (x, y) jest więc punktem krytycznym f na S jeśli

$$[1 \quad 2y] = \lambda [2x \quad 2y], \quad G(x, y) = 0.$$

W postaci układu równań warunki te zapisujemy jako:

$$\begin{cases} 1 = 2\lambda x \\ 2y = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

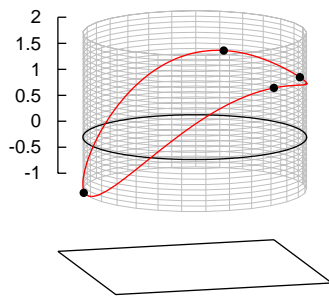
Drugie równanie oznacza, że

$$y(1 - \lambda) = 0, \quad \text{czyli} \quad y = 0 \quad \text{lub} \quad \lambda = 1.$$

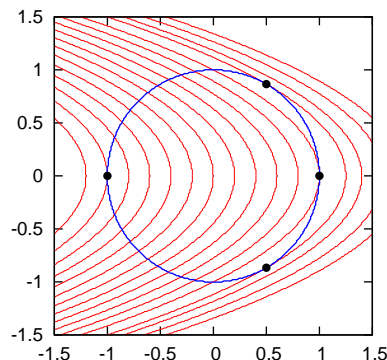
Gdy $y = 0$ współrzędna x musi być równa 1 lub -1 . Wtedy otrzymujemy λ odpowiednio $+\frac{1}{2}$, $-\frac{1}{2}$. Gdy zaś $\lambda = 1$ to $x = \frac{1}{2}$ i współrzędna y może przyjąć wartość $+\frac{\sqrt{3}}{2}$ lub $-\frac{\sqrt{3}}{2}$. Ostatecznie mamy cztery punkty krytyczne:

$$\begin{aligned} (1, 0), \quad \lambda &= \frac{1}{2}, \\ (-1, 0), \quad \lambda &= -\frac{1}{2}, \\ \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \lambda &= 1, \\ \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right), \quad \lambda &= 1. \end{aligned}$$

Zaznaczmy je na rysunku:



Drugi rysunek pokazuje, że punkty krytyczne obcięcia f do powierzchni S są tymi punktami, w których poziomice funkcji f są styczne do powierzchni S . Poziomice narysowane są na czerwono, a powierzchnia na niebiesko:



Z rysunków wynika, że punkty $(-1, 0)$ i $(0, 1)$ to minima a punkty $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ i $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$ to maksima. W ogólności musimy jednak znaleźć metodę badania rodzaju punktów krytycznych w przypadku gdy badamy funkcję obciętą do powierzchni.

Rachunek prowadzący do wzoru na pochodną funkcji obciętej do powierzchni jest dość skomplikowany. Zamieszczam go poniżej w dowodzie odpowiedniego twierdzenia. Dla celów praktycznych należy niestety zapamiętać wzór. Jeśli $p \in S$ jest punktem krytycznym funkcji f na S oraz h jest wektorem stycznym do S w punkcie p to forma kwadratowa reprezentująca drugą pochodną obcięcia funkcji f do S ma na h wartość

$$(f|_S)''(p)(h) = f''(p)(h) - \lambda G'''(p)(h),$$

gdzie λ jest wartością mnożnika Lagrange'a odpowiadającą punktowi krytycznemu p . Zbadamy określoność tej formy dla jednego z punktów krytycznych w naszym przykładzie. Wybierzmy punkt $p = (1, 0)$ z $\lambda = \frac{1}{2}$. Wyznaczamy macierz drugiej pochodnej funkcji f :

$$f''(p) = \begin{bmatrix} f_{xx}(p) & f_{xy}(p) \\ f_{xy}(p) & f_{yy}(p) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

Wyznaczamy także drugą pochodną G w punkcie p :

$$G''(p) = \begin{bmatrix} G_{xx}(p) & G_{xy}(p) \\ G_{xy}(p) & G_{yy}(p) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Macierz, którą mamy badać ma postać (bierzemy $\lambda = \frac{1}{2}$):

$$f''(p) - \lambda G''(p) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Określoność tej macierzy badamy jedynie na wektorach stycznych do powierzchni w punkcie $p = (1, 0)$. Pochodna G w $p = (1, 0)$ ma postać

$$G'(1, 0) = [2x \ 2y]_{|(1,0)} = [2 \ 0]$$

Wektory styczne do S w punkcie $(1, 0)$ spełniają warunek

$$G'(1, 0)(h) = 0, \quad [2 \ 0] \begin{bmatrix} h_x \\ h_y \end{bmatrix} = 2h_x = 0, \quad h_x = 0$$

Przestrzeń wektorów stycznych do okręgu S w punkcie $(1, 0)$ rozpięta jest przez wektor

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Obliczmy wartość formy $f''(p) - \lambda G''(p)$ na dowolnym wektorze proporcjonalnym do bazowego:

$$[0 \ \alpha] \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \alpha \end{bmatrix} = \alpha^2$$

Wartość drugiej pochodnej obciążenia na dowolnym wektorze stycznym do powierzchni jest dodatnia. Mamy więc do czynienia z minimum.

Zapiszmy jeszcze ogólne twierdzenie dotyczące wartości drugiej pochodnej obciążenia wraz z dowodem. Znajomość dowodu nie jest obowiązkowa, choć oczywiście wszystkich zachęcam do przeanalizowania rachunku.

Twierdzenie 4. *Przy założeniach twierdzenia Lagrange'a o punkcie krytycznym wartość drugiej pochodnej obciążenia f do powierzchni S w punkcie krytycznym na wektorach stycznych do S w tym punkcie wyraża się wzorem*

$$(f|_S)''(p)(h) = f''(p)(h) - \lambda G''(p)(h).$$

Dowód: (materiał nieobowiązkowy) Oznaczmy przez p punkt krytyczny obciążenia $f|_S$. Niech także $\Phi : \mathbb{R}^{n-1} \supset \mathcal{V} \rightarrow \mathbb{R}^n$ oznacza parametryzację otoczenia punktu p w S . Załóżmy, że parametryzacja została tak wybrana, że $\Phi(0) = p$. Pisząc 0 mam na myśli punkt zerowy w \mathbb{R}^{n-1} , czyli $(0, 0, \dots, 0)$. Jeśli p jest punktem krytycznym $f|_S$ to 0 jest punktem krytycznym $f \circ \Phi$. Obliczmy drugą pochodną złożenia $f \circ \Phi$ w otoczeniu punktu 0. Element \mathcal{V} z tego otoczenia oznaczmy y . Dla wektora przyrostu v zaczepionego w y mamy

$$f \circ \Phi'(y)(v) = f'(\Phi(y))(\Phi'(y)(v))$$

Drugą pochodną w 0 znajdziemy różniczkując powyższy wzór (traktując prawą stronę jako funkcję punktu y). Drugi przyrost oznaczmy u , różniczkujemy w $y = 0$:

$$(1) \quad (f \circ \Phi)''(0)(v, u) = f''(\Phi(0))(\Phi'(0)(v), \Phi'(0)(u)) + f'(\Phi(0))(\Phi''(0)(v, u))$$

Jeśli obrazy przyrostów v i u w \mathbb{R}^{n-1} oznaczmy odpowiednio h i k (tzn. h i k są wektorami stycznymi do S w punkcie p) to pierwszy składnik powyższego wzoru ma postać

$$(2) \quad f''(\Phi(0))(\Phi'(0)(v), \Phi'(0)(u)) = f''(p)(h, k).$$

Drugi składnik przekształcimy korzystając z tego, że punkt p jest punktem krytycznym. Funkcja $G \circ \Phi : \mathcal{V} \rightarrow \mathbb{R}$ jest funkcją stałą równą zero, gdyż G definiuje S . Jasne jest więc, że $(G \circ \Phi)''(0) = 0$. Drugą pochodną złożenia $G \circ \Phi$ obliczyć można według tych samych zasad co dla f , tyle, że teraz wynik musi być zero:

$$(G \circ \Phi)''(0)(v, u) = G''(\Phi(0))(\Phi'(0)(v), \Phi'(0)(u)) + G'(\Phi(0))(\Phi''(0)(v, u)) = 0.$$

Oznacza to, że

$$G''(\Phi(0))(\Phi'(0)(v), \Phi'(0)(u)) = -G'(\Phi(0))(\Phi''(0)(v, u)).$$

Oznaczając, jak poprzednio, $\Phi'(0)(v) = h$, $\Phi'(0)(u) = k$ i $\Phi(0) = p$ dostajemy

$$G''(p)(h, k) = -G'(p)(\Phi''(0)(v, u)).$$

Mnożąc obie strony powyższego równania przez λ i pamiętając, że w punkcie krytycznym p zachodzi

$$f'(p) = \lambda G'(p)$$

otrzymujemy

$$\lambda G''(p)(h, k) = -f'(p)(\Phi''(0)(v, u)),$$

czyli

$$(3) \quad f'(p)(\Phi''(0)(v, u)) = -\lambda G''(p)(h, k)$$

Wstawiając (3) i (2) do (1) otrzymujemy

$$\begin{aligned} (f \circ \Phi)''(0)(v, u) &= f''(\Phi(0))(\Phi'(0)(v), \Phi'(0)(u)) + f'(\Phi(0))(\Phi''(0)(v, u)) = \\ &= f''(p)(h, k) - \lambda G''(p)(h, k). \end{aligned}$$

Druga pochodna złożenia $f \circ \Phi$ w punkcie p jest drugą pochodną $f|_S$ w tym punkcie. Wstawiając dwa razy ten sam przyrost otrzymujemy formę kwadratową o którą chodziło w twierdzeniu. \square .

Rachunek przeprowadzony w dowodzie powyższego twierdzenia jest w ogólności skomplikowany. Słuchaczy, którzy nie chcą poprzestawać na zapamiętywaniu wzorów zachęcam do wykonania go w najprostszym przypadku funkcji określonej na \mathbb{R}^2 obciętej do jednowymiarowej powierzchni. Możemy myśleć tutaj o naszej funkcji f , powierzchni będącej okręgiem. Jako parametryzacji możemy użyć parametryzacji okręgu kątem biegunowym. Policzmy jak w dowodzie drugą pochodną złożenia $f \circ \Phi$:

$$f \circ \Phi(\varphi) = f(x(\varphi), y(\varphi))$$

W poniższym rachunku opuszczamy argumenty funkcji pamiętając jednak, że f zależy od φ poprzez x i y :

$$(f \circ \Phi)' = (f_x x_\varphi + f_y y_\varphi)' = (f_{xx} x_\varphi + f_{xy} y_\varphi) x_\varphi + f_x x_{\varphi\varphi} + (f_{yx} x_\varphi + f_{yy} y_\varphi) y_\varphi + f_y y_{\varphi\varphi}.$$

Pogrupujmy inaczej wyrazy:

$$(f \circ \Phi)' = (f_x x_\varphi + f_y y_\varphi)' = (f_{xx} x_\varphi + f_{xy} y_\varphi) x_\varphi + (f_{yx} x_\varphi + f_{yy} y_\varphi) y_\varphi + f_x x_{\varphi\varphi} + f_y y_{\varphi\varphi}.$$

Część czerwona jest to druga pochodna funkcji f działająca na wektor styczny do powierzchni otrzymany z przyrostu $\delta h = 1$:

$$\Phi'(\varphi) \delta h = \begin{bmatrix} x_\varphi \\ y_\varphi \end{bmatrix} \delta h = \begin{bmatrix} x_\varphi \delta h \\ y_\varphi \delta h \end{bmatrix}.$$

Inaczej zapiszemy więc część czerwoną jako

$$\begin{bmatrix} h_x & h_y \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} h_x \\ h_y \end{bmatrix}$$

pamiętając, że wektor przyrostu musi być styczny do powierzchni. Część niebieską zastąpimy wyrażeniem zawierającym G . Rachunek taki sam jak dla f moglibyśmy wykonać dla G , tyle, że druga pochodna wzdłuż powierzchni jest zawsze zero, bo funkcja G na powierzchni jest stała:

$$(G \circ \Phi)'' = (G_{xx}x_\varphi + G_{xy}y_\varphi)x_\varphi + (G_{yx}x_\varphi + G_{yy}y_\varphi)y_\varphi + G_{xx}x_{\varphi\varphi} + G_{yy}y_{\varphi\varphi} = 0$$

Przenosząc wyrażenia zawierające pierwszą pochodną G na drugą stronę dostajemy

$$(G_{xx}x_\varphi + G_{xy}y_\varphi)x_\varphi + (G_{yx}x_\varphi + G_{yy}y_\varphi)y_\varphi = -G_{xx}x_{\varphi\varphi} - G_{yy}y_{\varphi\varphi}$$

Wiadomo, że w punkcie krytycznym $f_x = \lambda G_x$ i $f_y = \lambda G_y$, dlatego po pomnożeniu obu stron powyższego równania przez λ możemy dokonać podstawienia:

$$\lambda[(G_{xx}x_\varphi + G_{xy}y_\varphi)x_\varphi + (G_{yx}x_\varphi + G_{yy}y_\varphi)y_\varphi] = -\lambda G_{xx}x_{\varphi\varphi} - \lambda G_{yy}y_{\varphi\varphi} = -f_{xx}x_{\varphi\varphi} - f_{yy}y_{\varphi\varphi}.$$

Otrzymaliśmy wyrażenie na część niebieską

$$f_{xx}x_{\varphi\varphi} + f_{yy}y_{\varphi\varphi} = -\lambda[(G_{xx}x_\varphi + G_{xy}y_\varphi)x_\varphi + (G_{yx}x_\varphi + G_{yy}y_\varphi)y_\varphi]$$

Podobnie jak dla funkcji f prawą stronę można zapisać jako działanie drugiej pochodnej G na wektor styczny do powierzchni:

$$\lambda \begin{bmatrix} h_x & h_y \end{bmatrix} \begin{bmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{bmatrix} \begin{bmatrix} h_x \\ h_y \end{bmatrix}.$$

Ostatecznie więc druga pochodna f obciętego do S ma w punkcie krytycznym w działaniu na wektor styczny h postać:

$$\begin{bmatrix} h_x & h_y \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} h_x \\ h_y \end{bmatrix} - \lambda \begin{bmatrix} h_x & h_y \end{bmatrix} \begin{bmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{bmatrix} \begin{bmatrix} h_x \\ h_y \end{bmatrix}.$$

Sama więc druga pochodna w punkcie krytycznym jest różnicą macierzy:

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} - \lambda \begin{bmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{bmatrix}.$$

Drugi człon w wyrażeniu na drugą pochodną pochodzi od różniczkowania funkcji złożonej z parametryzacją. Jest to wzór prawdziwy **jedynie w punkcie krytycznym**.

Potrzebujemy jeszcze dwóch przykładów: ekstremum funkcji określonej na \mathbb{R}^3 obciętej do dwuwymiarowej powierzchni i ekstremum funkcji na \mathbb{R}^3 obciętej do jednowymiarowej powierzchni, czyli krzywej.

Przykład 7. Znaleźć punkty krytyczne funkcji

$$f(x, y, z) = xyz$$

na sferze o promieniu $\sqrt{3}$ i środku w punkcie $(0, 0, 0)$. Zbadać dwa wybrane punkty krytyczne.

Sfera o promieniu $\sqrt{3}$ i środku w punkcie $(0, 0, 0)$ ma równanie

$$x^2 + y^2 + z^2 = 3,$$

weźmy więc $G(x, y, z) = x^2 + y^2 + z^2 - 3$. Zapisujemy warunek Lagranża:

$$f'(x, y, z) = \lambda G'(x, y, z), \quad [yz \ xz \ xy] = \lambda[2x \ 2y \ 2z].$$

Warunek na pochodne cząstkowe wraz z równaniem sfery dają układ równań:

$$\begin{cases} yz = 2\lambda x \\ xz = 2\lambda y \\ xy = 2\lambda z \\ x^2 + y^2 + z^2 = 3 \end{cases}$$

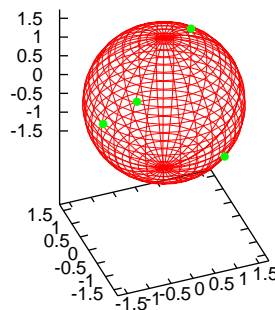
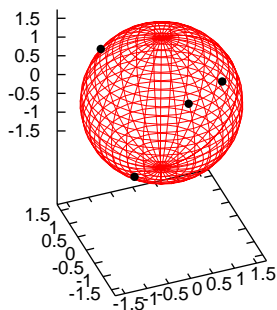
Powyższy układ równań ma całe mnóstwo rozwiązań. Jeśli założymy że żadna ze zmiennych nie jest równa 0 otrzymamy następujące rozwiązania:

$$\begin{aligned} (1, 1, 1), \lambda &= \frac{1}{2} \\ (-1, -1, 1), \lambda &= \frac{1}{2} \\ (-1, 1, -1), \lambda &= \frac{1}{2} \\ (1, -1, -1), \lambda &= \frac{1}{2} \\ (-1, 1, 1), \lambda &= -\frac{1}{2} \\ (1, -1, 1), \lambda &= -\frac{1}{2} \\ (1, 1, -1), \lambda &= -\frac{1}{2} \\ (-1, -1, -1), \lambda &= -\frac{1}{2} \end{aligned}$$

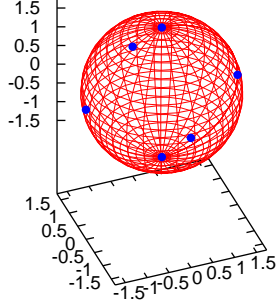
Założmy teraz, że $x = 0$. Wtedy pierwsze równanie wskazuje, że także y lub z jest 0. Gdy wszystkie trzy zmienne mają wartość 0 nie jest spełnione ostatnie równanie. Założmy więc, że tylko jedna z pozostałych zmiennych, np y jest także 0. Wtedy trzecie równanie wskazuje, że $\lambda = 0$. Jedynym warunkiem na zmienną z jest teraz ostatnie równanie, tzn $z = \pm\sqrt{3}$. Podobny wynik otrzymamy założywszy, że $y = 0$ lub $z = 0$. W ten sposób otrzymujemy dodatkowe sześć punktów krytycznych z wartością mnożnika równą 0:

$$(0, 0, \pm\sqrt{3}), \quad (0, \pm\sqrt{3}, 0), (\pm\sqrt{3}, 0, 0)$$

Na powierzchni zaznaczamy punkty krytyczne odpowiadające wartościom funkcji 1 (na zielono) i -1 (na czarno)



i punkty krytyczne odpowiadające wartości funkcji 0 (na niebiesko):



Punkty zaznaczone na niebiesko nie są ekstremalne, ponieważ w otoczeniu każdego z nich są punkty w których funkcja przyjmuje wartości dodatnie i ujemne.

Do badania wybieramy punkt $(1, 1, 1)$ dla $\lambda = \frac{1}{2}$. Forma kwadratowa, która określa rodzaj ekstremum to

$$f''(1, 1, 1) - \frac{1}{2}G''(1, 1, 1)$$

czyli

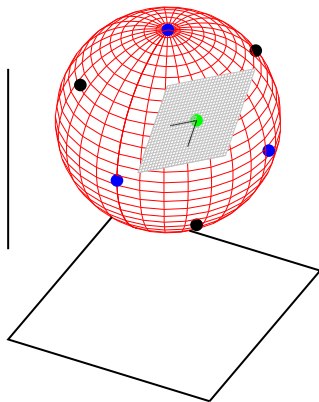
$$\begin{bmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{bmatrix}_{|(1,1,1)} - \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}_{|(1,1,1)} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

Szukamy wektorów stycznych do sfery w punkcie $(1, 1, 1)$. Są to wektory spełniające warunek

$$G'(1, 1, 1)h = 0, \quad \begin{bmatrix} 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} = 0.$$

Przestrzeń takich wektorów jest dwuwymiarowa, a przykładowa baza to:

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$



Obliczmy zatem formę kwadratową na wektorze postaci

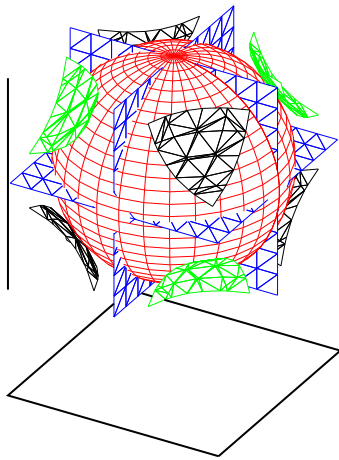
$$\alpha \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \alpha + \beta \\ -\alpha \\ -\beta \end{bmatrix}.$$

$$[(\alpha + \beta) \quad -\alpha \quad -\beta] \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \alpha + \beta \\ -\alpha \\ -\beta \end{bmatrix} = -2(\alpha + \beta)^2 - 2\alpha^2 - 2\beta^2$$

Porządkujemy wzór korzystając z metody Lagranża:

$$-2(\alpha + \beta)^2 - 2\alpha^2 - 2\beta^2 = -4(\alpha^2 + \beta^2 + \alpha\beta) = -4(\alpha + \frac{1}{2}\beta)^2 - 3\beta^2$$

Forma jest ujemnie określona, zatem punkt $(1, 1, 1)$ jest maksimum. Na rysunku zaznaczone są poziomice odpowiadające wartościom krytycznym funkcji. Na czarno minima, na zielono maksima a na niebiesko wartość 0, która nie jest ekstremalna. Poziomica $f(x, y, z) = 0$ nie jest powierzchnią w otoczeniu punktów krytycznych.



Przykład 8. Znaleźć punkty krytyczne funkcji

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

na przecięciu płaszczyzny $\{(x, y, z) : x + y + z = 0\}$ i walca $\{(x, y, z) : x^2 + y^2 = 2\}$. Zbadać jeden z nich.

Powierzchnia S zadana jest przez równania

$$G_1(x, y, z) = x^2 + y^2 - 2 = 0, \quad G_2(x, y, z) = x + y + z = 0.$$

Przestrzeń wektorów stycznych do powierzchni w punkcie p tej powierzchni składa się więc z wektorów spełniających także dwa równania: wektor h jest styczny jeśli

$$G'_1(p)h = 0 \quad \text{ i } \quad G'_2(p)h = 0.$$

Pochodna $f'(p)$ znika na tych wektorach jeśli jest kombinacją liniową pochodnych G_1 i G_2 . Tym razem potrzebować będziemy dwóch mnożników Lagranża λ i μ :

Obliczamy odpowiednie pochodne w $p = (x, y, z)$:

$$f'(x, y, z) = [2x \ 2y \ 2z] \quad G'_1(x, y, z) = [2x \ 2y \ 0] \quad G'_2(x, y, z) = [1 \ 1 \ 1]$$

i piszemy warunek Lagranża:

$$f'(x, y, z) = \lambda G'_1(x, y, z) + \mu G'_2(x, y, z), \quad [2x \ 2y \ 2z] = \lambda[2x \ 2y \ 0] + \mu[1 \ 1 \ 1]$$

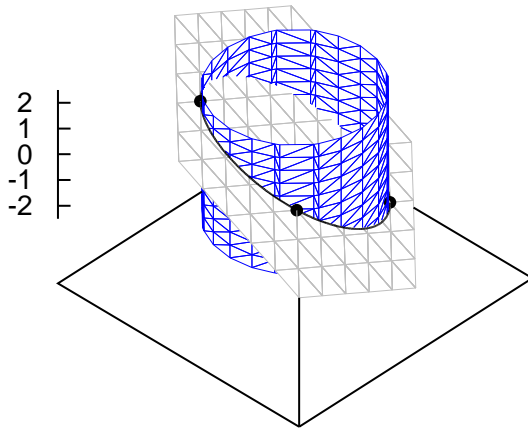
Wraz z równaniami powierzchni otrzymujemy układ równań:

$$\begin{cases} 2x = 2\lambda x - \mu \\ 2y = 2\lambda y - \mu \\ 2z = \mu \\ x^2 + y^2 = 2 \\ x + y + z = 0 \end{cases}$$

Otrzymujemy cztery punkty krytyczne

$$\begin{array}{lll} (1, -1, 0) & \lambda = 1 & \mu = 0 \\ (-1, 1, 0) & \lambda = 1 & \mu = 0 \\ (-1, -1, 2) & \lambda = 3 & \mu = 4 \\ (1, 1, -2) & \lambda = 3 & \mu = -4 \end{array}$$

Punkty te zaznaczamy na rysunku.



Wybieramy punkt $(1, 1, -2)$ do badania rodzaju punktu krytycznego. Bez dowodu podajemy fakt, iż jeśli powierzchnia zadana jest przez k równań postaci $G_i(x_1, \dots, x_n) = 0$ to w punkcie krytycznym p rodzaj ekstremum jest określony przez sygnaturę formy kwadratowej

$$f''(p) - \lambda_1 G''_1(p) - \dots - \lambda_k G''_k(p)$$

na przestrzeni stycznej do powierzchni w tym punkcie.

My mamy dwa równania, forma kwadratowa ma więc trzy składniki: druga pochodna f , druga pochodna G_1 i druga pochodna G_2 . Znajdujemy przestrzeń styczną do S w punkcie krytycznym:

$$G'_1(1, 1, -2) = [2 \ 2 \ 0] \quad G'_2(1, 1, -2) = [1 \ 1 \ 1]$$

Wektor styczny h ma być anihilowany przez obie powyższe macierze jednowierszowe. Prowadzi to do układu równań:

$$\begin{cases} 2h_x + 2h_y = 0 \\ h_x + h_y + h_z = 0 \end{cases}$$

Przestrzeń rozwiązań rozpięta jest przez wektor

$$h = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Badać musimy znak formy:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} - 3 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 4 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

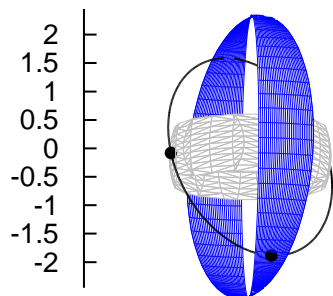
Ostatnia macierz jest zerowa, gdyż druga pochodna G_2 jest zerowa. Ostatecznie więc do badania mamy macierz

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

na przestrzeni wektorów proporcjonalnych do h :

$$\alpha[-1 \ 1 \ 0] \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \alpha^2[-1 \ 1 \ 0] \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = -2\alpha^2 < 0$$

Druga pochodna jest ujemna, mamy więc do czynienia z maksimum. Na rysunku zaznaczono fragmenty poziomic funkcji f stycznych w punktach krytycznych do powierzchni S :



Rachunek różniczkowy funkcji wielu zmiennych

Z uwagi na prostotę zapisu i łatwe interpretacje graficzne ograniczymy się jedynie do funkcji 2 lub 3 zmiennych. Naturalne uogólnienia wprowadzanych pojęć na funkcje k zmiennych zostawiamy czytelnikowi jako ćwiczenie.

Niech $R^2 = \{(x, y) : x \in R, y \in R\}$ oznacza zbiór par liczb rzeczywistych, a $R^3 = \{(x, y, z) : x \in R, y \in R, z \in R\}$ zbiór trójek (3-elementowych ciągów) takich liczb. Elementy zbiorów R^2 i R^3 możemy traktować jako

- punkty w 2 lub 3-wymiarowej przestrzeni z naturalnym układem współrzędnych
- wektory zaczepione w ustalonym punkcie będącym początkiem układu współrzędnych i końcach w zadanych punktach wodzącego
- wektory swobodne w 2 lub 3-wymiarowej przestrzeni, które można zaczepiać w dowolnym punkcie.

Traktując elementy (x_1, y_1) i (x_2, y_2) przestrzeni R^2 jako punkty – P_1 i P_2 możemy zdefiniować odległość (euklidesową) pomiędzy tymi punktami

$$\rho(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Przestrzeń (R^2, ρ) jest tzw. przestrzenią metryczną.

Metryka ρ spełnia następujące postulaty:

1. $\forall_{P_1, P_2 \in R^2} \rho(P_1, P_2) \geq 0$ i $\rho(P_1, P_2) = 0 \Leftrightarrow P_1 = P_2$
2. $\forall_{P_1, P_2 \in R^2} \rho(P_1, P_2) = \rho(P_2, P_1)$
3. $\forall_{P_1, P_2, P_3 \in R^2} \rho(P_1, P_3) \leq \rho(P_1, P_2) + \rho(P_2, P_3)$ (warunek trójkąta)

Granica ciągu o wartościach w R^k

Wiadomo, że ciąg o wartościach rzeczywistych to funkcja rzeczywista określona na zbiorze liczb rzeczywistych. Podobnie funkcję $P : N \ni n \rightarrow P(n) = P_n \in R^k$ nazwiemy ciągiem w przestrzeni R^k .

Przyjmując oznaczenia $P_n = (x_1^n, \dots, x_k^n)$ $P = (g_1, \dots, g_k)$ możemy zdefiniować granice ciągu w R^k

$$\boxed{\lim_{n \rightarrow \infty} P_n = P \Leftrightarrow \lim_{n \rightarrow \infty} \rho(P_n, P) = 0}$$

przy czym odległość $\rho(P, Q)$ pomiędzy punktami $P = (x_1, \dots, x_k)$ i $Q = (y_1, \dots, y_k)$ jest zadana

$$\text{wzorem } \rho(P, Q) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}.$$

Uwaga. Powyżej zdefiniowana zbieżność jest równoważna zbieżności „po współrzędnych”- uzasadnić.

Przykład. $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n, \frac{\sin \frac{1}{n}}{\frac{1}{n}}, \frac{2n}{n+1} = (e, 1, 2).$

Podobnie jak w rozważanym poprzednio przypadku funkcji $f: R \rightarrow R$ rzeczywistej zmiennej rzeczywistej możemy zdefiniować otoczenie i sąsiedztwo punktu w przestrzeni metrycznej

$$Ot(P_0, \delta) = K(P_0, \delta) = \{P \in D_f : \rho(P, P_0) < \delta\} ; S(P_0, \delta) = Ot(P_0, \delta) - \{P_0\}$$

Funkcja rzeczywista k zmiennych rzeczywistych

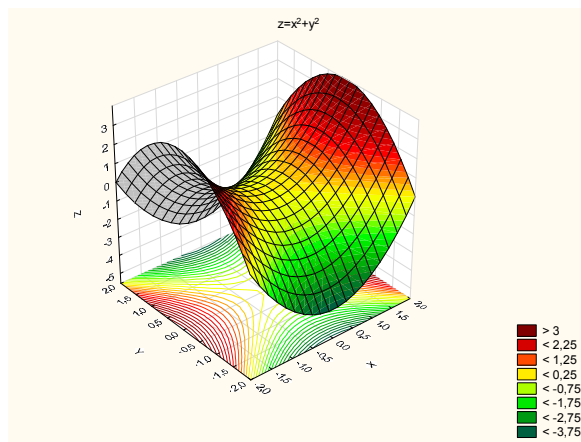
Utożsamiając punkt P z jego współrzędnymi $\mathbf{x} = (x_1, \dots, x_k)$ rozważać będziemy rzeczywistą funkcję punktu P jako funkcję wielu zmiennych (współrzędnych punktu)

$$f: R^k \supset D \ni \mathbf{x} \rightarrow f(\mathbf{x}) \in R$$

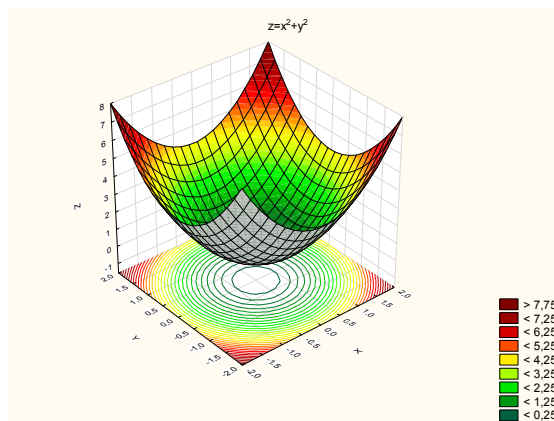
Przykład. $f(x, y) = \arcsin \frac{x}{a} + \arcsin \frac{y}{b}$ jest rzeczywistą funkcją dwóch zmiennych rzeczywistych określoną na prostokącie $[-a, a] \times [-b, b]$.

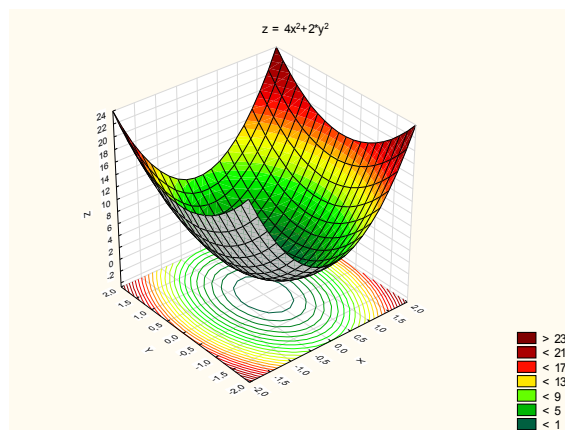
Przykładowe wykresy funkcji 2 zmiennych- przekroje- powierzchnie obrotowe np.:

- Paraboloida hiperboliczna $f(x, y) = x^2 - y^2$ (siodło)



- paraboloida obrotowa $f(x, y) = x^2 + y^2$





- paraboloida eliptyczna $f(x,y)=4x^2+9y^2$

Granica funkcji $f: R^k \supset D \ni P \rightarrow f(P) \in R$ w punkcie skupienia P_0 dziedziny D (punkt skupienia zbioru D nie musi należeć do zbioru D)

Heine $\lim_{P \rightarrow P_0} f(P) = g \Leftrightarrow \forall (P_n) \subset S(P_0, \delta) \lim_{n \rightarrow \infty} P_n = P_0 \Rightarrow \lim_{n \rightarrow \infty} f(P_n) = g$

Cauchy $\lim_{P_n \rightarrow P_0} f(P) = g \Leftrightarrow \forall \varepsilon \exists \delta \forall P \in D \quad 0 < \rho(P, P_0) \leq \delta \Rightarrow |f(P) - g| \leq \varepsilon$

Ciągłość funkcji $f: R^k \supset D \ni P \rightarrow f(P) \in R$ w punkcie $P_0 \in D$

Heine f jest ciągła w $P_0 \in D \Leftrightarrow \forall (P_n) \subset Ot(P_0, \delta) \lim_{n \rightarrow \infty} P_n = P_0 \Rightarrow \lim_{n \rightarrow \infty} f(P_n) = f(P_0)$

Cauchy f jest ciągła w $P_0 \in D$

$\Leftrightarrow \forall \varepsilon \exists \delta \forall P \in D \quad \rho(P, P_0) \leq \delta \Rightarrow |f(P) - f(P_0)| \leq \varepsilon$

Podobnie jak w przypadku funkcji jednej zmiennej funkcja $f: R^k \supset Ot(P_0, \delta) \ni P \rightarrow f(P) \in R$ jest ciągła w każdym punkcie izolowanym $P_0 \in D$, a w punkcie skupienia P_0 zbioru D jest ciągła jeżeli $\lim_{P \rightarrow P_0} f(P) = f(P_0)$.

Ciągłość punktowa i jednostajna funkcji na zbiorze D jest definiowana dokładnie tak jak poprzednio zastępując moduł różnicy dwóch liczb rzeczywistych przez odległość punktów przestrzeni metrycznej R^k .

Def. Funkcja $f: R^k \supset D \ni P \rightarrow f(P) \in R$ jest **jednostajnie ciągła** w D gdy

$\forall \varepsilon > 0 \exists \delta > 0 \forall P \in D \forall Q \in D \quad \rho(P, Q) \leq \delta \Rightarrow |f(P) - f(Q)| \leq \varepsilon$

Tw.

ciągłość jednostajna \Rightarrow ciągłość punktowa

POCHODNE CZĄSTKOWE I KIERUNKOWE

$$f: R^k \supset E \rightarrow R, \quad E - \text{otwarty}, \quad \mathbf{x} = (x_1, \dots, x_k) \in E$$

Def. Pochodną cząstkową funkcji f w punkcie \mathbf{x} względem i -tej zmiennej (czyli x_i) nazywamy skończoną (o ile istnieje) granicę

$$\lim_{h_i \rightarrow 0} \frac{f(x_1, \dots, x_i + h_i, \dots, x_k) - f(x_1, \dots, x_i, \dots, x_k)}{h_i} \stackrel{df}{=} \frac{\partial f}{\partial x_i}(\mathbf{x}) = f'_{x_i}(\mathbf{x})$$

Def. Pochodną kierunkową funkcji f w punkcie \mathbf{x} w kierunku wektora $\mathbf{k} \neq \mathbf{0}$ (zwykle przyjmuje się, że $\|\mathbf{k}\|=1$, czyli w kierunku wersora) nazywamy skończoną (o ile istnieje) granicę

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t\mathbf{k}) - f(\mathbf{x})}{t} = \mathbf{f}'_{\mathbf{k}}(\mathbf{x}).$$

Uwaga 1. $\varphi(t) = f(\mathbf{x} + t\mathbf{k})$ - funkcja jednej zmiennej

Uwaga 2. Pochodna cząstkowa jest szczególnym przypadkiem pochodnej kierunkowej w

$$\text{kierunku } i\text{-tego wersora bazy kanonicznej (wersora } i\text{-tej osi)} \quad \frac{\partial f}{\partial x_i}(\mathbf{x}) = f'_{e_i}(\mathbf{x})$$

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

- pochodne cząstkowe

$$\mathbf{x} = (x, y) \neq (0, 0) \quad \frac{\partial f}{\partial x}(x, y) = \frac{y^3(x^2 + y^6) - xy^3(2x)}{(x^2 + y^6)^2} = \frac{y^3(-x^2 + y^6)}{(x^2 + y^6)^2}$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{3y^2x(x^2 + y^6) - xy^3(6y^5)}{(x^2 + y^6)^2} = \frac{3xy^2(x^2 - y^6)}{(x^2 + y^6)^2}$$

$$(x, y) = (0, 0) \quad \frac{\partial f}{\partial x}(0, 0) = \lim_{\Delta x \rightarrow 0} \frac{f(0 + \Delta x, 0) - f(0, 0)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\frac{\Delta x \cdot 0}{\Delta x^2 + 0} - 0}{\Delta x} = 0 = \frac{\partial f}{\partial y}(0, 0)$$

- pochodna kierunkowa

$$\mathbf{x} = (x, y) \neq (0, 0) \quad \mathbf{k} = (k_x, k_y) \quad \|(k_x, k_y)\| = \sqrt{k_x^2 + k_y^2} = 1$$

$$\begin{aligned} f'_{(k_x, k_y)}(x, y) &= \lim_{t \rightarrow 0} \frac{f(\mathbf{x}^\circ + t\mathbf{k}) - f(\mathbf{x}^\circ)}{t} = \lim_{t \rightarrow 0} \frac{f(x + tk_x, y + tk_y) - f(x, y)}{t} = \\ &= \lim_{t \rightarrow 0} \frac{\frac{(x + tk_x)(y + tk_y)^3}{(x + tk_x)^2 + (y + tk_y)^6} - \frac{xy^3}{x^2 + y^6}}{t} = \frac{y^2(-x^2 + y^6)(yk_x - 3xk_y)}{(x^2 + y^6)^2} = \\ &= \frac{y^3(-x^2 + y^6)}{(x^2 + y^6)^2} k_x + \frac{3xy^2(x^2 - y^6)}{(x^2 + y^6)^2} k_y = \frac{\partial f}{\partial x}(x, y) k_x + \frac{\partial f}{\partial y}(x, y) k_y \end{aligned}$$

W punkcie $\mathbf{x} = (x, y) = (0, 0)$

$$f'_k(0,0) = \lim_{t \rightarrow 0} \frac{f((0,0) + t(k_x, k_y)) - f(0,0)}{t} = \lim_{t \rightarrow 0} \frac{f(tk_x, tk_y)}{t} = \lim_{t \rightarrow 0} \frac{tk_x (tk_y)^3}{(tk_x)^2 + (tk_y)^6} = 0$$

Funkcja f ma więc w każdym punkcie i w każdym kierunku pochodną.

Pytanie. Czy istnieje granica $\lim_{(x,y) \rightarrow (0,0)} f(x,y)$?

Powyższa granica nie istnieje, gdyż $(\frac{1}{n}, 0) \rightarrow (0,0)$ i $f(\frac{1}{n}, 0) = 0 \rightarrow 0$, a $(\frac{1}{n^3}, \frac{1}{n}) \rightarrow (0,0)$ i

$$f(\frac{1}{n^3}, \frac{1}{n}) = \frac{1}{2} \rightarrow \frac{1}{2}.$$

Wykazaliśmy więc, że **posiadanie pochodnej kierunkowej** w dowolnym kierunku (w szczególności posiadanie pochodnych cząstkowych) **nie zapewnia nawet ciągłości funkcji**.

Pochodna Frecheta (pochodna mocna)

Z rachunku różniczkowego funkcji jednej zmiennej wiemy, że dla funkcji różniczkowalnych prawdziwe jest twierdzenie o przedstawieniu przyrostu funkcji. Możliwość odpowiedniego przedstawienia przyrostu funkcji, może być wykorzystana jako alternatywny sposób zdefiniowania różniczkowalności. To podejście do różniczkowalności można łatwo uogólnić na funkcje wielu zmiennych.

Def. Funkcja $f: R^k \supset Ot(\mathbf{x}, \delta) \ni \mathbf{x} \rightarrow f(\mathbf{x}) \in R$ jest różniczkowalna w punkcie \mathbf{x} jeżeli istnieją stałe A_1, \dots, A_k takie, że dla dostatecznie małych przyrostów $\mathbf{h} = (h_1, \dots, h_k)$

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = A_1 h_1 + \dots + A_k h_k + r(\mathbf{x}, \mathbf{h}), \text{ przy czym } \lim_{\mathbf{h} \rightarrow 0} \frac{r(\mathbf{x}, \mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Wyrażenie $A_1 h_1 + \dots + A_k h_k$ można zapisać w postaci macierzowej $\mathbf{A} \mathbf{h}$, gdzie $\mathbf{A} = [A_1, \dots, A_k]$, $\mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_k \end{bmatrix}$.

Pochodną funkcji f w punkcie \mathbf{x} nazywamy odwzorowanie liniowe $\mathbf{L}: R^n \rightarrow R$ reprezentowane przez macierz \mathbf{A} , czyli warunek różniczkowalności można zapisać

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \mathbf{L} \mathbf{h} + r(\mathbf{x}, \mathbf{h}), \text{ przy czym } \lim_{\mathbf{h} \rightarrow 0} \frac{r(\mathbf{x}, \mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Fakty :

- jeżeli f jest różniczkowalna w punkcie \mathbf{x} , to jest ona ciągła w punkcie \mathbf{x}
- jeżeli f jest różniczkowalna w \mathbf{x} , to f ma w \mathbf{x} pochodne kierunkowe w dowolnym kierunku, więc ma także pochodne cząstkowe, przy czym $A_i = \frac{\partial f}{\partial x_i}(\mathbf{x})$
- jeżeli f ma w $Ot(\mathbf{x}, \delta)$ pochodne cząstkowe **ciągłe** w \mathbf{x} , to f jest różniczkowalna w \mathbf{x}
- jeżeli f jest różniczkowalna, to $f'_k(\mathbf{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}) k_i$ w zapisie macierzowym

$$f'_k(\mathbf{x}) = \mathbf{A} \mathbf{k}$$

Uzasadnienia powyższych faktów

Tw. Jeżeli f jest różniczkowalna w punkcie \mathbf{x} , to jest ona ciągła w punkcie \mathbf{x} .

Jest to natychmiastowa konsekwencja definicji różniczkowalności

Pochodne cząstkowe i kierunkowe a różniczkowalność.

f jest różniczkowalna w punkcie $\mathbf{x} \Rightarrow$

$$f(\mathbf{x} + t\mathbf{k}) - f(\mathbf{x}) = \mathbf{f}'(\mathbf{x}) \cdot t\mathbf{k} + r(\mathbf{x}, t\mathbf{k}) \quad /: t \quad r(\mathbf{x}, t\mathbf{k}) = o(\|t\mathbf{k}\|)$$

$$\frac{f(\mathbf{x} + t\mathbf{k}) - f(\mathbf{x})}{t} = \mathbf{f}'(\mathbf{x}) \cdot \mathbf{k} + \frac{r(\mathbf{x}, t\mathbf{k})}{t}$$

$$\left| \frac{r(\mathbf{x}, t\mathbf{k})}{t} \right| = \frac{|r(\mathbf{x}, t\mathbf{k})|}{\|t\mathbf{k}\|} \|\mathbf{k}\| \rightarrow 0, \text{ gdy } t \rightarrow 0 \Rightarrow f'_k(\mathbf{x}) = \mathbf{f}'(\mathbf{x}) \mathbf{k}$$

Niech \mathbf{e}_i - i-ty wektor bazy kanonicznej

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = f'_{\mathbf{e}_i}(\mathbf{x}) = [A_1, \dots, A_k] \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \leftarrow \{\text{i-te miejsce}\} = A_i$$

$$\mathbf{f}'(\mathbf{x}) = \mathbf{grad} f(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_k}(\mathbf{x}) \right] - \text{gradient funkcji } f \text{ w punkcie } \mathbf{x}$$

Ogólnie dla $\mathbf{f} : R^k \supset D \rightarrow R^m$

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} = \sum_{j=1}^m f_j(\mathbf{x}) \bar{\mathbf{e}}_j \quad \mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f_1}{\partial x_k}(\mathbf{x}) \\ \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f_m}{\partial x_k}(\mathbf{x}) \end{bmatrix}_{m \times k}$$

Pokazano więc fakt: Jeżeli funkcja $\mathbf{f} : R^k \supset D \rightarrow R^m$, $\mathbf{x} \in D$ - otwarty, jest różniczkowalna w punkcie $\mathbf{x} \in D$, to istnieją pochodne cząstkowe $\frac{\partial f_j}{\partial x_i}(\mathbf{x})$, $j=1, \dots, m$ $i=1, \dots, n$.

Twierdzenie odwrotne nie zachodzi, tzn. istnienie pochodnych cząstkowych nie gwarantuje różniczkowalności (a nawet nie gwarantuje ciągłości)

Tw. (Warunek wystarczający różniczkowalności) Jeżeli funkcja $f : R^k \supset D \rightarrow R$ (D -

otwarty, $\mathbf{x} \in D$) ma pochodne cząstkowe $\frac{\partial f}{\partial x_j} : \mathbf{x} \in D \rightarrow \frac{\partial f}{\partial x_j}(\mathbf{x})$ ($j=1, \dots, k$) w D , **ciągłe** w \mathbf{x} ,

to f jest różniczkowalna w \mathbf{x} .

Dow: dla $k=2$

$$\mathbf{x} = (x_1, x_2) \quad \mathbf{h} = (h_1, h_2) \quad f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2) =$$

$$f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) + f(x_1, x_2 + h_2) - f(x_1, x_2) =$$

z tw. Lagrange'a o wartości średniej

$$= \frac{\partial f}{\partial x_1}(\mathbf{c}_1)h_1 + \frac{\partial f}{\partial x_2}(\mathbf{c}_2)h_2 =$$

$$= \frac{\partial f}{\partial x_1}(\mathbf{x})h_1 + \frac{\partial f}{\partial x_2}(\mathbf{x})h_2 + \left(\frac{\partial f}{\partial x_1}(\mathbf{c}_1) - \frac{\partial f}{\partial x_1}(\mathbf{x}) \right)h_1 + \left(\frac{\partial f}{\partial x_2}(\mathbf{c}_2) - \frac{\partial f}{\partial x_2}(\mathbf{x}) \right)h_2$$

Wystarczy pokazać, że $r(\mathbf{x}, \mathbf{h}) = o(\|\mathbf{h}\|)$ czyli, że

$$\lim_{(h_1, h_2) \rightarrow (0,0)} \left| \frac{\partial f}{\partial x_1}(\mathbf{c}_1) - \frac{\partial f}{\partial x_1}(\mathbf{x}) \right| \frac{|h_1|}{\sqrt{h_1^2 + h_2^2}} + \left| \frac{\partial f}{\partial x_2}(\mathbf{c}_2) - \frac{\partial f}{\partial x_2}(\mathbf{x}) \right| \frac{|h_2|}{\sqrt{h_1^2 + h_2^2}} = 0.$$

Jest to prawda, gdyż

$$\lim_{(h_1, h_2) \rightarrow (0,0)} \left| \frac{\partial f}{\partial x_1}(\mathbf{c}_1) - \frac{\partial f}{\partial x_1}(\mathbf{x}) \right| = 0 \quad \text{ i } \quad \lim_{(h_1, h_2) \rightarrow (0,0)} \left| \frac{\partial f}{\partial x_2}(\mathbf{c}_2) - \frac{\partial f}{\partial x_2}(\mathbf{x}) \right| = 0 \quad \text{ z ciągłości pochodnych}$$

cząstkowych w \mathbf{x} a wyrażenia $\frac{|h_1|}{\sqrt{h_1^2 + h_2^2}}$ i $\frac{|h_2|}{\sqrt{h_1^2 + h_2^2}}$ są ograniczone,

Interpretacja geometryczna pochodnej

Niech $f : X \supset D \rightarrow Y$. Zbiór punktów $W = \{(\mathbf{x}, f(\mathbf{x})) \in X \times Y : \mathbf{x} \in D \subset X\}$ nazywamy **wykresem funkcji** $f : X \supset D \rightarrow Y$.

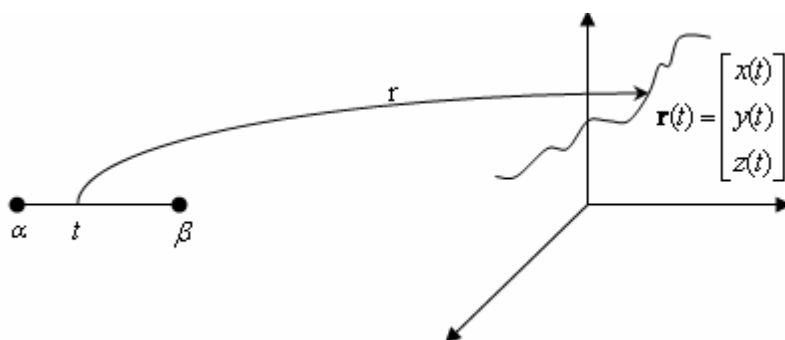
Fakt. Jeżeli funkcja $f : X \supset D \rightarrow Y$ jest różniczkowalna w punkcie $\mathbf{a} \in D$ to wektor $\mathbf{s} \in X \times Y$ jest styczny do wykresu W funkcji f w punkcie $(\mathbf{a}, f(\mathbf{a}))$ wtedy i tylko wtedy, gdy istnieje wektor $\mathbf{h} \in X$ taki, że

$$\mathbf{s} = (\mathbf{h}, f'(\mathbf{a}) \cdot \mathbf{h})$$

Tw. Jeżeli funkcja $\mathbf{f} : X \supset E \rightarrow Y$ jest różniczkowalna w punkcie $\mathbf{a} \in E$, to hiperpłaszczyzna w $X \times Y$ o równaniu $\mathbf{y} = \mathbf{f}(\mathbf{a}) + \mathbf{f}'(\mathbf{a})(\mathbf{x} - \mathbf{a})$ jest hiperpłaszczyzną styczną do wykresu funkcji \mathbf{f} w punkcie $(\mathbf{a}, \mathbf{f}(\mathbf{a}))$.

Przypadki szczególne

- $\mathbf{r} : [\alpha, \beta] \ni t \rightarrow \mathbf{r}(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} \in R^3$ jest funkcją wektorową, którą interpretujemy jako opis parametryczny krzywej w R^3 . Załóżmy, że funkcja ta jest różniczkowalna w punkcie $t_0 \in [\alpha, \beta]$



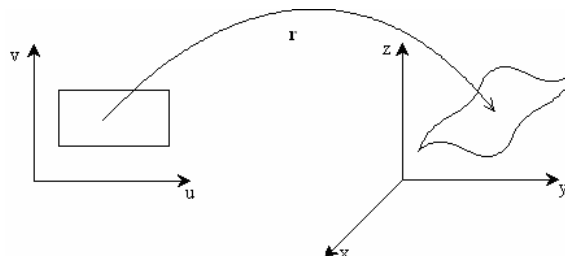
Pochodna $\mathbf{r}'(t_0)$ jest odwzorowaniem liniowym ciągłym z R w R^3 reprezentowanym przez

macierz $\mathbf{r}'(t_0) = \begin{bmatrix} x'(t_0) \\ y'(t_0) \\ z'(t_0) \end{bmatrix}$. Prosta o równaniu parametrycznym $\mathbf{r} = \mathbf{r}(t_0) + \mathbf{r}'(t_0)(t - t_0)$ jest

styczną do krzywej $\mathbf{r} = \mathbf{r}(t)$ w punkcie $\mathbf{r}_0 = \mathbf{r}(t_0)$.

- $\mathbf{r} : [a, b] \times [c, d] \ni (u, v) \rightarrow \mathbf{r}(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix} \in R^3$ jest funkcją wektorową, którą

interpretujemy jako opis parametryczny powierzchni w R^3 . Załóżmy, że funkcja ta jest różniczkowalna w punkcie (u_0, v_0)



Pochodna $\mathbf{r}'(u_0, v_0)$ jest odwzorowaniem liniowym ciągłym z R^2 w R^3 reprezentowanym

przez macierz $\mathbf{r}'(u_0, v_0) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix}_{(u_0, v_0)}$.

Płaszczyzna o równaniu parametrycznym $\mathbf{r} = \mathbf{r}(u_0, v_0) + \mathbf{r}'(u_0, v_0) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix}$, które można

także zapisać w postaci:

$$\mathbf{r} = \mathbf{r}(u_0, v_0) + \mathbf{r}_u(u - u_0) + \mathbf{r}_v(v - v_0),$$

gdzie \mathbf{r}_u i \mathbf{r}_v oznaczają kolumny macierzy reprezentującej pochodną $\mathbf{r}'(u_0, v_0)$ jest płaszczyzną styczną do powierzchni $\mathbf{r} = \mathbf{r}(u, v)$ w punkcie $\mathbf{r}_0 = \mathbf{r}(u_0, v_0)$.

Po rozpisaniu na współrzędne w postaci równanie płaszczyzny stycznej przybiera postać

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial u} \end{bmatrix} (u - u_0) + \begin{bmatrix} \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial v} \end{bmatrix} (v - v_0).$$

Po przemnożeniu obu stron równania $\mathbf{r} - \mathbf{r}_0 = \mathbf{r}_u(u - u_0) + \mathbf{r}_v(v - v_0)$ skalarnie przez wektor

$\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v$ ortogonalny do \mathbf{r}_u i \mathbf{r}_v rugujemy parametry u i v i otrzymujemy ogólną postać równania płaszczyzny stycznej

$$\mathbf{n} \circ (\mathbf{r} - \mathbf{r}_0) = 0, \text{ gdzie } \mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v.$$

W szczególności jeśli powierzchnia jest wykresem funkcji 2 zmiennych $z = f(x_1, x_2)$, to można ją zapisać parametrycznie przyjmując $u = x_1$ i $v = x_2$. Wówczas

$$\mathbf{r}_u = \begin{bmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial u} \end{bmatrix}, \quad \mathbf{r}_v = \begin{bmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial v} \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} -\frac{\partial f}{\partial u} \\ -\frac{\partial f}{\partial v} \\ 1 \end{bmatrix}. \quad \text{Stąd płaszczyzna o równaniu}$$

$$z - b = A_1(x_1 - a_1) + A_2(x_2 - a_2),$$

gdzie $A_1 = \frac{\partial f}{\partial x_1}(a_1, a_2)$, $A_2 = \frac{\partial f}{\partial x_2}(a_1, a_2)$ $b = f(a_1, a_2)$ jest styczna do powierzchni o

równaniu $z = f(x_1, x_2)$ w punkcie (a_1, a_2, b) .

Przykłady

1. Napisać równanie prostej stycznej do krzywej $\vec{r}(t) = \begin{cases} x(t) = 2 \cos t \\ y(t) = 2 \sin t, t \in \mathbb{R} \text{ w punkcie } A(0, 2, 0). \\ z(t) = \frac{t^2}{\pi} - \frac{\pi}{4} \end{cases}$

Rozwiązanie. Punktowi $A(0, 2, 0)$ odpowiada parametr $t_0 = \frac{\pi}{2}$. Ponadto $\vec{r}'(\frac{\pi}{2}) = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$. Wobec

tego poszukiwane równanie stycznej jest następujące $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} t$.

2. Napisać równanie płaszczyzny stycznej do powierzchni $z = x^2 + y^2$ w punkcie $A(1, 2, 5)$.

Rozwiązanie. Ogólnie $z - z(x_0, y_0) = \frac{\partial z}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial z}{\partial y}(x_0, y_0)(y - y_0)$. W rozważanym przypadku $z - 5 = 2(x - 1) + 4(y - 2)$.

Def. Mówimy, że funkcja f jest klasy C_D^1 jeżeli posiada w każdym punkcie zbioru D wszystkie pochodne cząstkowe ciągłe, czyli jest F - różniczkowalna w każdym punkcie zbioru E .

Pochodne cząstkowe wyższych rzędów.

Rozważmy funkcję rzeczywistą n zmiennych $f : R^n \supset D \rightarrow R$ określoną na zbiorze otwartym D .

Jeżeli w każdym punkcie zbioru D funkcja f ma pochodną cząstkową $\frac{\partial f}{\partial x_i}$, to mamy określoną

funkcją rzeczywistą $\frac{\partial f}{\partial x_i} : D \ni (x_1, \dots, x_n) \rightarrow \frac{\partial f}{\partial x_i}(x_1, \dots, x_n)$, dla której można w dowolnym ustalonym punkcie (x_1, \dots, x_n) zdefiniować (o ile istnieje) pochodną cząstkową po zmiennej x_j . W ten sposób określamy drugą pochodną cząstkową funkcji

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \left[\frac{\partial f}{\partial x_i} \right]$$

Tw. Schwarz. Jeżeli $f(x_1, \dots, x_n)$ ma w pewnym obszarze D ciągłe pochodne cząstkowe $\frac{\partial^2 f}{\partial x_j \partial x_i}$ i $\frac{\partial^2 f}{\partial x_i \partial x_j}$, to są one sobie równe w D .

Przykład Funkcja $f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$ jest ciągła w punkcie $(0, 0)$ i jej pochodne

cząstkowe są równe $\frac{\partial f}{\partial x}(0, y) = -y$; $\frac{\partial f}{\partial y}(x, 0) = x$. Stąd $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$ a $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$, gdyż $\frac{\partial^2 f}{\partial x_j \partial x_i}$ i $\frac{\partial^2 f}{\partial x_i \partial x_j}$ nie są ciągłe w punkcie

Pochodne cząstkowe funkcji złożonych.

Twierdzenie Jeżeli

1. $f(x_1, \dots, x_n)$ ma w pewnym obszarze $D \subset R^n$ ciągłe pochodne cząstkowe
 2. funkcje $x_i = x_i(u_1, \dots, u_m)$ $i=1, \dots, n$ mają ciągłe pochodne cząstkowe w pewnym obszarze $\Delta \subset R^m$
 3. $(x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m)) \in D$, gdy $(u_1, \dots, u_m) \in \Delta$
- to funkcja złożona $F(u_1, \dots, u_m) = f(x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m))$ ma ciągłe pochodne cząstkowe w każdym punkcie obszaru Δ i

$$\frac{\partial F}{\partial u_j}(u_1, \dots, u_m) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m)) \frac{\partial x_i}{\partial u_j}(u_1, \dots, u_m)$$

Dowód. dla przypadku szczególnego $n=2, m=1$ $F(t) = f(x(t), y(t))$

$$\begin{aligned} F'(t) &= \lim_{h \rightarrow 0} \frac{F(t+h) - F(t)}{h} = \lim_{h \rightarrow 0} \frac{f(x(t+h), y(t+h)) - f(x(t), y(t))}{h} = \\ &= \lim_{h \rightarrow 0} \frac{f(x(t+h), y(t+h)) - f(x(t), y(t+h)) + f(x(t), y(t+h)) - f(x(t), y(t))}{h} = \{\text{tw. Lagrange'a}\} = \\ &= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x}(c_1, y(t+h)) (x(t+h) - x(t))}{h} + \frac{\frac{\partial f}{\partial y}(x(t), c_2) (y(t+h) - y(t))}{h} = \{\text{ciągłość poch. cząstkowych}\} = \end{aligned}$$

$$= \frac{\partial f(x(t), y(t))}{\partial x} x'(t) + \frac{\partial f(x(t), y(t))}{\partial y} y'(t)$$

Przykłady

1. $f(x,y)=e^x \cos y$, $x(u,v)=3u+5v$, $y(u,v)=2u-v$, $F(u,v)=e^{3u+5v} \cos(2u-v)$. Obliczyć pochodne cząstkowe $\frac{\partial^2 F}{\partial u^2}$ i $\frac{\partial^2 F}{\partial v^2}$.
2. $F(u_1, u_2) = f(r(u_1, u_2))$ gdzie $r(u_1, u_2) = \sqrt{u_1^2 + u_2^2}$. Wykazać że $\frac{\partial^2 F}{\partial u_1^2} + \frac{\partial^2 F}{\partial u_2^2} = \frac{d^2 f}{dr^2} + \frac{1}{r} \frac{df}{dr}$

Różniczki

Oznaczenia $\mathbf{x}^0 = (x_1^0, \dots, x_n^0)$ $\Delta \mathbf{x} = (\Delta x_1, \dots, \Delta x_n)$ $f: Ot(\mathbf{x}^0, \delta) \rightarrow R$

Różniczką funkcji f w punkcie \mathbf{x}^0 dla przyrostu $\Delta \mathbf{x}$ nazywamy wyrażenie

$$df(\mathbf{x}^0, \Delta \mathbf{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}^0) \Delta x_i$$

Różniczki wyższych rzędów

Jeżeli przy ustalonym $\Delta \mathbf{x}$ funkcja $d(\cdot, \Delta \mathbf{x}): Ot(\mathbf{x}^0, \delta) \rightarrow R$ ma różniczkę, to nazywamy ją drugą różniczką funkcji f w punkcie \mathbf{x}^0 i oznaczamy $d^2 f(\mathbf{x}^0, \Delta \mathbf{x})$.

Przykład. Wyprowadzić wzór na drugą różniczkę funkcji dwóch zmiennych.

$$\begin{aligned} d((x_1, x_2), (\Delta x_1, \Delta x_2)) &= \frac{\partial f}{\partial x_1}(x_1, x_2) \Delta x_1 + \frac{\partial f}{\partial x_2}(x_1, x_2) \Delta x_2 \\ d^2((x_1, x_2), (\Delta x_1, \Delta x_2)) &= \\ &= \frac{\partial}{\partial x_1} \left[\frac{\partial f}{\partial x_1}(x_1, x_2) \Delta x_1 + \frac{\partial f}{\partial x_2}(x_1, x_2) \Delta x_2 \right] \Delta x_1 + \frac{\partial}{\partial x_2} \left[\frac{\partial f}{\partial x_1}(x_1, x_2) \Delta x_1 + \frac{\partial f}{\partial x_2}(x_1, x_2) \Delta x_2 \right] \Delta x_2 = \\ &= \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) \Delta x_1^2 + 2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) \Delta x_1 \Delta x_2 + \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) \Delta x_2^2 = \left(\frac{\partial}{\partial x_1} \Delta x_1 + \frac{\partial}{\partial x_2} \Delta x_2 \right)^2 f \end{aligned}$$

Druga różniczka funkcji wielu zmiennych w danym punkcie jest formą kwadratową przyrostów

$$d^2 f((x_1, \dots, x_n), (\Delta x_1, \dots, \Delta x_n)) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) \Delta x_i \Delta x_j$$

Uwagi o zapisie macierzowym drugiej różniczki.

$$d^2((x_1, x_2), (\Delta x_1, \Delta x_2)) = \begin{bmatrix} \Delta x_1 & \Delta x_2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) & \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$

Wzór na m -tą różniczkę funkcji n zmiennych

$$d^m f = \left(\frac{\partial}{\partial x_1} \Delta x_1 + \dots + \frac{\partial}{\partial x_n} \Delta x_n \right)^m f$$

Zadanie. Napisać wzór na drugą różniczkę funkcji trzech zmiennych i trzecią różniczkę funkcji dwóch zmiennych

Wzór Taylora

Jeżeli funkcja $f: R^n \supset Ot(\mathbf{x}^0, \delta) \ni \mathbf{x} \rightarrow f(\mathbf{x}) \in R$ ma ciągle pochodne cząstkowe do rzędu m w $Ot(\mathbf{x}^0, \delta)$, to istnieje $\theta \in (0,1)$ takie, że dla każdego $\mathbf{x} \in Ot(\mathbf{x}^0, \delta)$ prawdziwy jest wzór

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \frac{1}{1!} df(\mathbf{x}^0, \mathbf{x} - \mathbf{x}^0) + \frac{1}{2!} d^2 f(\mathbf{x}^0, \mathbf{x} - \mathbf{x}^0) + \dots + \frac{1}{(m-1)!} d^{m-1} f(\mathbf{x}^0, \mathbf{x} - \mathbf{x}^0) + r_m$$

$$\text{gdzie } r_m = \frac{1}{m!} d^m f(\mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0), \mathbf{x} - \mathbf{x}^0)$$

Uwaga: punkt „pośredni” $\mathbf{c} = \mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0)$ leży wewnątrz odcinka o końcach \mathbf{x}^0, \mathbf{x}

Dow. Parametryzujemy odcinek $\mathbf{x}(t) = \mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0)$; $t \in [0,1]$ i tworzymy funkcję jednej zmiennej $\varphi(t) = f(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0))$. Mamy więc rzeczywistą funkcję φ określoną na domkniętym odcinku $[0,1]$ spełniającą założenia twierdzenia Taylora dla funkcji jednej zmiennej. Stosując do niej wzór

$$\text{Maclaurina otrzymujemy } \varphi(1) = \varphi(0) + \frac{\varphi'(0)}{1!}(1-0) + \dots + \frac{\varphi^{(m-1)}(0)}{(m-1)!} 1^{m-1} + \frac{1}{m!} \varphi^{(m)}(\theta) 1^m \Leftrightarrow$$

$$\Leftrightarrow \varphi(1) = \varphi(0) + \frac{1}{1!} d\varphi(0,1) + \dots + \frac{1}{(m-1)!} d^{m-1} \varphi(0,1) + \frac{1}{m!} d^m \varphi(\theta,1)$$

$$\text{Ale } d^k \varphi(t_0, t) = d^k f(\mathbf{x}(t_0), \mathbf{x}(t) - \mathbf{x}(t_0)). \text{ Stąd } d^k \varphi(0,1) = d^k f(\mathbf{x}^0, \mathbf{x} - \mathbf{x}^0) \text{ itd.}$$

Przykład. Napisać wzór Taylora dla funkcji $f(x,y) = e^x \sin y$ w punkcie $(0,0)$ (wzór Maclaurina) dla $m=4$.

$$\text{Odp. } e^x \sin y = y + xy + \frac{1}{2} x^2 y - \frac{1}{6} y^3 + r_4$$

Zastosowanie różniczki w teorii błędów

Dana jest funkcja f wielu zmiennych. Wektorowy argument funkcji nie jest znany lecz dysponujemy jego pomiarem \mathbf{x} obarczonym błędem. Niech $\mathbf{x} + \Delta \mathbf{x}$ oznacza prawdziwą nieznaną wartość argumentu a błąd bezwzględny pomiaru (wektorowego) nie przekracza \mathbf{b} ($|\Delta \mathbf{x}| \leq \mathbf{b}$ nierówność wektorowa).

$$\text{Wówczas } |f(\mathbf{x} + \Delta \mathbf{x}) - f(\mathbf{x})| \approx |df(\mathbf{x}, \Delta \mathbf{x})| \leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(\mathbf{x}) \right| b_i.$$

Wy tłumaczenie przybliżonego charakteru wzoru, kiedy to przybliżenie jest „dobre” i jak postąpić gdy przybliżenia nie jest zadowalające.

Przykład. Oszacować metodą różniczki zupełnej błąd jaki popełniamy obliczając objętość prostopadłościanu o krawędziach 4.1, 3.2, 8.4 zmierzonych odpowiednio z dokładnością 0.1, 0.1, 0.2. Odp. (błąd bezwzględny ≤ 8.756 , błąd względny 8% (wy tłumaczyć jak rozumiemy błąd względny)

Ekstrema lokalne

Def. Funkcja $f: R^n \supset Ot(\mathbf{x}^0, \delta) \ni \mathbf{x} \rightarrow f(\mathbf{x}) \in R$ ma w punkcie \mathbf{x}^0 maksimum lokalne właściwe jeżeli
 $\forall \mathbf{x} \in S(\mathbf{x}^0, \delta) \quad f(\mathbf{x}) < f(\mathbf{x}^0)$ (Uwaga $S(\mathbf{x}^0, \delta) = Ot(\mathbf{x}^0, \delta) - \{\mathbf{x}^0\}$)

Podobnie minimum lokalne właściwe.

Warunek konieczny istnienia ekstremum

Jeżeli funkcja $f: R^n \supset Ot(\mathbf{x}^0, \delta) \ni \mathbf{x} \rightarrow f(\mathbf{x}) \in R$

- ma w punkcie \mathbf{x}^0 pochodne cząstkowe
- ma w punkcie \mathbf{x}^0 ekstremum

to $\frac{\partial f}{\partial x_i}(\mathbf{x}^0) = 0, i=1, \dots, n$.

Dowód. Dla prostoty przyjmijmy $n=2$. Jeżeli funkcja $f(x, y)$ ma w punkcie (x_0, y_0) ekstremum i pochodne cząstkowe, to różniczkowalna funkcja jednej zmiennej $f(x, y_0)$ ma w punkcie x_0 ekstremum.

Stąd $\frac{\partial f}{\partial x}(x_0, y_0) = 0$. Podobnie $\frac{\partial f}{\partial y}(x_0, y_0) = 0$.

Uwaga: punktami krytycznymi są także punkty w których pochodne cząstkowe nie istnieją.

Warunek wystarczający istnienia ekstremum

Jeżeli $f: R^n \supset Ot(\mathbf{x}^0, \delta) \ni \mathbf{x} \rightarrow f(\mathbf{x}) \in R$

- jest klasy $C^2(Ot(\mathbf{x}^0, \delta))$ (tzn. ma drugie pochodne cząstkowe ciągłe w $Ot(\mathbf{x}^0, \delta)$)
- $\frac{\partial f}{\partial x_i}(\mathbf{x}^0) = 0, i=1, \dots, n$
- $d^2f(\mathbf{x}^0, \Delta \mathbf{x}) > 0 (<), \forall \Delta \mathbf{x} \neq \mathbf{0}$

to f ma w punkcie \mathbf{x}^0 minimum (maksimum) lokalne właściwe

Dowód. Z wzoru Taylora dokładnie jak dla funkcji jednej zmiennej mamy

$f(\mathbf{x}) - f(\mathbf{x}^0) = \frac{1}{2!} d^2 f(\mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0), \mathbf{x} - \mathbf{x}^0) > 0 \quad \forall \Delta \mathbf{x} \neq \mathbf{0}$, gdy $d^2 f(\mathbf{x}^0, \Delta \mathbf{x}) > 0$, gdyż z ciągłości pochodnych rzędu drugiego, dla ustalonych przyrostów druga różniczka jest funkcją ciągłą w \mathbf{x}^0 więc z twierdzenia o lokalnym zachowaniu znaku $d^2 f(\mathbf{x}^0, \mathbf{x} - \mathbf{x}^0)$ i $d^2 f(\mathbf{x}^0 + \theta(\mathbf{x} - \mathbf{x}^0), \mathbf{x} - \mathbf{x}^0)$ mają ten sam znak dla $\forall \Delta \mathbf{x} \neq \mathbf{0}$

Problem. Jak badać warunek $d^2 f(\mathbf{x}^0, \Delta \mathbf{x}) > 0 (<), \forall \Delta \mathbf{x} \neq \mathbf{0}$. Dla $n=2$

$$\begin{aligned} d^2((x_1, x_2), (\Delta x_1, \Delta x_2)) &= \begin{bmatrix} \Delta x_1 & \Delta x_2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) & \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \\ &= \begin{bmatrix} \Delta x_1 & \Delta x_2 \end{bmatrix} \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = A\Delta x_1^2 + 2B\Delta x_1\Delta x_2 + C\Delta x_2^2 = \{\text{gdy } A \neq 0\} \\ &= A\{\Delta x_1^2 + 2\frac{B}{A}\Delta x_1\Delta x_2 + \frac{B^2}{A^2}\Delta x_2^2\} - \frac{B^2}{A}\Delta x_2^2 + C\Delta x_2^2 = A\left((\Delta x_1 + \frac{B}{A}\Delta x_2)^2 + \frac{AC - B^2}{A^2}\Delta x_2^2\right) \end{aligned}$$

Widać, że jeśli

- $D_1 = A > 0$ i $D_2 = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} > 0$ to druga różniczka jest dodatnia
- $D_1 = A < 0$ i $D_2 = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} > 0$ to druga różniczka jest ujemna

Uzupełnienia z algebry (nieobowiązujące)

- Formą kwadratową n zmiennych rzeczywistych nazywamy funkcję

$$\varphi: R^n \ni \mathbf{x} \rightarrow \varphi(\mathbf{x}) = \varphi(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j \quad \text{przy czym } a_{ij} = a_{ji}$$

- Forma kwadratowa ma przy ustalonej bazie przestrzeni $X=R^n$ ma dokładnie jedną reprezentację macierzową $\varphi(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle$ gdzie \mathbf{A} jest macierzą symetryczną.

Formę kwadratową φ nazywamy

- nieujemną - gdy $\forall \mathbf{x} \in R^n \quad \varphi(\mathbf{x}) \geq 0$
- niedodatnią - gdy $\forall \mathbf{x} \in R^n \quad \varphi(\mathbf{x}) \leq 0$
- dodatnią - gdy $\forall \mathbf{x} \in R^n, \mathbf{x} \neq \mathbf{0} : \varphi(\mathbf{x}) > 0$
- ujemną - gdy $\forall \mathbf{x} \in R^n, \mathbf{x} \neq \mathbf{0} : \varphi(\mathbf{x}) < 0$
- dodatnio określoną - gdy $\exists c > 0 \forall \mathbf{x} \in R^n : \varphi(\mathbf{x}) \geq c \|\mathbf{x}\|^2$
- ujemnie określoną - gdy $\exists c > 0 \forall \mathbf{x} \in R^n : \varphi(\mathbf{x}) \leq -c \|\mathbf{x}\|^2$
- nieokreśloną - gdy $\exists \mathbf{x} \in R^n \varphi(\mathbf{x}) > 0$ i $\exists \mathbf{y} \in R^n \varphi(\mathbf{y}) < 0$

Uwaga. W przypadku definiowania form kwadratowych na nieskończenie wymiarowej przestrzeni liniowej unormowanej należy odróżniać pojęcia.

dodatniość formy	i	dodatnia określoność formy
ujemność formy	i	ujemna określoność formy

W przypadku przestrzeni skończenie wymiarowych pojęcia te pokrywają się. Jest to konsekwencją zwartości sfery jednostkowej w przestrzeni skończenie wymiarowej.

Def. Minorami kątowymi (odpowiednich stopni) macierzy
$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$
 nazywamy

następujące wyznaczniki $D_1 = |a_{11}|$, $D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, ..., $D_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$.

Twierdzenie. (kryterium Sylwestera).

Forma kwadratowa $\varphi(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ jest dodatnia (dodatnio określona) $\Leftrightarrow D_i > 0 \quad \forall i$

Forma kwadratowa $\varphi(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ jest ujemna (ujemnie określona) $\Leftrightarrow (-1)^i D_i > 0 \quad \forall i$

Jeżeli wszystkie minory kątowe D_i są różne od 0 i zmieniają się inaczej niż w powyższych przypadkach, to forma jest nieokreślona

Przykłady Wyznaczyć ekstrema lokalne funkcji

a) $f(x,y) = x^4 + y^4 - 2x^2 + 4xy - 2y^2$

(punkty krytyczne $(0,0)$, $(\sqrt{2}, -\sqrt{2})$, $(-\sqrt{2}, \sqrt{2})$. WW nie rozstrzyga w $(0,0)$ ($f(x,x) - f(0,0) = 2x^4 > 0$, $f(x,0) - f(0,0) = x^2(2-x^2) < 0$ dla dostatecznie małych $x \neq 0 \Rightarrow$ brak ekstremum w $(0,0)$ w pozostałych punktach minimum lokalne (-8))

b) $f(x,y,z) = x^2 - 2x - y^3 + 3y + 5z^2$ (nieobowiązkowe)

Ekstrema globalne

Z twierdzenia Weierstrassa funkcja ciągła (wielu zmiennych) na zbiorze domkniętym i ograniczonym (a więc zwartym) D osiąga swoje kresy.

Algorytm

- szukamy punktów krytycznych we wnętrzu $\text{int } D$ ($\text{int } D$ int =interior= wnętrze)
- szukamy największej i najmniejszej wartości funkcji na brzegu D (to samo zagadnienie ale niższy wymiar)
- obliczamy wartości funkcji w wyznaczonych powyżej punktach i wybieramy (ze skończonej listy) wartość najmniejszą i największą

Przykład. Znaleźć najmniejszą i największą) wartość funkcji $f(x,y) = x^2 + y^2 - xy + x + y$ w trójkącie domkniętym ograniczonym przez proste $x=0$, $y=0$, $x+y+3=0$.

Odp. $\text{Max} = f(-3,0) = f(0,-3) = 6$ $\text{Min} = f(-1,-1) = -1$

Funkcje uwikłane

Przykład 1. Rozważmy równanie $f(x, y) = 0$ np. $x^2 + y^2 - 1 = 0$ i niech $f(x_0, y_0) = 0$.

Pytanie: Czy dla dowolnego $x \in [-1, 1]$ istnieje taki $y(x)$, że $f(x, y(x)) = 0$ i $y(x_0) = y_0$?

Niech $(x_0, y_0) = (0, 1)$. Wówczas funkcja $y(x) = \sqrt{1 - x^2}$; $-1 \leq x \leq 1$ spełnia powyższe warunki. Ale

spełnia je także funkcja $y(x) = \begin{cases} \sqrt{1 - x^2} & x \in [-1, 1] \cap Q \\ -\sqrt{1 - x^2} & x \in [-1, 1] \cap (R - Q) \end{cases}$. Dokładając warunek ciągłości funkcji

$y(x)$ eliminujemy ten przypadek.

Jeżeli $(x_0, y_0) = (1, 0)$, to nie istnieje funkcja $y(x)$ będąca rozwiązaniem problemu, ale istnieje funkcja $x(y)$ spełniająca warunki zadania.

Powód. W punkcie $(1, 0)$ nie istnieje styczna $Ax + By + C = 0$ dająca się rozwinąć ze względu na y ale daje się rozwinąć ze względu na x .

Intuicja. Funkcja różniczkowalna w punkcie (x_0, y_0) zachowuje się w otoczeniu tego punktu podobnie do swej liniowej aproksymacji.

Twierdzenie o funkcji uwikłanej . (Prosty dowód Leja F. Rachunek różniczkowy i całkowy) . Jeżeli

- funkcja $f(x, y)$ ma ciągle pochodne cząstkowe $\frac{\partial f}{\partial x}$ i $\frac{\partial f}{\partial y}$ w otoczeniu punktu (x_0, y_0)
- $f(x_0, y_0) = 0$ i $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$

to

1. dla każdej dostatecznie małej liczby $\varepsilon > 0$ istnieje taka liczba $\delta > 0$, że każdej wartości x z przedziału $(x_0 - \delta, x_0 + \delta)$ odpowiada dokładnie jedno rozwiązanie $y(x)$ równania $f(x, y) = 0$ należące do przedziału $(y_0 - \varepsilon, y_0 + \varepsilon)$
2. funkcja $y(x)$ jest ciągła w przedziale $(x_0 - \delta, x_0 + \delta)$ i ma w nim ciągłą pochodną wyrażoną wzorem

$$y'(x) = -\frac{\frac{\partial f}{\partial x}(x, y)}{\frac{\partial f}{\partial y}(x, y)}, \text{ gdzie } y = y(x)$$

Przykład.. Wykorzystując wzór Taylora znajdź przybliżenie funkcji uwikłanej $y = y(x)$ wielomianem stopnia trzeciego w otoczeniu punktu $(x_0, y_0) = (1, 0)$. Funkcja uwikłana zadana jest równaniem $\cos(xy) - x - 2y = 0$.

W otoczeniu punktu $(x_0, y_0) = (1, 0)$ są spełnione założenia tw. o funkcji uwikłanej więc równanie $\cos(xy) - x - 2y = 0$ określa w tym otoczeniu funkcję $y = y(x)$, przy czym $y(1) = 0$.

$$y(x) = y(1) + \frac{y'(1)}{1!}(x-1) + \frac{y''(1)}{2!}(x-1)^2 + \frac{y'''(1)}{3!}(x-1)^3 + r_4$$

Kolejne pochodne funkcji $y(x)$ w punkcie $x_0 = 1$ wyliczamy różniczkując równanie 1

$$(1) \quad \cos[xy(x)] - x - 2y[x] = 0,$$

i otrzymujemy

$$(2) \quad -\sin[xy(x)][y(x) + xy'(x)] - 1 - 2y'[x] = 0,$$

a stąd dla $x = 1$ uwzględniając że $y(1) = 0$ otrzymujemy $y'(1) = -\frac{1}{2}$.

Różniczkując równanie 2 otrzymujemy

$$(3) \quad -\cos[xy(x)][y(x) + xy'(x)]^2 - \sin[xy(x)][2y'(x) + xy''(x)] - 2y''(x) = 0$$

a stąd dla $x = 1$ uwzględniając że $y(1) = 0$ i $y'(1) = -\frac{1}{2}$ otrzymujemy $y''(1) = -\frac{1}{8}$.

Różniczkując równanie 3 otrzymujemy

$$(4) \quad \sin[xy(x)][y(x) + xy'(x)]^3 - 3\cos[xy(x)][y(x) + xy'(x)][2y'(x) + xy''(x)] - \sin[xy(x)][3y''(x) + xy'''(x)] - 2y'''(x) = 0,$$

a stąd dla $x = 1$ uwzględniając że $y(1) = 0$ i $y'(1) = -\frac{1}{2}$ $y''(1) = -\frac{1}{8}$ otrzymujemy $y'''(1) = -\frac{27}{32}$

Ostatecznie $y(x) = \frac{-\frac{1}{2}}{1!}(x-1) + \frac{-\frac{1}{8}}{2!}(x-1)^2 + \frac{-\frac{27}{32}}{3!}(x-1)^3 + r_4$

Ekstrema funkcji uwikłanych

- metoda rozwikłania ograniczeń (wyjaśnić)

Przykład (wprowadzający) Zbadać ekstrema funkcji uwikłanej $y = y(x)$ określonej równaniem $f(x, y) = 0$.

Zakładamy regularność funkcji f tak, aby wyliczone poniżej pochodne miały sens, czyli, że są spełnione założenia twierdzenia o funkcji uwikłanej.

$$f(x, y(x)) = 0 \quad \Bigg/ \frac{d}{dx}$$

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y'(x) = 0 \Rightarrow y'(x) = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \quad \text{o ile} \quad \frac{\partial f}{\partial y} \neq 0$$

$y(x)$ ma ekstremum w punkcie $x \Leftrightarrow$ (z WK) $y'(x) = 0 \Leftrightarrow \frac{\partial f}{\partial x} = 0$

Otrzymaliśmy więc WK istnienia ekstremum funkcji uwikłanej

$$\text{WK: } \begin{cases} \frac{\partial f}{\partial x} = 0 \\ f(x, y) = 0 \\ \frac{\partial f}{\partial y} \neq 0 \end{cases} \quad \begin{array}{l} \text{Rozwiązujemy pierwszy układ i dostajemy punkty krytyczne } P_i, \\ \text{dla których sprawdzamy ostatni warunek.} \end{array}$$

Badanie rodzaju ekstremum może przebiegać za pomocą badania znaku $y'(x)$ w otoczeniu punktu krytycznego P_i lub badania znaku $y''(x)$. Pierwszy sposób jest nieco kłopotliwy. Nawet badanie znaku formy kwadratowej wymagało specjalnego narzędzia – kryterium Sylwestera (są też inne).

Różniczkując ponownie otrzymamy

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y'(x) = 0 \quad \Bigg/ \frac{d}{dx}$$

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y \partial x} y'(x) + \left(\frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} y'(x) \right) y'(x) + \frac{\partial f}{\partial y} y''(x) = 0$$

$$\frac{\partial^2 f}{\partial x^2} + 2 \frac{\partial^2 f}{\partial y \partial x} y'(x) + \frac{\partial^2 f}{\partial y^2} (y'(x))^2 + \frac{\partial f}{\partial y} y''(x) = 0$$

w punktach krytycznych $y'(x) = 0$ więc

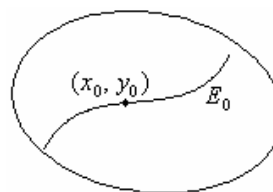
$$y''(x) = - \frac{\frac{\partial^2 f}{\partial x^2}}{\frac{\partial f}{\partial y}} - \text{tylko w punkcie krytycznym}$$

$y''(P_i) > 0 \Rightarrow y(x)$ ma w punkcie x_i minimum lokalne właściwe $y(x_i) = y_i$

Ekstrema warunkowe

Np.: $f: R^2 \supset E \rightarrow R$ E – otwarty
 $g: R^2 \supset E \rightarrow R$ ciągła

Oznaczmy $E_0 = \{(x, y) \in E : g(x, y) = 0\}$
 i niech $(x_0, y_0) \in E_0$, czyli E_0 – niepusty.



Rozpatrzmy funkcję f obciętą do E_0 :

Ekstremum funkcji f obciętej do E_0 nazywać będziemy ekstremum warunkowym funkcji f pod warunkiem $g(x, y) = 0$

Def. Funkcja f ma w punkcie (x_0, y_0) maksimum lokalne warunkowe właściwe

$$\Leftrightarrow \exists_{S(x_0, y_0)} \forall_{(x, y) \in S(x_0, y_0) \cap E} f(x, y) < f(x_0, y_0)$$

Jak znaleźć ekstremum warunkowe? (wprowadzenie do metody Lagrange'a)

Zakładając, że równanie $g(x, y) = 0$ określa funkcje uwikłaną $y(x)$ problem sprowadza się do szukania ekstremum funkcji jednej zmiennej $\varphi(x) = f(x, y(x))$. Zakładając regularność

z WK istnienia ekstremum otrzymujemy $\varphi'(x) = 0 \Leftrightarrow \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y'(x) = 0$,

a z warunku $g(x, y(x)) = 0 \quad \Bigg/ \frac{d}{dx} \Leftrightarrow \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} y'(x) = 0$ obliczamy $y'(x)$. Stąd otrzymujemy

$$\text{WK: } \begin{cases} \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} = 0 \Leftrightarrow \varphi'(x) = 0 \\ g(x, y(x)) = 0 \end{cases}$$

Można zauważyć, że równanie pierwsze WK jest wynikiem rugowania parametru λ z następującego

$$\text{układu równań: } \begin{cases} \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 \end{cases}, \text{ gdzie}$$

Lewe strony są pochodnymi cząstkowymi funkcji $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$

$$\text{WK jest więc } \Leftrightarrow \begin{cases} \frac{\partial L}{\partial x} = 0 \\ \frac{\partial L}{\partial y} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases}$$

Badanie sprowadza się do badania funkcji Lagrange'a $L(x, y, \lambda)$ z mnożnikiem Lagrange'a λ . Metodę tę można uogólnić dla funkcji wielu zmiennych.

Metoda mnożników Lagrange'a szukanie ekstremum funkcji $f(x, y)$ pod warunkiem $g(x, y) = 0$

$$f: R^2 \supset E \rightarrow R$$

$$g: R^2 \supset E \rightarrow R$$

$$E_0 = \{(x, y) \in E : g(x, y) = 0\}$$

Algorytm: Tworzymy funkcję Lagrange'a

$$L(x, y; \lambda) = f(x, y) + \lambda g(x, y)$$

$$\text{Warunek konieczny: } \begin{cases} \frac{\partial L}{\partial x} = 0 \\ \frac{\partial L}{\partial y} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases} \rightarrow P(x_i, y_i, \lambda_i)$$

Warunek wystarczający. Traktując mnożnik λ jako parametr, wyznaczyć w punktach krytycznych drugą różniczkę $d^2 L((x_i, y_i; \lambda_i), (\Delta x, \Delta y))$ przy czym przyrosty $(\Delta x, \Delta y)$ spełniają równanie

$$\frac{\partial g}{\partial x}(x_i, y_i; \lambda_i) \Delta x + \frac{\partial g}{\partial y}(x_i, y_i; \lambda_i) \Delta y = 0$$

Wyznaczając jeden z przyrostów (np. Δy jako funkcję drugiego przyrostu Δx) badamy określoność $d^2 L((x, \lambda), \Delta x)$.

Przykład. Wyznaczyć ekstrema warunkowe funkcji $f(x, y) = xy$ przy warunku $g(x, y) = x + y - 1 = 0$

Pokazać obie metody. W metodzie Lagrange'a zwrócić szczególną uwagę na konieczność krępowania przyrostów w WW.

- metoda rozwikłania ograniczeń : $g(x, y) = x + y - 1 = 0 \Leftrightarrow y = 1 - x$, $\varphi(x) = f(x, 1 - x) = x(1 - x)$

Funkcja φ ma w punkcie $x_0 = \frac{1}{2}$ maksimum lokalne równe $\frac{1}{4}$, więc funkcja f ma w punkcie $(x_0, y_0) = (\frac{1}{2}, \frac{1}{2})$ maksimum lokalne warunkowe.

- Metoda Lagrange'a $L(x, y; \lambda) = xy + \lambda(x + y - 1)$

$$\text{WK: } \begin{cases} \frac{\partial L}{\partial x} = 0 \\ \frac{\partial L}{\partial y} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases} \Leftrightarrow \begin{cases} y + \lambda = 0 \\ x + \lambda = 0 \\ x + y - 1 = 0 \end{cases} \Leftrightarrow P(\frac{1}{2}, \frac{1}{2}; -\frac{1}{2})$$

WW : $d^2L((x_1, y_1; \lambda_1), (\Delta x, \Delta y)) = 2\Delta x \Delta y$ przy czym $\Delta x + \Delta y = 0$. Wobec tego $d^2L((\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}), (\Delta x, -\Delta x)) = -2(\Delta x)^2 < 0$ dla $\Delta x \neq 0$ więc f ma w punkcie $(x_0, y_0) = (\frac{1}{2}, \frac{1}{2})$ maksimum lokalne warunkowe.

Przykładowe zadania z funkcji uwikłanych

1. Znaleźć ekstrema lokalne funkcji uwikłanej $y(x)$ określonej równaniem
 - a) $x^2 - 2xy + 2y^2 + 2x + 1 = 0$.
 - b) $x^2 e^y - y^4 + 1 = 0$.
2. W dostatecznie małym otoczeniu punktu $(1, y_0)$ narysować wykres funkcji uwikłanej $y(x)$ określonej równaniem $x^2 \ln y - y \ln x = 0$.