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 $f : \mathbb{R}^n \to \mathbb{R}$, jeśli istnieje otoczenie $\mathcal{O}$ punktu x_0 takie, że dla wszystkich $x \in \mathcal{O}$ zachodzi $f(x) \leqslant f(x_0)$ ($f(x) \geqslant f(x_0)$).

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. Mówią 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

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

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 $\varphi_i(t) \leqslant \varphi(0)$. 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:

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

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 $f : \mathbb{R}^n \to \mathbb{R}$ 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.

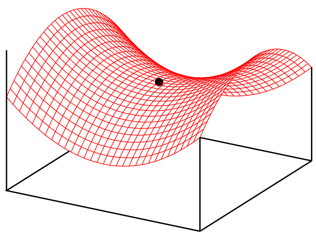
Przykład 1 (Siodło)**.** Najprostszy przykład to funkcja $f_1 : \mathbb{R}^2 \to \mathbb{R}$ zadana wzorem

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

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:



♣

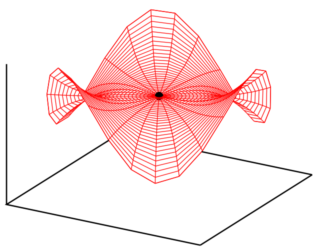
Przykład 2 (Małpie siodło)**.** Nieco ciekawszy kształt ma wykres funkcji

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

Jej pochodna

$$f_2'(0, 0) = [3x^2 - 3y^2 \;\; 6yx]$$

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

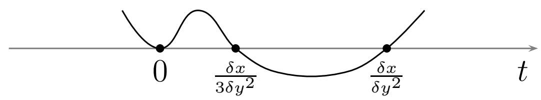
$$t \longmapsto \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$$

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

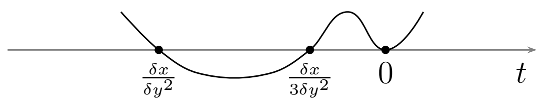
Załóżmy, że $\delta y \neq 0$ (co wyklucza na razie z rozważań prostą $y = 0$) wtedy zależność od t można opisać wzorem:

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

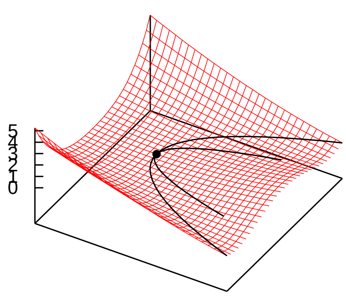
Widać więc, że wzdłuż każdej prostej funkcja ma w zerze minimum. Dla dodatnich δx sytuacja wygląda tak:

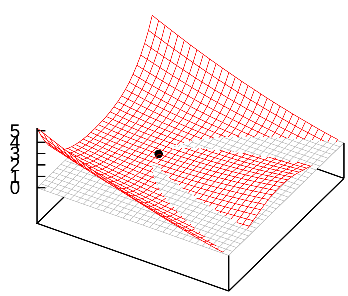


Dla ujemnych tak:



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 $F : \mathbb{R}^n \to \mathbb{R}^m$. 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

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

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:

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

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. *Załóżmy, że funkcja $f : \mathbb{R}^2 \to \mathbb{R}$ ma w otoczeniu $\mathcal{O}$ punktu (x_0, y_0) pochodne cząstkowe rzędu pierwszego i drugiego. Załóżmy 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 $\varphi : [a, b] \to \mathbb{R}$ 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:

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

Pochodną mieszaną

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

liczymy przechodząc do granicy

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

a pochodną mieszaną w drugiej kolejności licząc

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

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

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

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. □

Twierdzenie Taylora. Znajomość pochodnej funkcji $f:\mathbb{R}^n\to\mathbb{R}$ 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 $f:\mathbb{R}^n\to\mathbb{R}$ 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

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

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

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

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

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

i ich wartości w punkcie $(1,2)$:

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

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

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

Obliczamy drugie pochodne cząstkowe

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

i ich wartości w punkcie $(1,2)$:

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

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

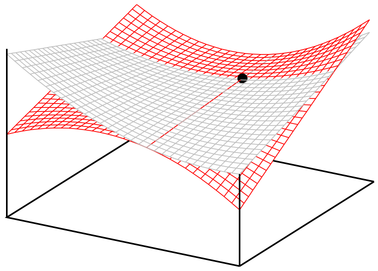
Funkcja f jest dosyć prosta, dlatego możemy *explicite* zapisać wzór na resztę:

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

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 eykres 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

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

Zaczynamy od wyznaczenia pierwszych pochodnych cząstkowych funkcji g:

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

$$\sin y[\cos x \sin(x + y) + \sin x \cos(x + y)] = \sin y \sin(2x + y).$$

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

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

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

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

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:

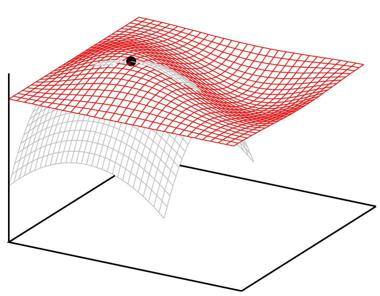
$$D_1 = -\sqrt{3}, \qquad D_2 = 3 - \frac{3}{4} = \frac{9}{4}.$$

Odpowiednie liczby wyznaczające sygnaturę to:

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

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:

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

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

Punkt krytyczny spełnia warunki

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

czyli

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

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

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

Z powyższych równań wyznaczamy a i b:

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

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

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

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

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

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

Macierz drugiej pochodnej ma postać

$$\left[\begin{array}{cc} 2\sum\limits_{i=1}^{n}(x_i)^2 & 2\sum\limits_{i=1}^{n}x_i \\ 2\sum\limits_{i=1}^{n}x_i & 2n \end{array}\right] = 2n\left[\begin{array}{cc} E(x^2) & E(x) \\ E(x) & 1 \end{array}\right].$$

Współczynnk $2n$ 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.♣